BRAIN CONNECTIVITY
Volume 00, Number 00, 2024

DOI: 10.1089/brain.2024.0032



# Frontal Cortex Acts as Causality Transition Hub from Mirror Network to Mentalizing Network During Action Intention Understanding

Li Zhang,[1] Lei Zhang,[1,2] Jing Wang,[3] and Yanmei Zhu[1,4]



## Abstract

***Introduction:*** While understanding other's action intention, mirror and mentalizing systems of human brain are successively activated in action perception and intention inference processes.
***Methods:*** To reveal the relationship between mirror and mentalizing systems during the two stages, this electroencephalogram study adopted the method of time-varying orthogonalized partial directed coherence (OPDC) to assess causal interaction between mirror and mentalizing networks during a "hand-cup interaction" action intention understanding task.
***Results:*** Task-related causal connectivity was found in gamma frequency band (30–45 Hz), primarily manifested as directed edges from sensorimotor to frontal areas in poststimulus 400–600 ms interval and directed links from frontal to parietal and temporal regions in 600–800 ms period. The analysis of event-related potential and source currents suggests that the change of inter-regional causality is related with functional transition of the brain from mirror matching to intention inference. The OPDC network modeling further finds that frontal area contains more inflow nodes in mirror network, whereas more outflow nodes in mentalizing network, with high betweenness centrality in temporally changing functional communities. Compared with intention-oriented actions, identification of unintelligible action intention particularly induces stronger OPDC from right superior frontal to inferior frontal gyrus and from sensorimotor to right frontotemporal regions during mentalizing inference process.
***Conclusion:*** These findings collectively suggest that, in the time ordering of information transfer within the directed networks, frontal area plays an important role of bridging hub between mirror and mentalizing systems, from maintaining and supervising perceptual information for mirror matching to controlling the mentalizing process for decoding other's action intention.

**Keywords:** action intention understanding; effective connectivity; mentalizing network; mirror network; orthogonalized partial directed coherence

## Impact Statement

From the perspective of neural mechanism of action intention understanding, this study extends previous research to decoding dynamic fluctuations in the brain network structure related to continuous cognitive subprocesses. In the field of human–machine interaction, the electroencephalogram (EEG) features extracted from this study for recognizing different types of action intention have potential application value for medical rehabilitation, such as motor dysfunction caused by stroke, spinal cord injury, and so on, by monitoring and analyzing patients' EEG signals. In addition, the EEG features can also be applied to controlling prosthetic limbs, exoskeletons, and other assistive devices to help patients restore or improve movement.

[1]School of Early-Childhood Education, Nanjing Xiaozhuang University, Nanjing, PR China.
[2]Department of Rheumatology and Immunology, Jinling Hospital, Affiliated Hospital of Medical School, Nanjing University, Nanjing, PR China.
[3]School of Computer Science and Information Technology, Xinyang Normal University, Xinyang, PR China.
[4]Ministry of Education, Key Laboratory of Child Development and Learning Science (Southeast University), Nanjing, PR China.

## Introduction

Action intention understanding is an essential ability of interpersonal communication for the human living in the society, which helps people better judge other's mental activity and motivation, so as to predict and explain their behavior (Oberman et al., 2007). The brain response of action intention understanding includes two crucial neurocognitive subprocesses, that is, recognition of action kinematics and inference of intrinsic motivation. In the two stages, the brain successively recruits mirror and mentalizing systems (Atique et al., 2011; Becchio et al., 2012; Brandone and Stout, 2023; Brass et al., 2007; Catmur, 2015; Ge et al., 2017; Oberman et al., 2007; Woodward and Gerson, 2014). Depending on low-level perceptual input, the mirror system mainly consisting of anterior intraparietal sulcus and premotor cortex is activated to map the visual information of actions onto memory representation (Buccino et al., 2007; Cheng et al., 2007; Rizzolatti et al., 2001). At the following stage of intention inference, the mentalizing system is activated to speculate other's mental state or intrinsic motivation according to perceived actions, which is primarily composed of temporoparietal junction (TPJ), medial prefrontal cortex (PFC), and precuneus (Blakemore and Decety, 2001; Catmur, 2015; Ge et al., 2019; Liew et al., 2011; Spunt et al., 2011; Van Overwalle and Baetens, 2009).

In previous neuroimaging studies, the concurrent activation of mirror and mentalizing areas was rarely detected (Van Overwalle and Baetens, 2009). However, the mirror system should provide rapid and intuitive input of perceptual information to the mentalizing areas to support inferential process for making correct judgment (Catmur, 2015; Gardner et al., 2015; Tidoni et al., 2013). Thus, the relevance between mirror matching and intention inference is still unclear, especially the role of mirror system in inferential process while the mentalizing network is being recruited. Past studies have found that the activity of mirror and mentalizing systems depends on the type of observed actions (Centelles et al., 2011; Ge et al., 2018; Liew et al., 2011; Van Overwalle and Baetens, 2009). While observing familiar or intention-oriented action, mirror neurons of the brain might directly respond to being aware of the goal of other's actions, without the need of initiating the mentalizing system. If the observed action is unusual or contextual information is absent, the mirror system might be inactive due to the lack of matchable action information, but the mentalizing area might be strongly motivated to fill in the "missing" information. The alternate activation of brain areas subserving one system and the areas subserving the other system was described as distinct but complementary function of mirror and mentalizing systems (Catmur, 2015; Chiavarino et al., 2012; De Lange et al., 2008; Cheng, 2007; Gardner et al., 2015; Zhang et al., 2018).

