NeuroImage 233 (2021) 117916

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Neural substrates for sharing intention in action during face-to-face imitation

Kohei Miyata, Takahiko Koike, Eri Nakagawa, Tokiko Harada, Motofumi Sumiya, Tetsuya Yamamoto, Norihiro Sadato*

Division of Cerebral Integration, Department of System Neuroscience, National Institute for Physiological Sciences, 38 Nishigonaka, Myodaiji, Okazaki, Aichi 444-8585, Japan

ARTICLE INFO

Keywords:
Imitation
Hyperscanning
fMRI
Shared representation of action
Internal model
Predictive coding

ABSTRACT

Face-to-face imitation is a unique social interaction wherein a shared action is executed based on the feedback of the partner. Imitation by the partner is the feedback to the imitatee's action, resulting in sharing actions. The neural mechanisms of the shared representation of action during face-to-face imitation, the core of inter-subjectivity, are not well-known. Here, based on the predictive coding account, we hypothesized that the pair-specific forward internal model is the shared representation of action which is represented by the inter-individual synchronization of some portion of the mirror neuron system. Hyperscanning functional magnetic resonance imaging was conducted during face-to-face interaction in 16 pairs of participants who completed an immediate imitation task of facial expressions. Paired participants were alternately assigned to either an imitator or an imitatee who was prompted to express a happy, sad, or non-emotional face. While neural activation elicited by imitating and being imitated were distinct with little overlap, on-line imitative interaction enhanced inter-brain synchronization in the right inferior parietal lobule that correlated with the similarity in facial movement kinematic profile. This finding indicates a critical role of the right inferior parietal lobule in sharing representation of action as a pair-specific forward internal model through imitative interaction.

1. Introduction

Imitation refers to reproducing a consequence of actions by copying actions with understanding the action's goal (Call and Carpenter, 2002). Face-to-face imitation is a unique social interaction wherein a shared action is executed based on the feedback of the partner. This interaction is critical in social cognition and its development (Decety and Sommerville, 2003; Hobson, 2003; Nadel-Brulfert and Baudonniere, 1982; Trevarthen, 1979). It is the avenue for developing early communication by recognizing 'self-other equivalences' or common acts (Gopnik and Meltzoff, 1994). Role-taking during imitation between an infant and his/her mother is a milestone in linking their subjective experiences (Decety and Sommerville, 2003) and promotes the development of an implicit sense of self as a social agent (Rochat, 1999). Furthermore, imitation, along with empathy and mindreading, is critical in inter-subjectivity by providing the same internal representation of action shared between the self and the other (shared representation of action [SRA]) (de Vignemont and Haggard, 2008). During mutual interaction, perceptual and motoric representations are linked such that perceiving another person's action activates the same representations as performing the action. This common coding, SRA, allows a person to embody the behaviors of others and infer their internal states (Barsalou et al., 2003). Thus, SRA is the core component of inter-subjectivity, providing the "like-me" framework (Gallese, 2003).

What is shared by SRA is supposed to be intention in action (de Vignemont and Haggard, 2008). According to Searle (1983), there are two types of intentions: prior intention and intention in action. Prior intention represents the goal of the action as a global unit (e.g. I intend to

Abbreviations: ACC, anterior cingulate cortex; BOLD, blood-oxygen-level-dependent; *C*, control for imitation; CSF, cerebrospinal fluid; dmPFC, dorsomedial prefrontal cortex; EPI, echo planar imaging; FA, flip angle; FoV, field of view; *H*, happy; HRF, hemodynamic response function; *I*, the initiator; IFG, inferior frontal gyrus; IPL, inferior parietal lobule; MNI, Montreal Neurological Institute; MTG, middle temporal gyrus; *N*, emotionally neutral with an open mouth; PMv, ventral premotor area; Pre-SMA, pre-supplementary motor area; *R*, the responder; rACC, rostral anterior cingulate cortex; *S*, sad; SMA, supplementary motor area; SRA, shared representation of action; TA, acquisition time; TE, echo time; TPJ, temporoparietal junction; TR, repetition time.

* Corresponding author at: Division of Cerebral Integration, Department of System Neuroscience, National Institute for Physiological Sciences, 38 Nishigonaka, Myodaiji, Okazaki, Aichi 444-8585, Japan.

E-mail addresses: kmiyata@idaten.c.u-tokyo.ac.jp (K. Miyata), tkoike@nips.ac.jp (T. Koike), enkgw@nips.ac.jp (E. Nakagawa), tokikoh@nips.ac.jp (T. Harada), msumiya@nips.ac.jp (M. Sumiya), tyama@nips.ac.jp (T. Yamamoto), sadato@nips.ac.jp (N. Sadato).

https://doi.org/10.1016/j.neuroimage.2021.117916
Received 24 April 2020; Received in revised form 17 February 2021; Accepted 19 February 2021
Available online 16 March 2021

drink) (de Vignemont and Haggard, 2008). Intention in action represents the action as a dynamic sequence of specific movements, initiating the action, guiding it and monitoring its effects. Intention in action corresponds to the motor plan linking the set of movements into meaningful sequences, prior to the dispatch of the final motor command to the muscles (Jeannerod, 1995). Prior intention represents the goal which may have multiple means to realize it, thus less specific to the movement to be performed. Intentions in action has a level of motor specificity that prior intention lacks (de Vignemont and Haggard, 2008). While the mirror system provides the neural substrate for imitation (Caspers et al., 2010; Iacoboni et al., 1999; Rizzolatti and Craighero, 2004), the neural mechanisms of the SRA during face-to-face imitation are not well-known (de Vignemont and Haggard, 2008).

The potential mechanisms of SRA during mutual imitation is a resonance of internal model. Imitating is understood as a prediction/control relation or forward/inverse internal model (Sasaki et al., 2012; Wolpert et al., 1998; Wolpert and Kawato, 1998). The internal model is defined as a set of input-output relations between motor commands and their sensory consequences (Penhune and Steele, 2012). The imitators decode the visually presented movements of the imitatees into their motor representation and generate a series of motor commands. This decoding process is regarded as the inverse internal model defined as the inference of the causes of sensory input, approximating the reverse transformation of sensory inputs to the causes during perceptual inference (cf. recognition models) (Kilner et al., 2007). The scheme of inverse models will only work when the processes generating the sensory inputs from the causes are invertible, which is not the case in general because the same sensory input can have many causes (Kilner et al., 2007). To solve this ill-posed problem, a predictive coding framework was proposed. In this schema, neuronal representations in higher levels of cortical hierarchies predict the representations in lower levels (Friston, 2008; Mumford, 1992; Rao and Ballard, 1999). The comparison of top-down predictions (forward model) with representations at the lower level forms a prediction error fed back up the hierarchy to update higher representations. This recursive exchange of signals suppresses prediction error at every level to provide a hierarchical explanation for sensory inputs (Friston and Frith, 2015). According to the predictive coding account, both self action optimization and action inference of others require forward model or top-down prediction (Kilner et al., 2007): The same forward model used to predict the sensorial effects of one's own actions can also be used as a constraint for decoding the actions of others (Friston, 2005; Kilner et al., 2007). That is, the inverse model of the imitator is the inversion of the forward (or generative) model of the imitatee with the imitator's forward model as a constraint. The imitatees also predict and monitor the partner's responses as consequences of their own actions to update their own forward model. As the imitator's motor output reflects the imitator's forward model, these iterative updating processes result in the inter-individual neuronal synchronization during reciprocal interaction of the agents, representing the shared forward internal models between the interacting agents (Friston and Frith, 2015). Based on this theoretical prediction, we expected the inter-individual synchronization of some portion of mirror neuron system of the two brains during mutual imitation, which represents the pair-specific forward internal models.

The predictive coding account also suggests that the imitatees predict and monitor the partner's responses as consequences of their own actions to detect contingency and similarity which elicits positive feeling (e.g., affiliation and trust) towards the imitator and prosocial behavior (Hale and Hamilton, 2016; Kühn et al., 2010). Thus, we expected the activation of the reward system of the imitatee.

To depict the neural substrates of imitating, being imitated, and SRA during face-to-face interaction, we conducted hyperscanning fMRI with an immediate imitation of facial expressions as a task. The roles of imitating and being imitated were alternated to enhance the shared representation. To eliminate the motor components common to both imitating and being imitated, we included a control condition that involved simultaneous facial movements. We first hypothesized that neural substrates of imitating and being imitated are distinct but linked by the SRA which is specified by pair-specific neural and behavioral synchronization. Second, we also hypothesized that being imitated elicits positive feelings with activation of the reward system.

2. Methods

2.1. Participants

Thirty-two healthy right-handed Japanese volunteers (22 women and 10 men; mean age, 22.42 years; range, 19–35 years) participated in this study. Eleven female pairs and 5 male pairs were randomly assigned by order of enrollment. They were not mutually acquainted before the experiment. All participants had normal or corrected-to-normal vision and had no history of neurological, major medical, or psychiatric disorders. Informed consent was obtained from all participants. This study was undertaken in compliance with national legislation and the Code of Ethical Principles for Medical Research Involving Human Subjects of the World Medical Association (the Declaration of Helsinki) and was approved by the ethical committee of the National Institute for Physiological Sciences.

2.2. Task and procedure

Each pair of participants engaged in real-time facial interaction via a 2-way video system while they lay supine in different MRI scanners (Fig. 1). Two MR-compatible video cameras captured the participants' faces, which were then displayed on a screen visible to their partner in a different scanner. The intrinsic delay of this live video system was up to 100 ms (Koike et al., 2019a). Video data of participants' faces were recorded at a rate of 30 frames per second for offline analysis (see 2.3.1. *Kinematic analysis of facial movements*). The task involved an imitation of 3 facial expressions: happy (*H*), sad (*S*), and emotionally neutral with an open mouth (*N*). Each of the paired participants was assigned to be either the initiator (*I*) or the responder (*R*). They were also asked to make the same facial expressions simultaneously as a control for imitation (*C*). The trial consisted of 3 phases: instruction, task, and subjective rating. Participants were instructed as to which imitation condition they had to perform (*I, R*, or *C*) at the beginning of each trial. The instructions, generated using Presentation software (Neurobehavioral System Inc., Albany, CA, USA), appeared for 2 s under the video of the partner's face that overlaid using a Picture-in-Picture system (NAC Image Technology and Panasonic System Solutions Japan Co. Ltd., Tokyo, Japan). When a participant was assigned as the initiator, the instruction indicating the required facial expression was highlighted in magenta. The partner, assigned as the responder, received the instruction "*imitate your partner's facial expression"* on the screen, highlighted in white. In the control condition, the instructions for the facial expression were presented in cyan to both participants. Following the instructions, the cue, a white frame, appeared on the screen for 2 s. Participants were asked to perform the task while the cue was visible and to return to their resting facial expression when it disappeared. Subsequently, the participants reported their current emotional state on a 7-point scale (−3 = negative emotion, 3 = positive emotion). The scale was presented for 3 s, during which participants were prompted to respond using 2 buttons on an MR-compatible button device (HHSC1 × 4-D; Current Designs Inc., Philadelphia, PA, USA) held in their right hands. The paired participants always expressed the same facial expression and could see each other's faces during the experiment. There were 1.5-s periods of rest between the instructions, the imitative interaction, and the subjective rating.

The current experiment used an event-related fMRI design consisting of 5 scanning runs with 36 trials each and a 0.5-s inter-trial interval (3 facial expressions × 3 imitation conditions × 4 repetitions). To optimize design efficiency, the orders of the facial expressions and the

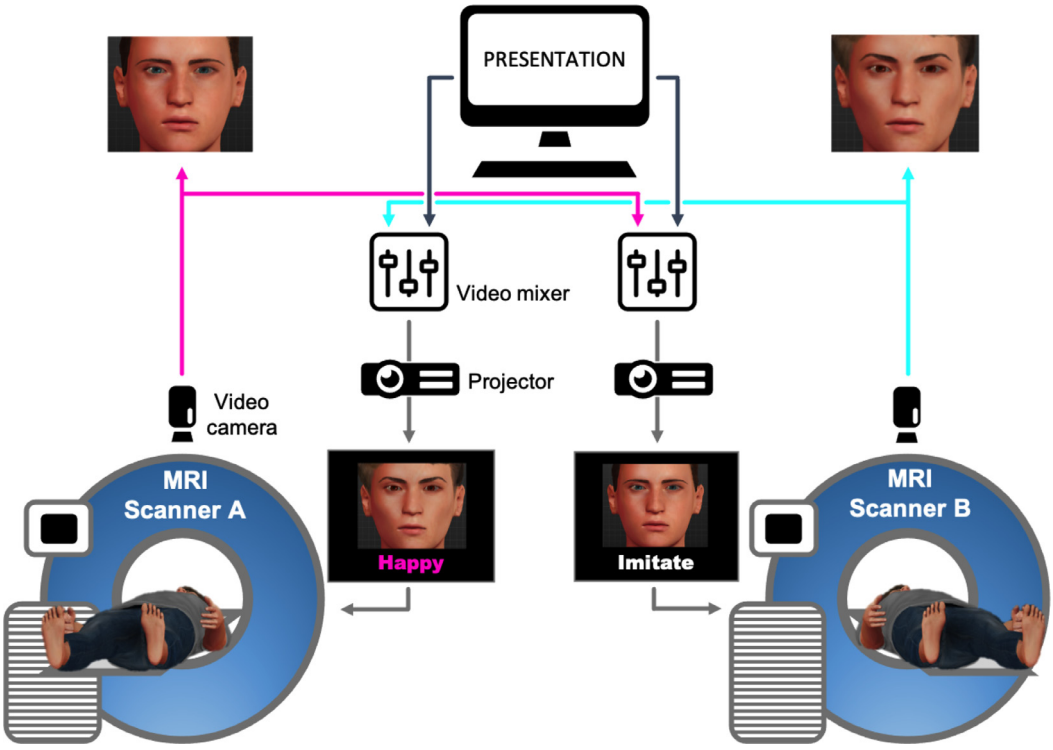

Fig. 1. Experimental setup. Each pair of participants engaged in real-time facial interaction via a 2-way video system while they lay supine in different MRI scanners. Two MR-compatible video cameras captured the participants' faces, which were then displayed on a screen visible to their partner in a different scanner. They were instructed as to which imitation condition they had to perform at the beginning of each trial. The instruction appeared under the video of the partner's face, generated using Presentation software, and overlaid centrally with a video image of the participant's face using a Picture-in-Picture system.

imitation conditions were arranged using a genetic algorithm (Wager and Nichols, 2003), and a null event was included where participants only watched the video of their partner's face to improve regressor orthogonalization. Consequently, 5 different trial order sequences were created that were then randomly assigned to 5 scanning runs across pairs. The experimental recording lasted approximately 45 min. Before the recording, participants practiced making facial expressions outside the scanners for 10 min. We gave participants feedback about the quality of their facial expressions using Face Reader 5 (Noldus Information Technology, Wageningen, The Netherlands). They also practiced the task inside the scanner for 5 min for familiarization with the task and general procedure.

2.3. Behavioral data analyses

2.3.1. Kinematic analysis of facial movements

To characterize the shared intention in action within each pair, similarities in the trial-by-trial variabilities of amplitude and speed during facial movement were evaluated. When the initiator's facial movement was imitated by the partner, their trial-by-trial variabilities were expected to be correlated. Thus, the correlation coefficients of trial-by-trial variabilities of movement amplitude and speed between paired participants were calculated and compared with those of pseudo pairs created by pairing participants from different pairs (240 pairs) to validate that inter-individual correlations were not a result of performing the same action but a result of real-time interaction.

First, the maximum amplitude and speed of lip corner movement were measured as they are assumed to represent facial movement (Schmidt et al., 2003). The data processing procedure is shown in Fig. 2 as an example of the movement speed correlation analysis. Individual facial features were identified and tracked automatically by a standard open-source face detection algorithm OpenFace software (Baltrušaitis et al., 2018). The change in the location of lip corners over time, relative to their default positions (i.e., the most frequent location through recording), was calculated as a movement displacement time series. The movement displacement time series were averaged across lip corners and low-pass filtered using a 7-Hz Butterworth filter. Speeds were then obtained using differential calculus of displacement.

The maximum amplitude and speed were calculated from 333 ms (10 frames) before the cue (white frame) onset to 1,667 ms (50 frames) after onset. The maximum amplitudes and speeds of each trial were aligned by facial expressions (*H, S*, and *N*) and by imitation roles (*I* and *R*) in each run so that the trial order was the same across pairs while retaining imitation roles within a pair (Fig. 2D; also see Inline Supplementary Fig. 1), then concatenated across runs. More specifically, we conducted the following procedures: (1) Within the run of each pair, the maximum amplitude and speed of each trial were re-ordered to the canonical order (T, stands for template) that determined the facial expression followed by the roles of the participants. (2) The 5 re-ordered trial sequences were concatenated as T1-T2-T3-T4-T5 following the run order. Note that the canonical order specified the role of the paired participants so that there were two patterns of the trial sequence depending on the role of the first trial. (3) Thus, we defined (T1-T2-T3-T4-T5)i and (T1-T2-T3-T4-T5)r, where i stands for initiator role, and r stands for responder role. (4) To generate the real and pseudo pairs, we made all combinations of the (T1-T2-T3-T4-T5)i and (T1-T2-T3-T4-T5)r. We made 16 real pairs, and 240 pseudo pairs.

This re-ordering process enabled us to make pseudo pairs in which the combination of imitation condition and facial expression between participants was equivalent to that of real pairs. The difference between real and pseudo pairs depended on whether they had a real-time interaction with each other. Pearson's correlation coefficients of movement amplitude and speed were then computed within the real and pseudo pairs. The correlation coefficients were z-transformed to follow the normal distribution using Fisher's z-transformation and compared between the real and pseudo pairs using unpaired *t*-tests. Thirty-nine of

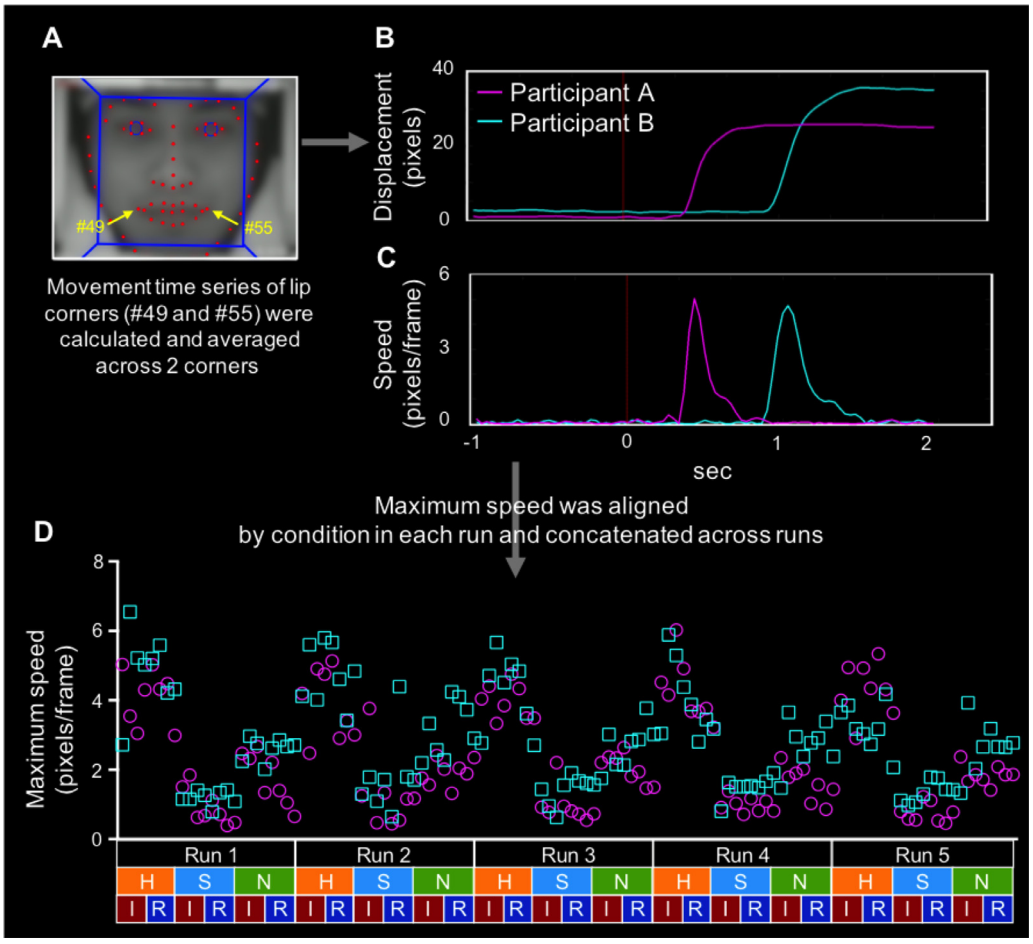

Fig. 2. Example of video data processing flow for calculating the inter-individual correlation of movement speed. (A) Lip corner markers were labeled (#49 and #55) and tracked using OpenFace software. (B) The distance of a lip corner marker from its default position (the most frequent location) was calculated to acquire movement displacement time series that were then averaged across lip corners. The vertical dotted line indicates the cue onset for a trial. (C) Velocities were obtained by differential calculus of displacement. The maximum speed was detected from 333 ms (10 frames) before the cue onset (vertical dotted line) to 1,667 ms (50 frames) after the onset. (D) The maximum speeds were aligned by condition in each run, then concatenated across runs. Pearson's correlation coefficients were calculated and z-transformed using Fisher's z-transformation. *H*, happy; *S*, sad; *N*, non-emotional; *I*, initiator; *R*, responder. We performed the same analysis for the control condition independently; therefore, the control condition was not included in this diagram.