Dynamic effective connectivity provides a new insight into temporal ordering of causal relation between the participating areas of a brain network, which primarily concerns on inferring the directions of neural interactions and information flow directly from data (Friston et al., 2013). For further estimating how one neural system exerts influence over another during mirror matching and intention inference processes, this study conducted an analysis of time-varying causal connectivity for the electroencephalogram (EEG) signals recorded from a "hand-cup interaction" action intention understanding task. The timing of mirror responses and mentalizing processing was determined by significant event-related potential (ERP) components. In the effective connectivity network modeling, orthogonalized partial directed coherence (OPDC) method was applied to gamma-band (30–45 Hz) EEG signals to construct inter-regional causal relationship and minimize the effect of mutual sources within the brain (Omidvarnia et al., 2014). Gamma response has been found to be strongly interwoven with a variety of sensory modalities (vision, sensation, and audition), attention, perception, high-order cognition, working memory load, basic motor response, and so on (Gaetz et al., 2013; Herrmann et al., 2010; Kumar et al., 2023; Vezoli et al., 2013). Besides, gamma-specific effective connectivity pattern between cortical areas has been found in memory-related brain function (Bayazi et al., 2021). Therefore, gamma-band OPDC can be considered as a key neural signature of information processing of action intention understanding in the human brain. Based on graph-theoretical analysis of time-sequential OPDC networks, spatiotemporal EEG features were further extracted and used for recognizing different types of action intention. According to changing causal flows over the task course, the conditions under which the mirror area affects mentalizing area and crucial hubs of information transfer in the transition from mirror network to mentalizing network were analyzed and discussed.

## Materials and Methods

### *EEG experiment and data preprocessing*

Subjects. The EEG experiment was approved by the Academic Committee of the Research Center for Learning Science, Southeast University, China. There were 30 college students recruited to perform the "hand-cup interaction" action intention understanding task. Exclusion criteria of subjects included left handedness, medical, neurological, or psychiatric illness, and history of brain injury or surgery. Each subject signed a fully informed consent form about the experiment and received the payment for the participation. Through further removing the signals seriously contaminated by artifacts and noise, effective EEG data from 24 subjects, including 10 males and 14 females, aged $22.4 \pm 2.3$ (mean $\pm$ standard deviation) were retained finally.

Experimental paradigm. The experimental task comprised three conditions in which the actions presented in the stimuli pictures with different intention types, including two intention-oriented actions quite usual in our life, that is, grasping a cup for using it (Ug) and grasping a cup for transporting it (Tg), and an unintelligible action, that is, simple contact (Sc) without a clear intention (Fig. 1A). The task was composed of 294 trials, in which each condition included 98 trials. For the trials with the same hand motion, the cup was marked with one of seven colors with equal probability; therefore, each color of cup was presented 14 times in each condition. The trials from the three conditions were cross-presented, with a constrain that consecutive presentation of the same action was less than four times, as well as the same color of cup.

At the beginning of a trial, the symbol "+" was presented at the center of screen for 150 ms. Then a cup was shown for 500 ms. After that, the screen presented a hand interacting

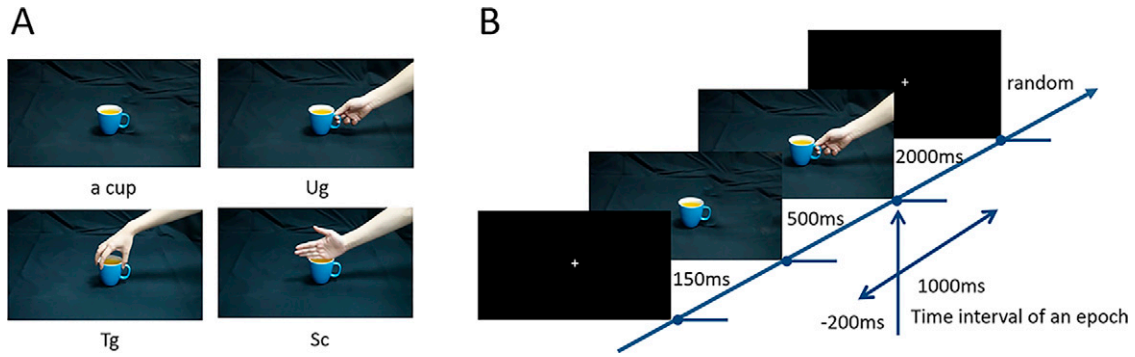


**FIG. 1.** Experimental paradigm of "hand-cup interaction" action intention understanding task. **(A)** Stimulus materials: a cup at the center of screen; a hand grasping a cup for using it (Ug); a hand grasping a cup for transporting it (Tg); and simple contact (Sc). **(B)** Timeline of stimulus presentation and time interval of an epoch of EEG data. EEG, electroencephalogram.

with the cup for 2000 ms. In the meantime, the subjects were asked to judge the intention of the action performer just in their brain without the need of pressing any button. Then, the symbol "+" was presented with a random time length at the end of the trial (Fig. 1B). The total time for an experiment execution was about 24 min.

EEG recording and preprocessing. An international 10–20 system of neuroscan was used to record EEG signals with sampling rate at 500 Hz, in which 60 electrodes were set to cover frontal, parietal, temporal, and occipital regions. In addition, two reference electrodes were placed on bilateral mastoids to provide reference signals, and four surface electrodes were used to simultaneously record electrooculographic (EOG) signals, with two electrodes placed on higher and lower eyelids and the other two electrodes placed 1 cm lateral to the outer corner of the left and right orbits.