1920 trials (2.03%) were excluded from this video analysis because we could not identify landmarks.

2.3.2. *Subjective rating*

The subjective rating was averaged across trials within each condition and participant. A two-way repeated-measures ANOVA with 2 within-subject factors—facial expression (*H, S,* and *N*) and imitation condition (*I, R,* and *C*)—was performed on the subjective rating using the statistical software SPSS, version 23 (IBM Corporation, Armonk, NY, USA). Subjective rating values of the initiator were normalized by subtracting the value of the control condition from that of each facial expression condition and averaging across facial expressions for the fMRI correlation analysis (see 2.5.1. *Individual brain activation*). The Greenhouse–Geisser correction was used in cases where the result of Mauchly's test of sphericity was significant. Bonferroni correction was used to correct for multiple comparisons. The level for statistical significance was set at $p < .05$.

2.4. *fMRI data acquisition and preprocessing*

The functional images of the paired participants were acquired simultaneously using 2 identical MRI scanners (Magnetom Verio 3T; Siemens Medical System, Erlangen, Germany) that were located 10.5 m apart in a room. A 24-channel phased array coil constructed by replacing the top 12-channel component of a Siemens Verio standard 32-channel phased array coil with a 4-channel flex coil (Siemens) was used to avoid covering the participants' faces (Koike et al., 2016). To obtain functional images with less noise due to facial movement, sparse sampling was used with a T2*-weighted echo planar imaging (EPI) sequence sensitive to blood-oxygen-level-dependent (BOLD) contrast [repetition time (TR) = 3.5 s, acquisition time (TA) = 0.5 s, echo time (TE) = 35 ms, flip angle (FA) = 90°, acquired matrix = 64 × 64, field of view (FoV) = 192 × 192 mm, slice thickness = 3 mm with slice gap 0.51 mm, and 42 slices]. We sampled EPI data immediately before each phase of the trial: instruction, task, and subjective rating (see Inline Supplementary Fig. 2). Six slices were acquired simultaneously by using a multi-band sequence (Moeller et al., 2010) to maximize the acquisition speed. The slices were collected obliquely parallel to the anterior-posterior commissure plane. Each run lasted approximately 8 min, resulting in 128 volumes. Also, high-resolution T1-weighted images were acquired for anatomical localization using the standard 32-channel phased array coil with 3-dimensional magnetization-prepared rapid acquisition gradient echo sequencing (TR = 1,800 ms, TE = 1.98 ms, FoV = 256 × 256 × 176 mm, voxel size = 1 × 1 × 1 mm, FA = 9°).

The fMRI data were analyzed with statistical parametric mapping using SPM12 software (Wellcome Trust Centre for Neuroimaging, Uni-

versity College London, UK) implemented in MATLAB (Version 2016a, Mathworks Inc., Natick, MA, USA). The first 3 volumes of each run were discarded from the analysis. All EPI volumes were realigned to the first volume in the first run by rigid body transformations. A mean image for all EPI volumes was then created, according to which individual volumes were again spatially realigned. The realigned EPI volumes were co-registered with the T1-weighted structural image. The structural image was normalized into standardized Montreal Neurological Institute (MNI) coordinate space using "Diffeomorphic Anatomical Registration using Exponentiated Lie algebra" (DARTEL) normalization (Ashburner, 2007) with a template generated from an independent set of T1-weighted structural images of 512 healthy volunteers (256 women and 256 men; mean age, 23.29 years) using the same MRI system as that in the present study (Rajaei et al., 2018). The normalization parameters were applied to the EPI volumes to ensure an anatomically informed normalization. The normalized EPI volumes were smoothed using an isotropic Gaussian kernel filter of 8 mm with full width at half maximum.

2.5. Statistical analyses

2.5.1. Individual brain activation

The first-level statistical analysis was performed using a general linear model to estimate the parameters for each experimental condition. Vectors containing the white frame appearance onset (duration = 2 s) were convolved with the canonical hemodynamic response function (HRF) to form the main regressors in the design matrix. The design matrix included 9 regressors of interest (*HI, HR, HC, SI, SR, SC, NI, NR*, and *NC*), representing variable combinations for facial expressions [happy (*H*), sad (*S*), or neutral (*N*)] and imitation conditions [initiator/imitatee (*I*), responder/imitator (R), or control (*C*)]. The model also included pushing buttons and visual stimuli for instructions and subjective ratings which were convolved with the HRF, the time series of 6 realignment parameters, and the average signal intensities in white matter and cerebrospinal fluid (CSF) as regressors of no interest. The estimation methods were ordinary least squares. The data were high-pass filtered with a cut-off period of 128 s to remove low-frequency signal drifts. The AR(1) autocorrelation model was globally applied over the brain. Since there were no significant interactions with facial expressions in the subjective rating, the 3 facial expressions were pooled to focus on the main effects of being imitated and imitating. To explore the brain areas associated with imitating and being imitated, 2 contrast images of imitating ($R > C$) and being imitated ($I > C$) were constructed from each participant and used for the second-level analysis.

The second-level random effects analysis was performed to allow inferences across participants using a flexible factorial design. The main effect of the initiator ($I > C$) and responder ($R > C$) were investigated, followed by a conjunction analysis ($I > C \cap R > C$) to examine the common neural activation between imitating and being imitated. To identify the region associated with positive emotions evoked by being imitated, we examined parametric modulation of neural activity based on individual subjective ratings using the contrast image and subjective rating of the initiator ($I > C$).

2.5.2. Inter-brain synchronization

To investigate inter-brain synchronization related to imitation, a beta-series correlation technique initially used to map activation similarity between regions within a brain (Rissman et al., 2004) was extended to an inter-brain beta-series correlation analysis (Koike et al., 2019b). The design matrix was remodeled such that the magnitude of the task-related BOLD response was estimated separately by modeling each trial as a separate regressor. This procedure resulted in 180 regressors of interest (3 facial expressions × 3 imitation conditions × 4 repetitions × 5 runs). For regressors of no interest, we used the same regressors as in the individual brain activation analysis (*2.5.1. Individual brain activation*). The estimation at the first level analysis produced 180 beta maps. Each map represents neural activation during each trial. To compare the real pairs and pseudo pairs, the beta estimate images of each trial were re-ordered with the identical procedure to that for the behavioral analysis. The beta maps were aligned in each run by facial expressions (*H, S*, and *N*) and imitation roles (*I* and *R*). To retain imitation roles within a pair, the order of imitation roles was different within a pair (*I* to *R* or *R* to *I*) (see Inline Supplementary Fig. 1). The reordering process enabled us to postulate pseudo pairs performed the task in the same way as real pairs did. Inter-brain correlations of beta-series were calculated voxel-wise in the same coordinate positions using Pearson's correlation coefficient and z-transformed using Fisher's z-transformation. Following the similar procedure, we also made the re-ordered sequence of the beta contrast images of the control condition to calculate the pair-specific correlation as the control for the imitation related correlation.

To examine the group effect, the second-level analysis for beta-series correlation maps was performed using a flexible factorial design that modeled run and group effects. A conjunction contrast [(real pair > pseudo pair) ∩ (real pair > 0)] was computed using the correlation maps of real pairs and all possible pseudo pairs.

To explore whether the correlation coefficient was also higher when participants were engaged in real-time interaction in the control condition, a conjunction analysis [(real pair > pseudo pair) ∩ (real pair > 0)] and an unpaired t-test was employed to compare the z-transformed correlation coefficient between real pair and pseudo pair. The unpaired t-test was performed at the peak coordinate (x = 68, y = −34, z = 38) that was significant in the imitation condition.

2.5.3. Neural and behavioral synchronization

To assess the relationship between inter-brain activation similarity and facial movement similarity, Pearson's correlation coefficients were computed between the z-transformed correlation coefficient of facial movement amplitude or speed and at the significant peak coordinate (x = 68, y = −34, z = 38) in the beta-series correlation analysis. A permutation test was then performed where 5000 samples were generated by computing correlation coefficients of pseudo pairs. We iterated calculation of correlation coefficients of 16 pairs who were randomly selected from 240 pseudo pairs by the MATLAB function "randperm". We then computed the probability of the correlation coefficient in real pairs against that in randomly selected pseudo pairs.

2.5.4. Statistical thresholding of imaging results and anatomical labeling

For all analyses, the resulting statistical values were height-thresholded with $p < .001$ (uncorrected), and a significant effect was reported when the volume of the cluster survived a family-wise error correction at the cluster level ($p < .05$) in line with recommendations for controlling false positives in the cluster-based correction approach (Eklund et al., 2016). The anatomical locations were determined using the SPM Anatomy Toolbox (Eickhoff et al., 2007, 2006, 2005), and the locations were checked using a paper atlas (Mai et al., 2015). The significant clusters in white matter and CSF areas were masked out. The datasets of fMRI in the current study are available from the corresponding author upon request.

3. Results

To characterize behavioral similarities, the correlation coefficient of the trial-by-trial variabilities of the maximum amplitude and speed of facial movement between paired participants were calculated (Fig. 2). The z-transformed correlation coefficient was higher in real pairs than in pseudo pairs in both movement amplitude [$t(15) = 4.32$, $p < .001$] and speed [$t(15) = 5.46$, $p < .001$], indicating that paired participants tended to match their movements (Fig. 3).

Correspondingly, neural correlates of inter-individual similarity in the trial-by-trial variability of the task-related activity were sought using beta-series correlation analysis. There were correlations between the inter-individual jittering of the imitation-related neural activation of the

Table 1
Neural correlates with inter-individual similarity in the trial-by-trial variability of the task-related activity.

	peak		*MNI coordinates*					
Cluster size (mm^3)	*p-Value (FWE-corr)*	*T*	*x*	*y*	*z*	*Hem*	*Anatomical region*	*(Probability by Anatomy Toolbox)*
984	0.896	4.16	68	−34	38	R	SupraMarginal Gyrus	
	0.941	4.09	56	−46	38	R	Inferior Parietal Lobule	PFm (67%), PF (18%)
	0.971	4.03	52	−44	36	R	SupraMarginal Gyrus	PFm (38%), Area hIP2 (IPS) (13%), Area hIP1 (IPS) (12%)
	0.993	3.93	64	−36	34	R	SupraMarginal Gyrus	PF (67%), PFm (15%)
	1.000	3.58	60	−38	38	R	SupraMarginal Gyrus	PF (64%), PFm (34%)
	1.000	3.52	68	−30	40	R	SupraMarginal Gyrus	

The locations of local maxima were defined by the SPM Anatomy Toolbox v2.2c (Eickhoff et al., 2007, 2006, 2005) and the paper atlas (Mai et al., 2015). Hem, hemisphere; R, right; x, y, z = location (in mm) with the three axes.

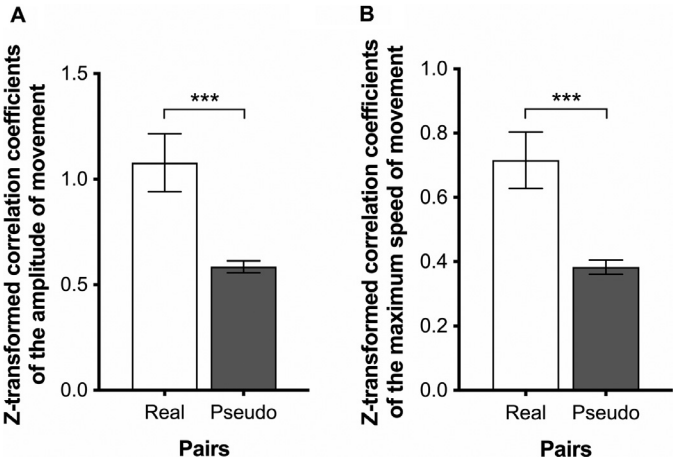

Fig. 3. Inter-individual correlation coefficients of facial movements. (A) The correlation coefficients of facial movement amplitude were higher in the real pairs than in the pseudo pairs. (B) The correlation coefficients of facial movement speed were also higher in the real pairs than in the pseudo pairs. Error bars represent the standard error of the mean. ***: $p < .001$.

pairs in the right IPL (Fig. 4A and Table 1); these values were higher than those of the pseudo pairs. The z-transformed correlation coefficient of the real pairs at the significant peak coordinate (x = 68, y = −34, z = 38) was positive and higher than that of the pseudo pairs in the imitation condition but such a clear difference was not observed in the control condition [$t(254) = 0.44$, $p = .660$] (Fig. 4B and Fig. 4C). Moreover, the z-transformed correlation coefficient at the peak coordinate was correlated with that of movement speed ($r = .48$, $p = .037$) but not with amplitude ($r = .37$, $p = .107$) (Fig. 4D). Therefore, the higher the similarity between the speed of facial movements, the higher the similarity between the activation patterns of the right IPL.

Compared to the control condition (to eliminate movement execution), the responder condition activated the midcingulate cortex, the pre-supplementary motor area (pre-SMA), supplementary motor area (SMA) (Mayka et al., 2006), the bilateral putamen and precuneus, the left ventral premotor area (PMv) extending to the primary motor cortex, and the right middle occipital and inferior temporal gyri (Fig. 5 and Table 2). On the contrary, compared to the control condition, the initiator condition activated the dorsomedial prefrontal cortex (dmPFC) extending to the anterior cingulate cortex (ACC), the pre-SMA, the bilateral superior temporal sulcus, the cerebellum, the temporoparietal junction (TPJ), and the right IFG extending to the PMv (Fig. 5 and Table 2).

Regarding the subjective rating, participants reported more positive emotions when being imitated than when imitating or simultaneously making facial expressions with their partner (Fig. 6). ANOVA yielded significant main effects for the imitation condition [$F(2, 62) = 10.19$, $p < .001$] and facial expression [$F(1.36, 42.01) = 49.42$, $p < .001$] but no significant 2-way interaction [$F(4, 124) = 1.20$, $p = .315$]. Simple effect tests revealed that the rating value of the initiator was significantly higher than that of the responder ($p = .029$) and the control condition ($p = .001$). There was no significant difference between the responder and control conditions. Thus, the experience of being imitated induced positive emotions irrespective of the facial expression. Individuals who reported more positive emotions after being imitated showed higher activity in the rostral anterior cingulate cortex (rACC), evidenced by a positive correlation between the subjective ratings and rACC activity (peak coordinates: x = 8, y = 26, z = 18) of the initiator compared to the control (Fig. 6), suggesting that the rACC is associated with the positive emotions evoked by being imitated.

4. Discussion

4.1. Behaviorally relevant inter-brain synchronization of the right IPL

In this experiment, we aimed to depict the shared representation of action, that is, what is shared during imitation at behavioral and neural levels. We found that the kinematic parameters are shared at the behavioral level and the task-specific BOLD responses of the right IPL at the neural level. Inter-brain synchronization in the right IPL was correlated with facial movement kinematic profile similarity (Fig. 4D). Thus, neural representation of the kinematic parameters was shared in the right IPL. The issue is what level of the hierarchy of the motor control system the shared neural representation of the right IPL corresponds to: prior intention, intention in action, or motor command (de Vignemont and Haggard, 2008).

Prior intention represents the goal of the action as a global unit. There may be several means to realize the same prior intention so that it is not sufficient to specify the movement (de Vignemont and Haggard, 2008). Therefore, the inter-brain synchronization of the right IPL is unlikely to represent prior intention. As there was no inter-brain neural synchronization in the primary motor cortex, which dispatches the final motor command to the muscles, it is unlikely the right IPL represents the motor command *per se*. Intention in action, on the other hand, represents the action as a dynamic sequence of specific movement. It initiates the action, guides it, and monitors its effects. It corresponds to the motor plan, prior to the dispatch of the final motor command to the muscles (de Vignemont and Haggard, 2008). Therefore, intention in action has a motor specificity that prior intention lacks. For this reason, intention in action can be identified with the forward model, which computes the sensory outcome of the motor commands needed to achieve the desired state (de Vignemont and Haggard, 2008; Wolpert et al., 1995).

The trial-by-trial variability in the maximum amplitude and speed of facial movement showed a higher correlation within real than pseudo pairs. The responder copies both categorical facial expressions (happiness and sadness) and their kinematic profiles which describe the action as a dynamic sequence, i.e., intention in action. The intention in action was only shared within pairs who interacted in real time, whereas the category of facial expressions associated with the goal of the action, i.e., prior intention, was shared by both the real and pseudo pairs. Thus, the

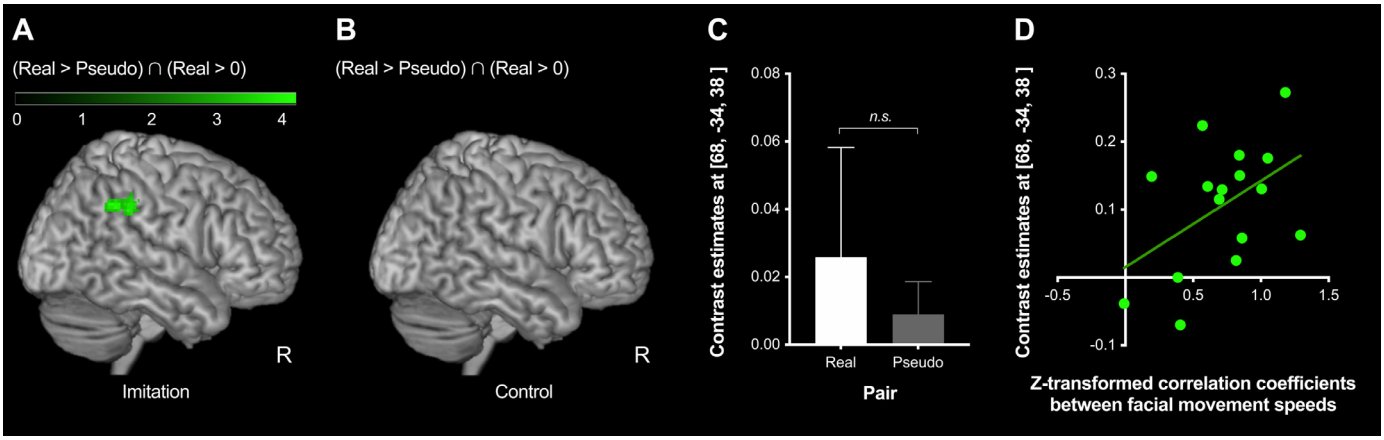

Fig. 4. Pair-specific inter-brain correlation. (A) The results of inter-brain beta-series correlation analysis during imitative interaction. Significant clusters were overlaid on the MNI template image. Statistical thresholds were set at $p < .001$ (uncorrected) at the voxel level and $p < .05$ (FWE-corrected) at the cluster level. (B) The results of inter-brain beta-series correlation analysis in the control condition. There was no significant voxel. (C) The z-transformed correlation coefficient at the significant peak coordinate (x = 68, y = −34, z = 38) was not different between the real and pseudo pairs in the control condition. Error bars represent the standard error of the mean. (D) A significant correlation between the z-transformed correlation coefficient at the significant peak coordinate (x = 68, y = −34, z = 38) and that of facial movement speed ($r = .48$, $p = .037$). The trend line represents a linear regression line. MNI, Montreal Neurological Institute; FWE, family-wise error.

pair-specific inter-individual correlation of kinematic profiles represents the sharing of intention in action.

This notion is consistent with the view of imitation as a prediction/control relation or forward/inverse internal model of motor control (Wolpert et al., 1998; Wolpert and Kawato, 1998); the imitatee made the facial expression which is controlled by the prediction of one's own action (forward model), while imitating is the control of self-action in the context of another's action (inverse model). The forward and inverse models are tightly coupled during both their acquisition and use (Wolpert and Kawato, 1998), because the same forward model used to predict the sensorial effects of our own actions can also be used as a constraint for predicting the actions of others (Friston, 2005; Kilner et al., 2007). Considering the inter-individual causal relationship between imitating and being imitated, the neural substrates of the pair-specific inter-individual synchronization represent the shared forward model, the set of the commands specifying facial expression movements. The forward model of facial movement is shared between the pair of individuals who play different roles. Therefore, the observed pair-specific synchronization can be explained by the shared forward model generated by inter-individual causal relationship. This interpretation is further supported by the fact that the neural synchronization of the right IPL disappeared during the control condition, which lacked the prediction/control components.