The raw EEG signals were preprocessed by Scan 4.3 procedure. At first, the EEG signals were band-pass filtered into 1–60 Hz range. Then, each trial was segmented into a 1200 ms epoch, including 200 ms prestimulus and 1000 ms poststimulus intervals, and the baseline correction was conducted for the prestimulus time interval. Through visual inspection, the trials contaminated by the artifacts, such as subject's movement and sweat, were rejected first. After that, ocular artifacts, including ocular movements and eye blinks, were cleared according to the simultaneously recorded EOG signals. The trials contaminated by electromyogram and electrocardiogram noises were rejected with a threshold set in the range of 50–75 $\mu$v. Furthermore, other residual artifacts, such as the components of ocular and muscle movements, were further removed through the independent component analysis running in a trial-by-trial manner in the EEGLAB toolbox.

After the preprocessing procedure, 1146, 1174, and 1139 trials were retained for Ug, Tg, and Sc conditions, respectively, in which 36–68 trials were retained for each subject. By concatenating the trials from all subjects for each task condition, this study constructed three sets in which the size of each set equals (number of channels) × (number of time points) × (number of trials).

### *ERP and source current analyses*

An ERP analysis was conducted to determine the timing of psychological processing in the task course. The EEG signals from midline electrodes, including Fz (frontal cortex), FCz (frontal-central sulcus), and Cz (central sulcus), were filtered into 1–30 Hz and group-averaged for further comparison among task conditions.

Besides, the source currents of the ERPs were estimated using a cortical source estimation procedure implemented in the Brainstorm toolbox (http://neuroimage.usc.edu/brainstorm; Tadel et al., 2011). In this procedure, the EEG signals were assumed to be generated from a block of electric dipoles at the cortical surface. The noise of the scalp sensors was removed through computing the noise covariance matrix of the signals in the prestimulus interval. Cortical currents were estimated through an inverse kernel matrix, which was produced from the forward model constructed by OpenMEEG software (http://www-sop.inria.fr/athena/software/ OpenMEEG/; Gramfort et al., 2010). For the group-averaged EEG data from each condition, dipole modeling method was adopted that adjusted the parameters of a single current dipole fitted to the sensor data, to realize source current transformation at each point in time. As a result, in each 200 ms time interval of the task, a linearly constrained minimum variance-type map was produced, and an equivalent current dipole was fitted at the strongest peak location of that map.

### *Time–frequency analysis of OPDC*

Orthogonalized partial directed coherence. Directed interactions over time and frequency between EEG signals were measured by OPDC, which is an orthogonalized version of classical partial directed coherence (PDC) usually used to detect causal influences in multivariate stochastic systems (Huang et al., 2016; Omidvarnia et al., 2014; Zhang et al., 2020). The classical PDC can provide a measure in frequency domain on the basis of Granger causality, through modeling time signals by multivariate autoregressive (MVAR) processes. For a time series $y(t)$ with $L$ number of samples $(t = 1, \ldots\ldots, L)$, a strictly causal MVAR model of order $p$ with dimension $M$ (Here $M$ is the number of channels) is defined as

$$\begin{bmatrix} y_1{}^{(t)} \\ \cdots \\ y_M{}^{(t)} \end{bmatrix} = \sum_{r=1}^{p} A_r \begin{bmatrix} y_1{}^{(t-r)} \\ \cdots \\ y_M{}^{(t-r)} \end{bmatrix} + \begin{bmatrix} \omega_1{}^{(t)} \\ \cdots \\ \omega_M{}^{(t)} \end{bmatrix} \quad (1)$$

where $A_1, A_2 \ldots A_p$ are $M \times M$ coefficient matrices,

$$A_r = \begin{bmatrix} a_{11}^r \dots a_{1M}^r \\ \dots \\ a_{M1}^r \dots a_{MM}^r \end{bmatrix}, (r = 1, \dots, p) \tag{2}$$

and $[\omega_1 \dots \omega_M]^T = \omega$ is a normally distributed real-valued zero-mean white noise vector with diagonal covariance matrix $\Sigma_w = \langle ww^T \rangle = diag\{\lambda_{kk}^2\}$. Here, $\langle \cdot \rangle$ is the expected value operator. In $A_r$, the real-valued parameter $a_{kl}^r$ reflects the linear relationship between channels $k$ and $l$ at the delay $r$.

Based on time-varying version of the model in Equation (1) in which matrices $A_r(t)$ are time varying, the time-varying PDC from signal $l$ to signal $k$ at frequency $f$ is described as

$$PDC_{kl}(t,f) = \pi_{kl}(t,f) \triangleq \frac{|A_{kl}(t,f)|}{\sqrt{a_l^H(t,f)a_l(t,f)}} \tag{3}$$

Here, $A_{kl}(t,f)$ refers to the $kl$th element of $A(t,f)$, and $a_l^H$ represents the Hermitian transpose of the vector $a_l$. The measure $\pi_{kl}(t,f)$ takes values between 0 and 1, which is directional, that is, $\pi_{kl}(t,f)$ is not equal to $\pi_{lk}(t,f)$.

EEG-based interactive relationship analysis faces with a significant challenge of volume conduction (i.e., several signals are often produced by a common brain source), which might lead to false connectivity. Hipp et al. (2012) suggest that the orthogonal parts can be extracted by orthogonalizing power envelope of surface EEG signals in the complex domain, to reduce the covariability. The OPDC was thus proposed to be used in directed connectivity analysis through combining time-varying PDC analysis, orthogonalization, and imaginary part of coherence function. Notably, the main idea of OPDC is to alleviate the effect of mutual sources at the level of MVAR coefficients, instead of performing the orthogonalization process at the amplitude level. In the computation of OPDC, EEG signals are supposed to be generated through a linear superposition of $K$ independent source signals. In the frequency domain, the transformation relationship can be described as

$$Y_i(f) = \sum_{k=1}^{K} v_{ik} S_k(f) \tag{4}$$

This means that multichannel EEG signals $Y(f) = VS(f)$. Here, $Y(f) \in C^M$, $S(f) \in C^K$ represents multivariate source signal, and $V \in R^{M \times K}$ refers to the matrix containing all source weights. By fitting a strictly causal MVAR mode on the multichannel EEG signal $y(t)$ in the time domain, the Equation (1) can be transformed into the frequency domain as follows:

$$Y(f) = \sum_{r=1}^{p} A_r e^{-j2\pi fr} Y(f) + W(f) = B(f)Y(f) + W(f) \tag{5}$$

where $B(f) = \sum_{r=1}^{p} A_r e^{-j2\pi fr}$ and $B_{kl}(f) = \sum_{r=1}^{p} a_{kl}^r e^{-j2\pi fr}$.

According to Equation (5), $Y(f) = B(f)VS(f) + W(f)$. The cross-spectral density matrix $C(f)$ of $Y(f)$ can be acquired by

$$\begin{aligned} C(f) &= \langle Y(f)Y^H(f) \rangle = \langle (B(f)VS(f) + W(f)) \\ &\times \left(S^H(f)V^H B^H(f) + W^H(f)\right) \rangle \end{aligned} \tag{6}$$

Between $Y_i(f)$ and $Y_j(f)$, the cross-spectral density function, $C_{ij}(f)$, can be computed according to

$$\begin{aligned} C_{ij}(f) = \Bigg\langle \sum_{n_1=1}^{M} \sum_{n_2=1}^{M} \sum_{k_1=1}^{K} \sum_{k_2=1}^{K} B_{in_1}(f) B_{jn_2}^*(f) v_{n_1k_1} v_{n_2k_2} \\ \times S_{k_1}(f) S_{k_2}^*(f) \Bigg\rangle + \langle W_i(f) W_j^*(f) \rangle \end{aligned} \tag{7}$$

In Equation (7), $S_{k_1}(f)$ and $S_{k_2}(f)$ are independent; therefore, all terms containing $\langle S_{k_1}(f) S_{k_2}^*(f) \rangle$, $k_1 \neq k_2$ are 0, which makes Equation (7) be transformed into

$$\begin{aligned} C_{ij}(f) = \Bigg\langle \sum_{n_1=1}^{M} \sum_{n_2=1}^{M} \sum_{k=1}^{K} B_{in_1}(f) B_{jn_2}^*(f) v_{n_1k_1} v_{n_2k_2} |S_k(f)|^2 \Bigg\rangle \\ + \langle W_i(f) W_j^*(f) \rangle \end{aligned} \tag{8}$$

The imaginary part of $C_{ij}(f)$ can be described as follows:

$$\begin{aligned} \mathrm{Imag}\{C_{ij}(f)\} = \sum_{n_1=1}^{M} \sum_{n_2=1}^{M} \sum_{k=1}^{K} \\ \times \{v_{n_1k_1} v_{n_2k_2} |S_k(f)|^2 \mathrm{Imag}\{B_{in_1}(f) B_{jn_2}^*(f)\}\} \end{aligned} \tag{9}$$

Since the effect of mutual sources is excluded, true interaction between EEG signals can be reflected by $\mathrm{Imag}\{B_{in_1}(f) B_{jn_2}^*(f)\}$, that is,

$$\begin{aligned} \mathrm{Imag}\{B_{in_1}(f) B_{jn_2}^*(f)\} &= \sum_{r_1=1}^{p} \sum_{r_2=1}^{p} a_{in_1}^{r_1} a_{jn_2}^{r_2} sin\left(2\pi f(r_1 - r_2)\right) \\ &= \sum_{r_1=1}^{p} \sum_{r_2=1}^{p} a_{in_1}^{r_1} a_{jn_2}^{r_2} sin\left(2\pi f(r_1)\right) cos\left(2\pi f(r_2)\right) \\ &+ \sum_{r_1=1}^{p} \sum_{r_2=1}^{p} a_{in_1}^{r_1} a_{jn_2}^{r_2} sin\left(2\pi f(r_2)\right) cos\left(2\pi f(r_1)\right) \end{aligned} \tag{10}$$

When the mutual sources have been excluded, the orthogonalized components of $a_{kl}^r e^{-j2\pi fr}, k = 1, \dots, M, \; l = 1, \dots, M$ at different delays, including the real part $a_{kl}^r cos(2\pi fr)$ and the imaginary part $a_{kl}^r sin(2\pi fr)$, play a salient role in estimating the true connectivity between channels.

Based on this principle, the time-varying OPDC strength of the interaction of a given signal $l$ with regard to signal $k$ can be represented by

$$\Psi_{kl}(t,f) \triangleq \frac{\left| \sum_{r_1=1}^{p} \sum_{r_2=1}^{p} a_{kl}^{r1}(t) a_{kl}^{r2}(t) cos(2\pi fr_1) sin(2\pi fr_2) \right|}{a_l^H(t,f) a_l(t,f)} \tag{11}$$

$$= \frac{\left|\sum_{r_1=1}^{p}\sum_{r_2=1}^{p} Real\{a_{kl}^{r1}(t)e^{-i2\pi fr_1}\} Imag\{a_{kl}^{r2}(t)e^{-i2\pi fr_2}\}\right|}{a_l^H(t,f)a_l(t,f)} \quad (12)$$

In Equation (11), the sum of weighted sine and cosine terms puts on a trend varying appearance to the OPDC measure along the frequency axis. It can be seen that Equations (11) and (12) are equivalent to the following decomposition of $\pi_{kl}(t,f)$, that is,

$$\Psi_{kl}(t,f) = \frac{|Real\{A_{kl}(t,f)\}|}{\sqrt{a_l^H(t,f)a_l(t,f)}} \cdot \frac{|Imag\{A_{kl}(t,f)\}|}{\sqrt{a_l^H(t,f)a_l(t,f)}} \quad \text{if } k \neq l \quad (13)$$

In Equation (13), each factor ranges from 0 to $\pi_{kl}(t,f)$; thus, the measure $\Psi_{kl}(t,f)$ will always take values between 0 and 1.