Involvement of the right parietal region was shown by the previous hyperscanning fMRI studies of joint attention (Bilek et al., 2015) and joint force production (Abe et al., 2019), and hyperscanning electroencephalography studies of face-to-face imitation (Dumas et al., 2010), interpersonal rhythmic movement synchrony (Tognoli et al., 2007), and joint musical performance (Novembre et al., 2016), suggesting less specific function such as self-other differentiation. The present study is distinct from the previous studies in that (1) the kinematic profile of action is shared between two participants in a face-to-face situation, (2) the task-specific activity in the right IPL was inter-individually synchronized, and (3) the neural synchronization was correlated with the similarity in facial movement kinematic profile. These specific findings lead us to conclude that the right IPL codes intention in action.

The IPL is related to motor planning (Quiroga et al., 2006) and motor imagery (Jeannerod, 2006; Rizzolatti and Matelli, 2003; Sirigu et al., 1996; Wheaton and Hallett, 2007) and is considered as the area responsible for intentions, emerging through neural computation within the motor network for action preparation and planning (Desmurget et al., 2009; Desmurget and Sirigu, 2009). The parietal lesion in humans is known to cause ideomotor apraxia, the inability to pantomime, imitate, and use tools properly (Wheaton and Hallett, 2007). The important role of the posterior parietal cortex in sensorimotor transformation and integration is well established (Andersen, 1995; Andersen and Buneo, 2002). A non-human primate study has shown that cells in the lateral intraparietal area have activity specifically related to movements the animal intends to make, thus representing the intention in action (Snyder et al., 1997).

The literature on the coding of intentions indicates the involvement of the frontal regions. By manipulating the context during functional MRI, Iacoboni et al. (2005) added the different intentional component of the same grasping action (either intention of drinking or cleaning). They found that the premotor mirror system is understanding the prior intentions of others. de Lange et al. (2008) asked participants to judge whether the intention of the action was ordinary, whereas on other trials, participants had to judge whether the action was carried out in an ordinary manner. They found that observing actions with extraordinary intentions was associated with increased activation (compared to ordinary actions) in the IFG irrespective of whether participants paid attention to the intention or manner. Thus, de Lange et al. (2008) also concluded that the intention of others (i.e., prior intention) is coded in the frontal mirror system.

With the distinction of the two types of intention, the results of Iacoboni et al. (2005) and de Lange et al. (2008) indicate the role of the frontal MNS for prior intention, and those of the present study suggest the role of the parietal MNS for intention in action. This difference comes from the task design: present study requires participant to imitate the facial expression of the partner, thus specific motor parameters needs to be shared. These findings lead us to conclude that the inter-individual neural synchronization of the right IPL represents the shared intention in action, that is, the SRA.

4.2. *Different neural activation reflecting different psychological processes*

We used the condition of simultaneously making the same facial expressions as a control for imitation. Neural activations in this condition represent the shared process between imitating and being imitated, such as making facial expressions and observing a partner's facial expressions. Therefore, the remaining activations (after subtracting the activation in the control) were specific to imitating and being imitated. As expected, neural activation elicited by imitating and being imitated was distinct

Table 2
Neural correlates with main effect of initiator (top) and responder (bottom).

Initiator > Control								
	peak		*MNI coordinates*					
Cluster size (mm³)	*p-Value (FWE-corr)*	*T*	*x*	*y*	*z*	*Hem*	*Anatomical region*	*(Probability by Anatomy Toolbox)*
16616	0.000	7.74	46	−24	−2	R	Superior Temporal Gyrus	
	0.007	5.63	56	−42	10	R	Middle Temporal Gyrus	
	0.012	5.49	60	−44	42	R	SupraMarginal Gyrus	PFm (63%), PF (22%)
	0.035	5.20	60	−38	38	R	SupraMarginal Gyrus	PF (64%), PFm (34%)
	0.128	4.81	52	−48	26	R	SupraMarginal Gyrus	PFm (28%), PGa (20%)
	0.974	3.75	66	−44	22	R	Superior Temporal Gyrus	PF (38%), PFm (37%), PGa (16%)
	1.000	3.24	56	−56	12	R	Middle Temporal Gyrus	PGa (45%), PGp (24%)
1920	0.001	6.08	−52	−26	2	L	Superior Temporal Gyrus	
	0.004	5.80	−48	−30	0	L	Middle Temporal Gyrus	
2376	0.004	5.76	10	12	72	R	Posterior-Medial Frontal	
12040	0.009	5.56	−54	−48	40	L	Inferior Parietal Lobule	PFm (51%), PF (26%), PGa (15%)
	0.012	5.48	−60	−50	30	L	SupraMarginal Gyrus	PFm (50%), PF (18%), PGa (13%)
	0.023	5.31	−58	−48	22	L	Superior Temporal Gyrus	PFm (19%)
	0.026	5.28	−56	−50	28	L	SupraMarginal Gyrus	PFm (28%), PGa (26%), PF(21%)
	0.050	5.10	−58	−54	28	L	SupraMarginal Gyrus	PFm (46%), PGa (40%), PF(14%)
	0.055	5.07	−42	−58	20	L	Middle Temporal Gyrus	
	0.073	4.99	−38	−54	22	L	Angular Gyrus	
	0.096	4.90	−52	−60	20	L	Middle Temporal Gyrus	PGp (30%), PGa (25%), PFm (15%)
5096	0.018	5.39	−6	48	28	L	Superior Medial Gyrus	
	0.612	4.21	10	50	24	R	Superior Medial Gyrus	
1496	0.107	4.87	30	−88	−20	R	Area hOc3v [V3v]	Area hOc3v [V3v] (46%), Area hOc4v [V4(v)] (27%)
	0.714	4.12	22	−80	−28	R	Cerebelum	Lobule VIIa crusI (94%)
2984	0.155	4.75	50	20	−4	R	IFG (p. Orbitalis)	Area 45 (31%)
	0.194	4.68	52	22	0	R	IFG (p. Triangularis)	Area 45 (48%), Area 44 (18%)
	0.364	4.45	50	22	20	R	IFG (p. Triangularis)	Area 45 (17%)
1648	0.294	4.53	−26	−76	−20	L	Cerebelum	Lobule VIIa crusI (25%), Area hOc4v [V4(v)] (16%)
	0.515	4.30	−24	−80	−24	L	Cerebelum	Lobule VIIa crusI (92%)
	0.773	4.06	−20	−78	−34	L	Cerebelum	Lobule VIIa crusI (82%), Lobule VIIa crusII (15%)
1488	0.539	4.28	4	34	24	R	ACC	
	0.998	3.53	2	26	32	R	MCC	
Responder > Control								
	peak		*MNI coordinates*					
Cluster size (mm³)	*p-Value (FWE-corr)*	*T*	*x*	*y*	*z*	*Hem*	*Anatomical region*	*(Probability by Anatomy Toolbox)*
11056	0.000	6.56	−6	−78	48	L	Precuneus	Area 7P (SPL) (72%), Area 7A (SPL) (12%)
	0.006	5.68	−2	−66	60	L	Precuneus	Area 7A (SPL) (40%)
	0.033	5.21	0	−74	52	L	Precuneus	Area 7A (SPL) (30%), Area 7P (SPL) (26%)
	0.054	5.07	14	−68	46	R	Precuneus	
5264	0.001	6.25	−26	−4	14	L	Putamen	
	0.551	4.27	−16	2	16	L	Caudate Nucleus	
8456	0.001	6.06	−6	−4	68	L	Posterior-Medial Frontal	
	0.625	4.20	−10	16	38	L	MCC	
	0.992	3.64	6	−6	64	R	Posterior-Medial Frontal	
4416	0.002	5.95	24	−4	10	R	Putamen	
4768	0.009	5.56	−54	−4	40	L	Postcentral Gyrus	
	0.028	5.26	−56	−2	34	L	Precentral Gyrus	
	0.571	4.25	−58	−12	36	L	Postcentral Gyrus	Area 3b (41%), Area 4p (11%)
	0.976	3.74	−58	−12	24	L	Postcentral Gyrus	Area 3b (29%)
2976	0.014	5.45	−4	−18	34	L	MCC	
	0.051	5.09	6	−20	34	R	MCC	
3704	0.175	4.71	40	−78	0	R	Middle Occipital Gyrus	Area hOc4la (54%), Area hOc4lp (18%)
	0.399	4.41	30	−96	10	R	Middle Occipital Gyrus	Area hOc3d [V3d] (55%), Area hOc4lp (20%)
	0.593	4.23	48	−60	−10	R	Inferior Temporal Gyrus	Area FG2 (11%)
	0.984	3.70	42	−72	−6	R	Inferior Temporal Gyrus	Area hOc4la (63%), Area FG1 (13%)
	0.998	3.53	40	−92	4	R	Middle Occipital Gyrus	Area hOc4lp (62%), Area hOc3v [V3v] (14%), Area hOc4d [V3A] (12%)
	1.000	3.43	22	−102	16	R	Area hOc2 [V2]	

The locations of local maxima were defined by the SPM Anatomy Toolbox v2.2c (Eickhoff et al., 2007, 2006, 2005) and the paper atlas (Mai et al., 2015). Hem, hemisphere; L, left; R, right; x, y, z = location (in mm) with the three axes; ACC: anterior cingulate cortex; IFG; inferior frontal gyrus; MCC: midcingulate cortex.

with little overlap. This finding might reflect the different psychological processes of imitating and being imitated (Hale and Hamilton, 2016).