Time-frequency detection of OPDC change. According to the definition, the OPDC analysis was conducted to measure inter-regional directed interactions over time and frequency of the EEG signals. The EEG channels located at frontal and central sensorimotor regions were selected to calculate OPDC between pairwise signals. The dual extended Kalman filter was adopted to estimate time-varying MVAR parameters. A statistical testing procedure was applied on the data from all the subjects to obtain time-resolved OPDC that quantifies the directed influences between selected channels. In this procedure, the prestimulus 200 ms period was used to construct the null distribution as being considered as baseline. The time-varying OPDC measures were extracted from the poststimulus task period of each trial and compared with a distribution of the same measures extracted from the baseline intervals of each trial. The highest score at the 99th percentile of the distribution of time-frequency bin over prestimulus epochs is computed, which produces a thresholding plane. To find the time-frequency bins with significant values over the baseline interval, a time-frequency threshold was applied to each trial. As a result, average of the thresholded OPDC was computed as the mean connectivity representation between channels.

### *Construction of time-sequential causal networks*

To identify how mirror matching and intention inference modulate inter-regional influence, OPDC-based effective connectivity networks were constructed in task-related time windows and frequency band. Within a brain functional network, inter-regional interactions present a quasi-stable state (about 100–200 ms), which makes functional network flexibly adapt and continuously reorganize from one coordination system to another (Baker et al., 2014; Bola and Sabel, 2015). Besides, each time span containing significant ERP component is about 200 ms in this study. Considering the dynamic fluctuation of connectivity affected by the duration of inter-regional interactions and the ERP response interval, an equal time segment was set for EEG data to construct time-sequential directed networks. Accordingly, each trial was decomposed into one prestimulus interval and five continuous poststimulus time windows, that is, 0–200 ms (TW1), 200–400 ms (TW2), 400–600 ms (TW3), 600–800 ms (TW4), and 800–1000 ms (TW5) intervals.

OPDC-based directed network construction. Through further time-frequency detection for OPDC change with respect to baseline, EEG signals were further filtered into 30–45 Hz gamma frequency band. For 60-channel EEG signals, each channel serves as a "node" in a directed network. In every 200 ms time window of a trial, OPDC was calculated for pairwise EEG channels, resulting in a $60\times60$ association matrix of OPDC measures with asymmetry characteristic. Thus, a sequence composed of six continuously evolving OPDC association matrices was formed for each trial.

For every OPDC association matrix, an adjacent matrix (i.e., a directed network) can be generated by setting a threshold following the two criteria: (1) For emphasizing local and regional variations between the topological structures of directed networks organized for different cognitive functions or different cognitive conditions, a fixed connection density was set for all OPDC matrices to get the adjacent matrices, according to Erdos–Renyi model (Erdős and Rényi, 1964), that is, $p_{den} \approx 2ln(n)/n$, where $n$ is the number of nodes in a network; and (2) while tracing the change of connectivity strength of a network along the time course, a fixed threshold was set for time-sequential OPDC matrices. The threshold is the minimum OPDC measure of the adjacency matrices in prestimulus baseline period with a fixed connection density abiding by criterion (1).

Local and global matrices of directed graphs. Graph-theoretical analysis was performed to quantify the topological properties of OPDC-directed networks (Bullmore and Sporns, 2009). In the dynamic reorganization of directed networks, the change in local specialization was assessed by nodal degree, nodal strength, and betweenness centrality, and global integration was quantified by modularity. For the directed graphs of OPDC networks, the Louvain algorithm with rapid convergence properties and high hierarchical partitioning was adopted for community detection (Blondel et al., 2008). According to the definition from graph theory, $N$ is the set of all the nodes in a network, and $(i \rightarrow j)$ represents the directed link from node $i$ to node $j$, $(i,j \in N; i \neq j)$. If there is directed connection status from node $i$ to node $j$, $b_{i\rightarrow j} = 1$; otherwise, $b_{i\rightarrow j} = 0$.

Nodal degree is the number of links connected to the node. For a directed network, the indegree is the number of inward links, and the outdegree refers to the number of outward links (Opsahl et al., 2010).

$$d_{i(in)} = \sum_{j\in N, i\neq j} b_{j\rightarrow i} \quad (14)$$

$$d_{i(out)} = \sum_{j\in N, i\neq j} b_{i\rightarrow j} \quad (15)$$

In a directed network, a node with high outdegree centrality can be viewed as a "cause" hub, which reflects the "outflow" in information transfer. On the contrary, a node with high indegree centrality serves as an "effect" hub, indicating the "inflow" in inter-regional information communication (Seth, 2010; Sporns et al., 2007).

Nodal strength is given by the sum of the OPDC measures of the weighted adjacent edges connected to the node. For node $i$, the inflow strength is the sum of OPDC measures of

inward links and the outflow strength refers to that of outward links.

$$str_{i(in)} = \sum_{j \in N, i \neq j} b_{j \to i} \times OPDC_{ji} \tag{16}$$

$$str_{i(out)} = \sum_{j \in N, i \neq j} b_{i \to j} \times OPDC_{ij} \tag{17}$$

Betweenness centrality of nodes was computed to assess the communication role of a node within a network, which is defined as the fraction of all shortest paths in the network that pass through a given node $i$.

$$bc_i = \frac{1}{(n-1)(n-2)} \sum_{\substack{h,j \in N \\ h \neq j, h \neq i, j \neq i}} \frac{\rho_{hj}(i)}{\rho_{hj}} \tag{18}$$

where $\rho_{hj}$ is the number of the shortest paths between nodes $h$ and $j$, and $\rho_{hj}(i)$ is the number of the shortest paths between nodes $h$ and $j$ that pass through node $i$.

To measure the strength of division of a directed network into modules, modularity of the network can be computed according to

$$Q = \sum_{u \in G} \left[ e_{uu} - \left( \sum_{v \in G} e_{uv} \right)^2 \right] \tag{19}$$

where $G$ is a set of nonoverlapping modules that the network can be completely divided into, and $e_{uv}$ is determined by the ratio of the number of connections that link the nodes in module $u$ with the nodes in module $v$ to the total number of the connections in the network.

### *Cosine similarity of nodal degree between OPDC networks*

For each trial, six time-sequential directed networks were produced along the task course. According to the definition of nodal degree, an indegree vector and an outdegree vector with 60 dimensions (the number of nodes) were extracted from a directed graph. Degree distribution of all nodes in a network is an important indicator of network construction and resilience. Complex brain network has a characteristic of "heavy-tailed" degree distribution, that is, a small number of nodes with numerous connections (hub nodes), but most of the nodes with few connections (van den Heuvel and Sporns 2013). Therefore, nodal indegree and outdegree in a directed network were thought to ideally represent local maxima of "inflow" and "outflow" connections in global network. For the indegree (resp. outdegree) vectors of two OPDC networks, cosine similarity can be used to estimate topological similarity in localization of inflow (resp. outflow) hubs, by measuring the consistency of directions between two vectors.

Given two nodal degree vectors, $E$ and $F$, extracted from two networks, their cosine similarity is defined as

$$similarity = cos(\theta) = \frac{E \cdot F}{||E||\,||F||} = \frac{\sum_{i=1}^{n} E_i \times F_i}{\sqrt{\sum_{i=1}^{n} (E_i)^2} \times \sqrt{\sum_{i=1}^{n} (F_i)^2}} \tag{20}$$

Here, $E_i$ and $F_i$ represent the elements of vectors $E$ and $F$, respectively. The cosine value of the angle between the two vectors determines whether they roughly point in the same direction. When two vectors have the same direction, the cosine similarity value is 1; when the angle between two vectors is 90°, the cosine similarity value is 0; and when two vectors point in completely opposite directions, the cosine similarity value is $-1$.

For detecting the hub reorganization of directed networks along the timeline, cosine similarity was calculated between the nodal indegree (resp. outdegree) vectors of every two time-sequential directed graphs from averaged adjacent matrices of each subject. Besides, the last OPDC network formed in TW5 was compared with that in the baseline period by the cosine similarity test, to reveal whether the dynamic network at the end of a trial was returned to the level of resting-state network.

### *Statistical and discriminant analyses for spatiotemporal EEG features*

First, in each time window with ERP evoked compared with baseline, the amplitudes of single-trial evoked potentials were statistically compared among the three conditions by one-way analysis of variance (ANOVA). Moreover, in each time window, internodal OPDC measures (i.e., elements of the association matrices) were compared one by one between task conditions by the ANOVA, to isolate the direction and strength of information flows related to understanding other's intention-oriented and unintelligible actions. A false discovery rate procedure was conducted to correct for the multiple hypothesis testing, with significance level set to 0.05. The null hypothesis is that the difference between conditions is zero.

Besides, EEG features of the topological hub reorganization from mirror network to mentalizing network were extracted according to the significant differences in the OPDC networks. As nodes with high outdegree and indegree represent the "cause" and "effect" hubs, the changes of outflow strength and inflow strength of nodes at crucial brain areas from mirror response to mentalizing process constitute two-dimensional input features. For each subject, single-trial feature samples were recognized individually by different classification algorithms with 10-fold cross validation, including linear discriminant analysis, support vector machine, and Naive Bayes, to reveal functional transition of the nodes in dynamic network reorganization and determine distinguishable EEG channels and features for understanding other's different types of action intentions.

## Results

### *ERP responses and cortical current distribution*

From the ongoing waveforms in Ug, Tg, and Sc conditions, it can be found that, in poststimulus period, the five consecutive 200 ms time windows well contain the span of important ERP components (Fig. 2). Significant difference among the amplitudes of ERPs from Ug, Tg, and Sc conditions can be first found in N170 (TW1) related to motion component of visual stimulus. Besides, P3b in TW3 and the late positive potential in TW4 also show significant differences among the three conditions (Table 1).