4.2.1. Imitating

In the present study, there was significant activation in the left PMv and pre-SMA/SMA during imitating compared with the control condition; both are reportedly involved in action observation and imitation (Caspers et al., 2010). In addition to cortical activation in the occipito-parieto-frontal region, the striatum, a part of the basal ganglia, was also activated. The imitating condition required participants to make facial expressions from a non-verbal cue: the partner's facial expression. In contrast, the control condition required participants to make a facial expression cued by a verbal command. Thus, the imitating condition required participants to convert visual input into a motor output. Notably, ideomotor apraxia, characterized by a failure to imitate actions (Tessari et al., 2007), was observed in patients with basal ganglia dys-

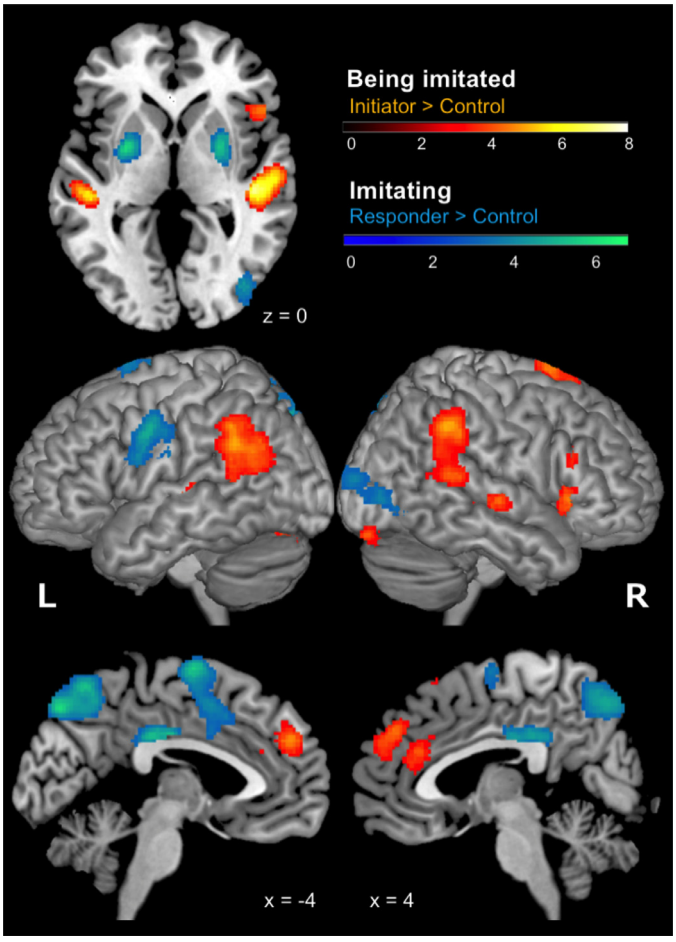

Fig. 5. Neural activations by being imitated and by imitating. A significant cluster was overlaid on the MNI template image. Statistical thresholds were set at $p < .001$ (uncorrected) at the voxel level and $p < .05$ (FWE-corrected) at the cluster level. L, left; R, right; MNI, Montreal Neurological Institute; FWE, family-wise error.

function (Leiguarda, 2001; Wheaton and Hallett, 2007; Zadikoff and Lang, 2005). The basal ganglia participate in the early stages of processing biological movements, possibly by selecting suitable motor programs to match the stimulus (Kessler et al., 2006). It is also involved in selecting kinematic parameters and responses by inhibiting inappropriate responses to behaviorally relevant stimuli (Lawrence et al., 1999; Rushworth et al., 1998). This represents the functions of the basal ganglia in the visuomotor conversion mechanism necessary for imitation (Tessari et al., 2007).

4.2.2. *Being imitated*

In contrast with the control condition, being imitated showed activation of the parieto-premotor mirror system, comprising the right IFG and IPL, and the mentalizing network, comprising the dmPFC and bilateral TPJ. The involvement of these two systems is consistent with the model of being imitated proposed by Hale and Hamilton (2016) featuring perception-action matching and self-other processing. The mirror system participates in perception-action matching (Fogassi et al., 2005; Rizzolatti et al., 1997; Rizzolatti and Craighero, 2004; Rizzolatti and Sinigaglia, 2016). The mentalizing network is related to the self-other distinction critical for contingency detection (Decety and Sommerville, 2003; Farrer and Frith, 2002; Ruby and Decety, 2001; Uddin et al., 2006). Previous studies have demonstrated that the mirror system and the mentalizing network have distinct but complementary functions in understanding the prior intentions of others (de Lange et al., 2008; van Overwalle and Baetens, 2009). During the mutual imitation, the IFG and the IPL, enhanced functional connectivity with the mentalizing network (Sperduti et al., 2014). The excessive connectivity between the mirror system and the mentalizing network was observed in young adolescent with ASD compared with the matched control (Fishman et al., 2014). The default mode network, a system for psychological self-relevant processing and mentalizing, and the mirror system may interact through densely connected "hubs" such as the anterior insula and the posterior cingulate cortex/precuneus (Molnar-Szakacs and Uddin, 2013). Considering the proximity of the IPL to the TPJ, the right TPJ might also be a critical node that links the 2 distinct systems of the parieto-premotor mirror system and the mentalizing network for processing the action of being imitated. A further study eval-

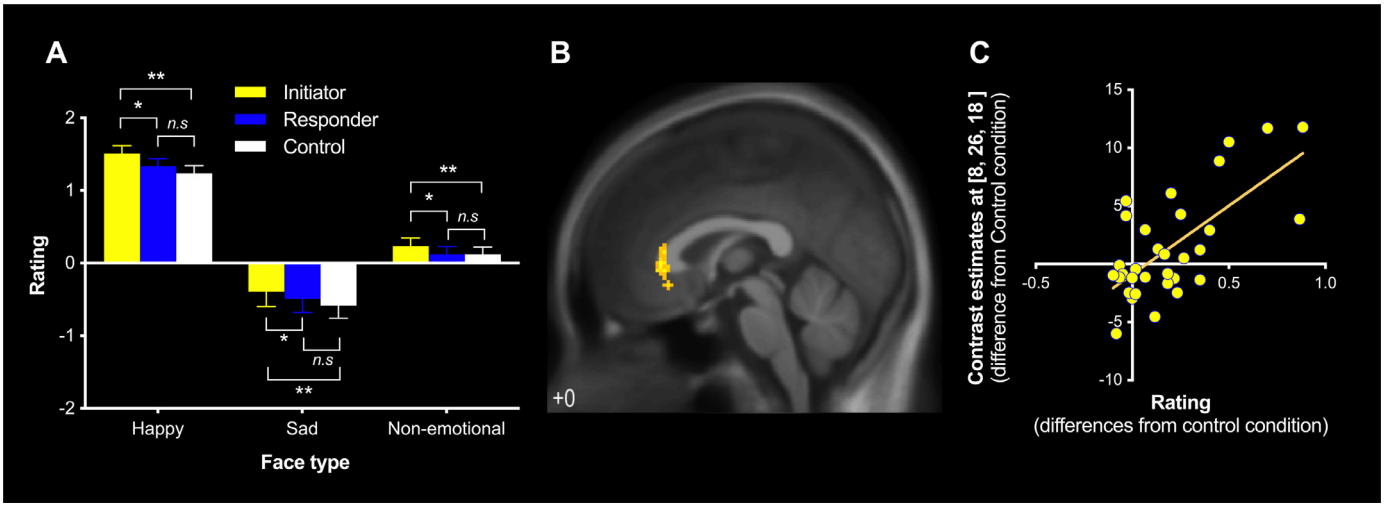

Fig. 6. Association between the subjective rating of emotional valence and rACC activity. (A) The subjective rating results. Positive rating values indicate positive emotions, and negative values indicate negative emotions. The experience of being imitated increased subjective ratings irrespective of facial expression, as indicated by positive values. Error bars represent the standard error of the mean. (B) A statistical parametric map illustrating the cluster (yellow) which was significantly correlated with normalized subjective ratings (*I–C*) in the initiator using the contrast of the initiator ($I > C$). Significant clusters were overlaid on the mean T1-weighted image of all 32 participants. Statistical thresholds were set at $p < .001$ (uncorrected) at the voxel level and $p < .05$ (FWE-corrected) at the cluster level. (C) The scatter plot demonstrates the correlation between contrast estimates at [8, 26, 18] (rACC) and subjective ratings in the initiator. **: $p < .01$, *: $p < .05$, *n.s*: not significant. *I*, initiator; *R*, responder; *C*, control; rACC: rostral anterior cingulate cortex; FWE, family-wise error.

uating the connectivity between the mirror system and the mentalizing network is warranted.

The model of being imitated proposed by Hale and Hamilton (2016) also includes the reward system. Positive affect and prosocial behaviors are induced by the detection of being imitated (Chartrand and Bargh, 1999; van Baaren et al., 2003; Van Baaren et al., 2004; Kühn et al., 2010; Hale and Hamilton, 2016). Positive consequences are reported in both unconscious and conscious detection of being imitated (Catmur and Heyes, 2013). A reward system is involved in these positive responses (Kokal et al., 2011; Kühn et al., 2010). In the present study, positive emotions elicited by being imitated were positively correlated with rACC activation. The rACC is a proven affective division of the ACC related to positive emotional states (Etkin et al., 2011, 2006) such as emotional laughter (Caruana et al., 2015; Szameitat et al., 2010). The rACC also plays an important role in future reward estimation (Rushworth et al., 2011; Tanaka et al., 2016). Krueger et al. (2009) proposed that the rACC may act as an event simulator that is involved in reward estimation by integrating information from the dorsal mPFC and the orbitofrontal cortex. Previously, Matsunaga et al. (2016) found that the rACC is related to both a temporary positive emotion and a trait-like long-term sense of being happy; the latter was shown through a positive correlation between gray matter density of the rACC and trait-like subjective happiness. Our findings, along with those of Matsunaga et al. (2016), underscore the importance of the mutuality of imitation in social settings, whereby being imitated could be a reward induced by the self-contingent responses from others (social contingency).

4.3. Limitations

In real social interaction, the turn-taking between imitatee and imitator occurs dynamically and spontaneously and may lead to sharing of emotions. In our experimental paradigm, however, the spontaneity and dynamicity of the mutual imitation were sacrificed because we introduced explicit turn-taking to depict the processes of imitating and being imitated separately. Furthermore, we tested only two emotional expressions because happy and sad faces are more likely to be shared in everyday human interaction. Further studies utilizing hyperscanning EEG / NIRS (near infrared spectroscopy) are warranted to test if our findings can be generalized to other emotional expressions in more ecological conditions.

4.4. Conclusion

While neural activation elicited by imitating and being imitated was distinct with little overlap, on-line imitative interaction enhanced interbrain synchronization in the right IPL that correlated with the similarity in facial movement kinematic profile. We conclude that the interindividual neural synchronization of the right IPL during face-to-face interactions represents shared representation of action as pair-specific forward internal model, corresponding to a dynamic sequence of movement, that is, intention in action.

Declaration of Competing Interest

None.

Acknowledgments

The authors wish to express their gratitude to Mr. Yoshikuni Ito, Mrs. Reiko Kimura, Mrs. Megumi Iwase, and Mrs. Kuniko Takenaka for their kind supports in conducting the experiment.

Funding

This study was supported, in part, by a Grant-in-Aid for Scientific Research A #15H01846 (N.S.), Grant-in-Aid for Young Scientists B #17K13177 (K.M.), Challenging Exploratory Research grants #15K12775 (T.K.) from the Japan Society for the Promotion of Science, and the Japan Agency for Medical Research and Development [grant numbers JP18dm0107152 (N.S.) and JP18dm0307005 (N.S.)].

CRediT author statement

Kohei Miyata: Conceptualization, Investigation, Formal analysis, Writing – original draft, Visualization, Project administration, and Funding acquisition. **Takahiko Koike:** Conceptualization, Investigation, Formal analysis, Writing – reviewing & editing, and Funding acquisition. **Eri Nakagawa:** Investigation and Writing – reviewing & editing. **Tokiko Harada:** Investigation and Writing – reviewing & editing. **Motofumi Sumiya:** Investigation and Writing – reviewing & editing. **Tetsuya Yamamoto:** Software and Formal analysis. **Norihiro Sadato:** Conceptualization, Investigation, Writing – reviewing & editing, Supervision, Funding acquisition.