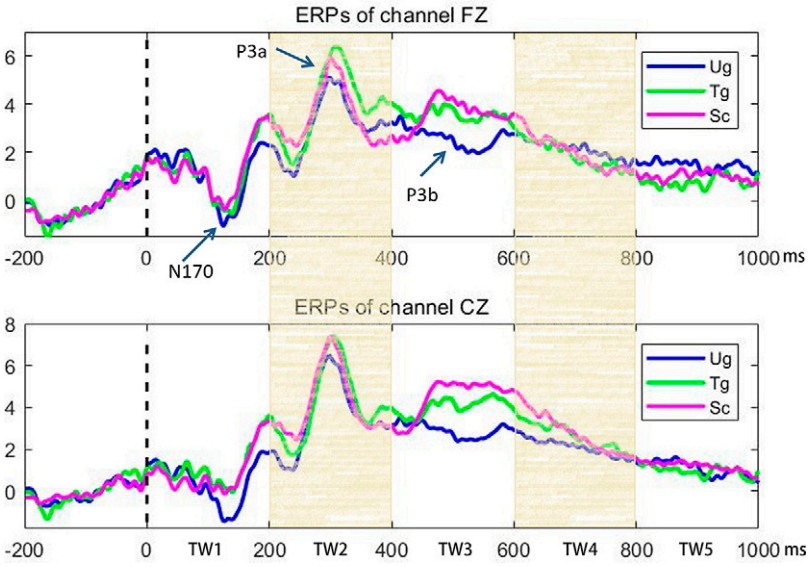


**FIG. 2.** Task-elicited potential change over time course. Grand average of ERPs for Ug, Tg, and Sc conditions from EEG channels **(A)** Fz at frontal area and **(B)** Cz at central sulcus area. Time = 0 corresponds to the onset of "hand-cup interaction" presentation. The figure shows that each condition has elicited significant ERP components N170, P3a, and P3b. The blue, green, and purple solid lines represent the Ug, Tg, and Sc conditions, respectively. ERP, event-related potential; EEG, electroencephalogram; Fz, frontal cortex; Cz, central sulcus.



It is notable that the source currents in TW3 interval under the three conditions are distributed on sensorimotor area and extended to posterior parietal cortices, which are the crucial participating components of the mirror system (Van Overwalle and Baetens 2009). In TW4 interval, the source distribution is changed to TPJ and the PFC, that is, the primary components of mentalizing system (Bonini et al., 2010; Van Overwalle and Baetens 2009) (Fig. 3). Although the three conditions elicit spatially similar cortical current distribution, understanding intention-oriented actions in Ug and Tg conditions induces higher activity in TW3, but less cortical activation in TW4 interval, compared with unintelligible action in Sc condition. The variation in cortical activation strength affected by action intention types well coincides with the complementary effect of mirror response and mentalizing function (Catmur, 2015; De Lange et al., 2008; Decety, 2007; Gardner et al., 2015). Combining the psychological mechanism of the ERP components and cortical source localization, TW3 is supposed to be highly involved in mirror matching of other's action onto one's memory system and TW4 is very likely the time interval comprising the mentalizing process.

### *Interchannel OPDC changes in gamma frequency band*

Corresponding to the change of cortical currents, the OPDC in the frequencies around 30–45 Hz presents time-varying and complementary causal interactions between frontal and sensorimotor areas. As shown in Figure 4, interchannel OPDC significantly occurs from sensorimotor (channels C5, C3, and C1) to frontal (channels FPz, Fz, and F1) regions in TW1, TW2, and TW3. Subsequently, the directions of causal interaction are reversed, manifested as directed links from frontal (channels FPz, Fz, and F1) to sensorimotor regions (channels C5, C3, and C1). Compared with the time-frequency OPDC measures in Ug and Tg conditions, Sc condition induces lower OPDC from sensorimotor to frontal regions in TW3 (400–600 ms) interval, but higher OPDC from frontal to sensorimotor regions in the later TW4 (600–800 ms) interval. The complementary effect of interchannel OPDC provides further evidence for the timing of mirror response and intention inference.

### *Task effects on spatiotemporal evolution of OPDC networks*

Except the network formed in TW4 interval, directed networks in prestimulus, TW1, TW2, TW3, and TW5 intervals show highly similar topological architectures, primarily manifested as directed edges from sensorimotor area to frontal area (Fig. 5). The result means that, in the five time intervals, central sensorimotor regions play the role of "cause" hub of directed information flow, whose "effect" is exerted on the frontal area. Notably, the OPDC networks in TW4 show significantly different topological structures compared with those in other five time windows. The nodes at frontal regions particularly exhibit higher outdegree in the directed networks in TW4, indicating that the frontal cortex plays the

TABLE 1. SIGNIFICANCE VALUES (*P*) IN ANOVA TESTS FOR AMPLITUDES OF ERPS OF MIDLINE CHANNELS AMONG UG, TG, AND SC CONDITIONS

| *Channel time window* | *TW1 (N170)* | *TW2 (P3a)* | *TW3 (P3b)* | *TW4* | *TW5* |
|---|---|---|---|---|---|
| Fz | 0.0083** | 0.4572 | 0.0085** | 0.0487* | 0.0952 |
| FCz | 0.0018** | 0.6725 | 0.0108* | 0.0176* | 0.1030 |
| Cz | 0.0014** | 0.3888 | 0.0402* | 0.0121* | 0.1494 |

*p*-Significance value; * represents $p < 0.05$ and ** denotes $p < 0.01$.
ANOVA, analysis of variance; ERP, event-related potential; TW, time window; Fz, frontal cortex; FCz, frontal–central sulcus; Cz, central sulcus.