Data and code availability statement

The data that support the findings of this study are available from the corresponding author, N.S., upon reasonable request. Some data are not publicly available due to their containing information that could compromise the privacy of participants.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2021.117916.

References

Abe, M.O., Koike, T., Okazaki, S., Sugawara, S.K., Takahashi, K., Watanabe, K., Sadato, N., 2019. Neural correlates of online cooperation during joint force production. Neuroimage 191, 150–161. doi:10.1016/j.neuroimage.2019.02.003.

Andersen, R.A., 1995. Encoding of intention and spatial location in the posterior parietal cortex. Cereb. Cortex 5, 457–469. doi:10.1093/cercor/5.5.457.

Andersen, R.A., Buneo, C.A., 2002. Intentional maps in posterior parietal cortex. Annu. Rev. Neurosci. 25, 189–220. doi:10.1146/annurev.neuro.25.112701.142922.

Ashburner, J., 2007. A fast diffeomorphic image registration algorithm. Neuroimage 38, 95–113. doi:10.1016/j.neuroimage.2007.07.007.

Baltrušaitis, T., Zadeh, A., Lim, Y.C., Morency, L.P., 2018. OpenFace 2.0: facial behavior analysis toolkit. In: IEEE International Conference on Automatic Face and Gesture Recognition, Piscataway, New Jersey. IEEE, pp. 59–66.

Barsalou, L.W., Niedenthal, P.M., Barbey, A.K., Ruppert, J.A., 2003. Social embodiment. In: Psychology of Learning and Motivation. Elsevier Science, New York, pp. 43–92. doi:10.1016/S0079-7421(03)01011-9.

Bilek, E., Ruf, M., Schäfer, A., Akdeniz, C., Calhoun, V.D., Schmahl, C., Demanuele, C., Tost, H., Kirsch, P., Meyer-Lindenberg, A., 2015. Information flow between interacting human brains: identification, validation, and relationship to social expertise. Proc. Natl. Acad. Sci. U. S. A. 112, 5207–5212. doi:10.1073/pnas.1421831112.

Call, J., Carpenter, M., 2002. Three sources of information in social learning. In: Dautenhahn, K., Nehaniv, C.L. (Eds.), Imitation in Animals and Artifacts. The MIT Press, pp. 211–228.

Caruana, F., Avanzini, P., Gozzo, F., Francione, S., Cardinale, F., Rizzolatti, G., 2015. Mirth and laughter elicited by electrical stimulation of the human anterior cingulate cortex. Cortex 71, 323–331. doi:10.1016/j.cortex.2015.07.024.

Caspers, S., Zilles, K., Laird, A.R., Eickhoff, S.B., 2010. ALE meta-analysis of action observation and imitation in the human brain. Neuroimage 50, 1148–1167. doi:10.1016/j.neuroimage.2009.12.112.

Catmur, C., Heyes, C., 2013. Is it what you do, or when you do it? The roles of contingency and similarity in pro-social effects of imitation. Cogn. Sci. 37, 1541–1552. doi:10.1111/cogs.12071.

Chartrand, T.L., Bargh, J.A., 1999. The chameleon effect: The perception-behavior link and social interaction. J. Pers. Soc. Psychol. 76, 893–910. doi:10.1037/0022-3514.76.6.893.

de Lange, F.P., Spronk, M., Willems, R.M., Toni, I., Bekkering, H., 2008. Complementary systems for understanding action intentions. Curr. Biol. 18, 454–457. doi:10.1016/j.cub.2008.02.057.

de Vignemont, F., Haggard, P., 2008. Action observation and execution: What is shared? Soc. Neurosci. 3, 421–433. doi:10.1080/17470910802045109.

Decety, J., Sommerville, J.A., 2003. Shared representations between self and other: A social cognitive neuroscience view. Trends Cogn. Sci. 7, 527–533. doi:10.1016/j.tics.2003.10.004.

Desmurget, M., Reilly, K.T., Richard, N., Szathmari, A., Mottolese, C., Sirigu, A., 2009. Movement intention after parietal cortex stimulation in humans. Science 324 (80-), 811–813. doi:10.1126/science.1169896.

Desmurget, M., Sirigu, A., 2009. A parietal-premotor network for movement intention and motor awareness. Trends Cogn. Sci. 13, 411–419. doi:10.1016/j.tics.2009.08.001.
Dumas, G., Nadel, J., Soussignan, R., Martinerie, J., Garnero, L., 2010. Inter-brain synchronization during social interaction. PLoS One 5, e12166. doi:10.1371/journal.pone.0012166.
Eickhoff, S.B., Heim, S., Zilles, K., Amunts, K., 2006. Testing anatomically specified hypotheses in functional imaging using cytoarchitectonic maps. Neuroimage 32, 570–582. doi:10.1016/j.neuroimage.2006.04.204.
Eickhoff, S.B., Paus, T., Caspers, S., Grosbras, M.H., Evans, A.C., Zilles, K., Amunts, K., 2007. Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. Neuroimage 36, 511–521. doi:10.1016/j.neuroimage.2007.03.060.
Eickhoff, S.B., Stephan, K.E., Mohlberg, H., Grefkes, C., Fink, G.R., Amunts, K., Zilles, K., 2005. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. Neuroimage 25, 1325–1335. doi:10.1016/j.neuroimage.2004.12.034.
Eklund, A., Nichols, T.E., Knutsson, H., 2016. Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. Proc. Natl. Acad. Sci. U. S. A. 113, 7900–7905. doi:10.1073/pnas.1602413113.
Etkin, A., Egner, T., Kalisch, R., 2011. Emotional processing in anterior cingulate and medial prefrontal cortex. Trends Cogn. Sci. 15, 85–93. doi:10.1016/j.tics.2010.11.004.
Etkin, A., Egner, T., Peraza, D.M., Kandel, E.R., Hirsch, J., 2006. Resolving emotional conflict: A role for the rostral anterior cingulate cortex in modulating activity in the amygdala. Neuron 51, 871–882. doi:10.1016/j.neuron.2006.07.029.
Farrer, C., Frith, C.D., 2002. Experiencing oneself vs another person as being the cause of an action: the neural correlates of the experience of agency. Neuroimage 15, 596–603. doi:10.1006/nimg.2001.1009.
Fishman, I., Keown, C.L., Lincoln, A.J., Pineda, J.A., Müller, R.A., 2014. Atypical cross talk between mentalizing and mirror neuron networks in autism spectrum disorder. JAMA Psychiatry 71, 751–760. doi:10.1001/jamapsychiatry.2014.83.
Fogassi, L., Ferrari, P.F., Benno, G., Rozzi, S., Chersi, F., Rizzolatti, G., 2005. Parietal lobe: from action organization to intention understanding. Science 308 (80-), 662–667. doi:10.1126/science.1106138.
Friston, K., 2008. Hierarchical models in the brain. PLoS Comput. Biol. 4. doi:10.1371/journal.pcbi.1000211.
Friston, K., 2005. A theory of cortical responses. Philos. Trans. R. Soc. B Biol. Sci. 360, 815–836. doi:10.1098/rstb.2005.1622.
Friston, K., Frith, C., 2015. A duet for one. Conscious. Cogn. 36, 390–405. doi:10.1016/j.concog.2014.12.003.
Gallese, V., 2003. The roots of empathy: the shared manifold hypothesis and the neural basis of intersubjectivity. Psychopathology 36, 171–180. doi:10.1159/000072786.
Gopnik, A., Meltzoff, A.N., 1994. Minds, bodies, and persons: young children's understanding of the self and others as reflected in imitation and theory of mind research. In: Parker, U.T., Mitchell, R.W., Boccia, M.L. (Eds.), Self-Awareness in Animals and Humans: Developmental Perspectives. Cambridge University Press, New York, pp. 166–186.
Hale, J., Hamilton, A.F., 2016. Cognitive mechanisms for responding to mimicry from others. Neurosci. Biobehav. Rev. 63, 106–123. doi:10.1016/j.neubiorev.2016.02.006.
Hobson, R.P., 2003. The Cradle of Thoughts. Oxford University Press, New York.
Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J.C., 2005. Grasping the intentions of others with one's own mirror neuron system. PLoS Biol 3, 0529–0535. doi:10.1371/journal.pbio.0030079.
Iacoboni, M., Woods, R.P., Brass, M., Bekkering, H., Mazziotta, J.C., Rizzolatti, G., 1999. Cortical mechanisms of human imitation. Science 286 (80-), 2526–2528. doi:10.1126/science.286.5449.2526.
Jeannerod, M., 2006. Motor Cognition. Oxford University Press, New York.
Jeannerod, M., 1995. Mental imagery in the motor context. Neuropsychologia 33, 1419–1432.
Kessler, K., Biermann-Ruben, K., Jonas, M., Roman Siebner, H., Bäumer, T., Münchau, A., Schnitzler, A., 2006. Investigating the human mirror neuron system by means of cortical synchronization during the imitation of biological movements. Neuroimage 33, 227–238. doi:10.1016/j.neuroimage.2006.06.014.
Kilner, J.M., Friston, K.J., Frith, C.D., 2007. Predictive coding: an account of the mirror neuron system. Cogn. Process. 8, 159–166. doi:10.1007/s10339-007-0170-2.
Koike, T., Sumiya, M., Nakagawa, E., Okazaki, S., Sadato, N., 2019a. What makes eye contact special? Neural substrates of on-line mutual eye-gaze: a hyperscanning fMRI study. eNeuro 6, 1–18. doi:10.1523/ENEURO.0284-18.2019.
Koike, T., Tanabe, H.C., Adachi-Abe, S., Okazaki, S., Nakagawa, E., Sasaki, A.T., Shimada, K., Sugawara, S.K., Takahashi, H.K., Yoshihara, K., Sadato, N., 2019b. Role of the right anterior insular cortex in joint attention-related identification with a partner. Soc. Cogn. Affect. Neurosci. 14, 1131–1145. doi:10.1093/scan/nsz087.
Koike, T., Tanabe, H.C., Okazaki, S., Nakagawa, E., Sasaki, A.T., Shimada, K., Sugawara, S.K., Takahashi, H.K., Yoshihara, K., Bosch-Bayard, J., Sadato, N., 2016. Neural substrates of shared attention as social memory: a hyperscanning functional magnetic resonance imaging study. Neuroimage 125, 401–412. doi:10.1016/j.neuroimage.2015.09.076.
Kokal, I., Engel, A., Kirschner, S., Keysers, C., 2011. Synchronized drumming enhances activity in the caudate and facilitates prosocial commitment—if the rhythm comes easily. PLoS One 6. doi:10.1371/journal.pone.0027272.
Krueger, F., Barbey, A.K., Grafman, J., 2009. The medial prefrontal cortex mediates social event knowledge. Trends Cogn. Sci. 13, 103–109. doi:10.1016/j.tics.2008.12.005.
Kühn, S., Müller, B.C.N., van Baaren, R.B., Wietzker, A., Dijksterhuis, A., Brass, M., 2010. Why do I like you when you behave like me? Neural mechanisms mediating positive consequences of observing someone being imitated. Soc. Neurosci. 5, 384–392. doi:10.1080/17470911003633750.
Lawrence, A.D., Sahakian, B.J., Rogers, R.D., Hodges, J.R., Robbins, T.W., 1999. Discrimination, reversal, and shift learning in Huntington's disease: mechanisms of impaired response selection. Neuropsychologia 37, 1359–1374. doi:10.1016/S0028-3932(99)00035-4.
Leiguarda, R., 2001. Limb apraxia: cortical or subcortical. Neuroimage 14, 137–141. doi:10.1006/nimg.2001.0833.
Mai, J.K., Majtanik, M., Paxinos, G., 2015. Atlas of the Human Brain, 4th Ed. Academic Press, Cambridge, MA.
Matsunaga, M., Kawamichi, H., Koike, T., Yoshihara, K., Yoshida, Y., Takahashi, H.K., Nakagawa, E., Sadato, N., 2016. Structural and functional associations of the rostral anterior cingulate cortex with subjective happiness. Neuroimage 134, 132–141. doi:10.1016/j.neuroimage.2016.04.020.
Mayka, M.A., Corcos, D.M., Leurgans, S.E., Vaillancourt, D.E., 2006. Three-dimensional locations and boundaries of motor and premotor cortices as defined by functional brain imaging: a meta-analysis. Neuroimage 31, 1453–1474. doi:10.1016/j.neuroimage.2006.02.004.
Moeller, S., Yacoub, E., Olman, C.A., Auerbach, E.J., Strupp, J., Harel, N., Uğurbil, K., 2010. Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. Magn. Reson. Med. 63, 1144–1153. doi:10.1002/mrm.22361.
Molnar-Szakacs, I., Uddin, L.Q., 2013. Self-processing and the default mode network: interactions with the mirror neuron system. Front. Hum. Neurosci. 7, 1–11. doi:10.3389/fnhum.2013.00571.
Mumford, D., 1992. On the computational architecture of the neocortex - II the role of cortico-cortical loops. Biol. Cybern. 66, 241–251. doi:10.1007/BF00198477.
Nadel-Brulfert, J., Baudonniere, P.M., 1982. The social function of reciprocal imitation in 2-year-old peers. Int. J. Behav. Dev. 5, 95–109.
Novembre, G., Sammler, D., Keller, P.E., 2016. Neural alpha oscillations index the balance between self-other integration and segregation in real-time joint action. Neuropsychologia 89, 414–425. doi:10.1016/j.neuropsychologia.2016.07.027.
Penhune, V.B., Steele, C.J., 2012. Parallel contributions of cerebellar, striatal and M1 mechanisms to motor sequence learning. Behav. Brain Res. 226, 579–591. doi:10.1016/j.bbr.2011.09.044.
Quiroga, R.Q., Snyder, L.H., Batista, A.P., Cui, H., Andersen, R.A., 2006. Movement intention is better predicted than attention in the posterior parietal cortex. J. Neurosci. 26, 3615–3620. doi:10.1523/jneurosci.3468-05.2006.
Rajaei, N., Aoki, N., Takahashi, H.K., Miyaoka, T., Kochiyama, T., Ohka, M., Sadato, N., Kitada, R., 2018. Brain networks underlying conscious tactile perception of textures as revealed using the velvet hand illusion. Hum. Brain Mapp. 39, 4787–4801. doi:10.1002/hbm.24323.
Rao, R.P.N., Ballard, D.H., 1999. Hierarchical predictive coding model hierarchical predictive coding of natural images. Nat. Neurosci. 2, 79.
Rissman, J., Gazzaley, A., D'Esposito, M., 2004. Measuring functional connectivity during distinct stages of a cognitive task. Neuroimage 23, 752–763. doi:10.1016/j.neuroimage.2004.06.035.
Rizzolatti, G., Craighero, L., 2004. The mirror-neuron system. Annu. Rev. Neurosci. 27, 169–192. doi:10.1146/annurev.neuro.27.070203.144230.
Rizzolatti, G., Fogassi, L., Gallese, V., 1997. Parietal cortex: from sight to action. Curr. Opin. Neurobiol. 7, 562–567. doi:10.1016/s0959-4388(97)80037-2.
Rizzolatti, G., Matelli, M., 2003. Two different streams form the dorsal visual system: anatomy and functions. Exp. Brain Res. 153, 146–157. doi:10.1007/s00221-003-1588-0.
Rizzolatti, G., Sinigaglia, C., 2016. The mirror mechanism: a basic principle of brain function. Nat. Rev. Neurosci. 17, 757–765. doi:10.1038/nrn.2016.135.
Rochat, P., 1999. Early Social Cognition: Understanding Others in the First Months of Life. Psychology Press, New York.
Ruby, P., Decety, J., 2001. Effect of subjective perspective taking during simulation of action: A PET investigation of agency. Nat. Neurosci. 4, 546–550. doi:10.1038/87510.
Rushworth, M.F.S., Nixon, P.D., Wade, D.T., Renowden, S., Passingham, R.E., 1998. The left hemisphere and the selection of learned actions. Neuropsychologia 36, 11–24. doi:10.1016/s0028-3932(97)00101-2.
Rushworth, M.F.S., Noonan, M.A.P., Boorman, E.D., Walton, M.E., Behrens, T.E., 2011. Frontal cortex and reward-guided learning and decision-making. Neuron 70, 1054–1069. doi:10.1016/j.neuron.2011.05.014.
Sasaki, A.T., Kochiyama, T., Sugiura, M., Tanabe, H.C., Sadato, N., 2012. Neural networks for action representation: a functional magnetic-resonance imaging and dynamic causal modeling study. Front. Hum. Neurosci. 6, 1–17. doi:10.3389/fnhum.2012.00236.
Schmidt, K.L., Cohn, J.F., Tian, Y., 2003. Signal characteristics of spontaneous facial expressions: automatic movement in solitary and social smiles. Biol. Psychol. 65, 49–66. doi:10.1016/s0301-0511(03)00098-x.
Searle, J., 1983. Intentionality: An Essay in the Philosophy of Mind. Cambridge University Press., Cambridge, UK.
Sirigu, A., Duhamel, J., Cohen, L., Pillon, B., Dubois, B., Agid, Y., 1996. The mental representation of hand movements after parietal cortex damage. Science 273 (80-), 1564–1568. doi:10.1126/science.273.5281.1564.
Snyder, L.H., Batista, A.P., Andersen, R.A., 1997. Coding of intention in the posterior parietal cortex. Nature 386, 167–170. doi:10.1038/386167a0.
Sperduti, M., Guionnet, S., Fossati, P., Nadel, J., 2014. Mirror neuron system and mentalizing system connect during online social interaction. Cogn. Process. 15, 307–316. doi:10.1007/s10339-014-0600-x.
Szameitat, D.P., Kreifelts, B., Alter, K., Szameitat, A.J., Sterr, A., Grodd, W., Wildgruber, D., 2010. It is not always tickling: distinct cerebral responses during perception of different laughter types. Neuroimage 53, 1264–1271. doi:10.1016/j.neuroimage.2010.06.028.
Tanaka, S.C., Doya, K., Okada, G., Ueda, K., Okamoto, Y., Yamawaki, S., 2016. Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. Nat. Neurosci. 7, 887–893. doi:10.1038/nn1279.

Tessari, A., Canessa, N., Ukmar, M., Rumiati, R.I., 2007. Neuropsychological evidence for a strategic control of multiple routes in imitation. Brain 130, 1111–1126. doi:10.1093/brain/awm003.

Tognoli, E., Lagarde, J., de Guzman, G.C., Kelso, J.A.S., 2007. The phi complex as a neuromarker of human social coordination. Proc. Natl. Acad. Sci. U. S. A. 104, 8190–8195. doi:10.1073/pnas.0611453104.

Trevarthen, C., 1979. Communication and cooperation in early infancy. In: Bullowa, M. (Ed.), Before Speech: The Beginning of Human Communication. Cambridge University Press, London, pp. 321–347.

Uddin, L.Q., Molnar-Szakacs, I., Zaidel, E., Iacoboni, M., 2006. rTMS to the right inferior parietal lobule disrupts self-other discrimination. Soc. Cogn. Affect. Neurosci. 1, 65–71. doi:10.1093/scan/nsl003.

van Baaren, R.B., Holland, R.W., Kawakami, K., van Knippenberg, A., 2004. Mimicry and prosocial behavior. Psychol. Sci. 15, 71–74. doi:10.1111/j.0963-7214.2004.01501012.x.

van Baaren, R.B., Holland, R.W., Steenaert, B., van Knippenberg, A., 2003. Mimicry for money: behavioral consequences of imitation. J. Exp. Soc. Psychol. 39, 393–398. doi:10.1016/S0022-1031(03)00014-3.

van Overwalle, F., Baetens, K., 2009. Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. Neuroimage 48, 564–584. doi:10.1016/j.neuroimage.2009.06.009.

Wager, T.D., Nichols, T.E., 2003. Optimization of experimental design in fMRI: a general framework using a genetic algorithm. Neuroimage 18, 293–309. doi:10.1016/S1053-8119(02)00046-0.

Wheaton, L.A., Hallett, M., 2007. Ideomotor apraxia: a review. J. Neurol. Sci. 260, 1–10. doi:10.1016/j.jns.2007.04.014.

Wolpert, D.M., Ghahramani, Z., Jordan, M.I., 1995. An internal model for sensorimotor integration. Science 269 (80-), 1880–1882.

Wolpert, D.M., Kawato, M., 1998. Multiple paired forward and inverse models for motor control. Neural Networks 11, 1317–1329. doi:10.1016/s0893-6080(98)00066-5.

Wolpert, D.M., Miall, R.C., Kawato, M., 1998. Internal models in the cerebellum. Trends Cogn. Sci. 2, 338–347. doi:10.1016/S1364-6613(98)01221-2.

Zadikoff, C., Lang, A.E., 2005. Apraxia in movement disorders. Brain 128, 1480–1497. doi:10.1093/brain/awh560.