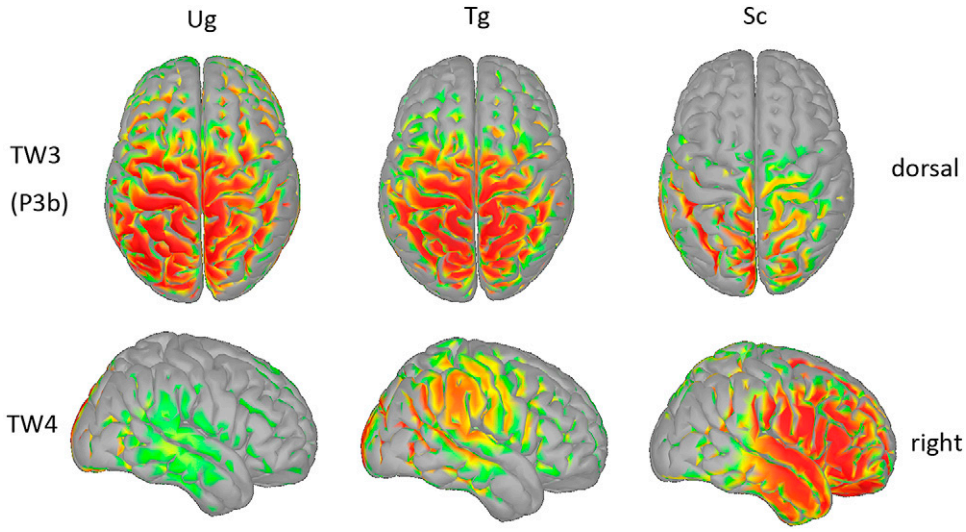


**FIG. 3.** Source currents of ERPs displayed on cortical surface in the time intervals of TW3 (400–600 ms) and TW4 (600–800 ms) under Ug, Tg, and Sc conditions. ERPs, event-related potentials. TW, time window.

role of "cause" hub at this time, when the "effect" is exerted on the posterior parietal and occipital regions.

Further graph-theoretical analysis reveals the temporal evolution of global topology of the effective connectivity networks. While a fixed threshold is set for the OPDC matrices, the connection density of the directed network in TW4 is significantly increased, with the number nearly doubling those of the OPDC networks in the other five time intervals (Fig. 6A). Corresponding to this result, modularity of the OPDC network in TW4 is significantly decreased (Fig. 6B). Less modularity of a network indicates higher functional integration among anatomically discrete brain areas (Bullmore and Sporns, 2009). Thus, the OPDC network constructed in TW4 can be seen as highly integrated and usually has more efficient information transfer.

Besides, from the baseline period to the intervals of TW1, TW2, and TW3, the number of "effect" nodes with indegree higher than outdegree in the OPDC networks is almost invariant over the time course, but obviously increases in the OPDC network of TW4 interval and then returns to the initial level at the end of task (TW5) (Fig. 6C). In the meantime, the number of "cause" nodes with higher outdegree in the OPDC networks is remarkably reduced (Fig. 6D). In TW4 interval, quantitatively, more inflow nodes mean that more "effect" regions are receiving incoming information flow from anterior neural system, and less outflow nodes

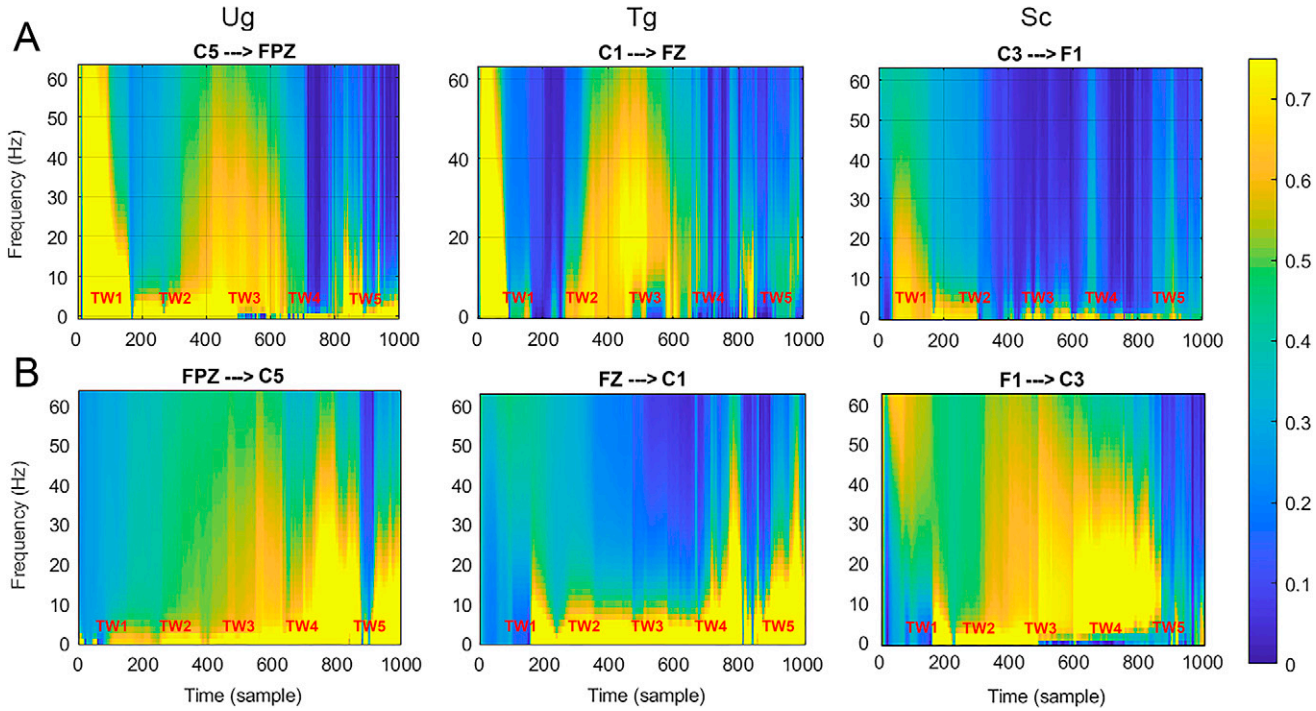


**FIG. 4.** Time-frequency diagram of OPDC between EEG channels located at frontal and sensorimotor regions in Ug, Tg, and Sc conditions. In each subdiagram, horizontal axis represents poststimulus time course of the task, and vertical axis depicts EEG frequencies. OPDC, orthogonalized partial directed coherence; EEG, electroencephalogram.

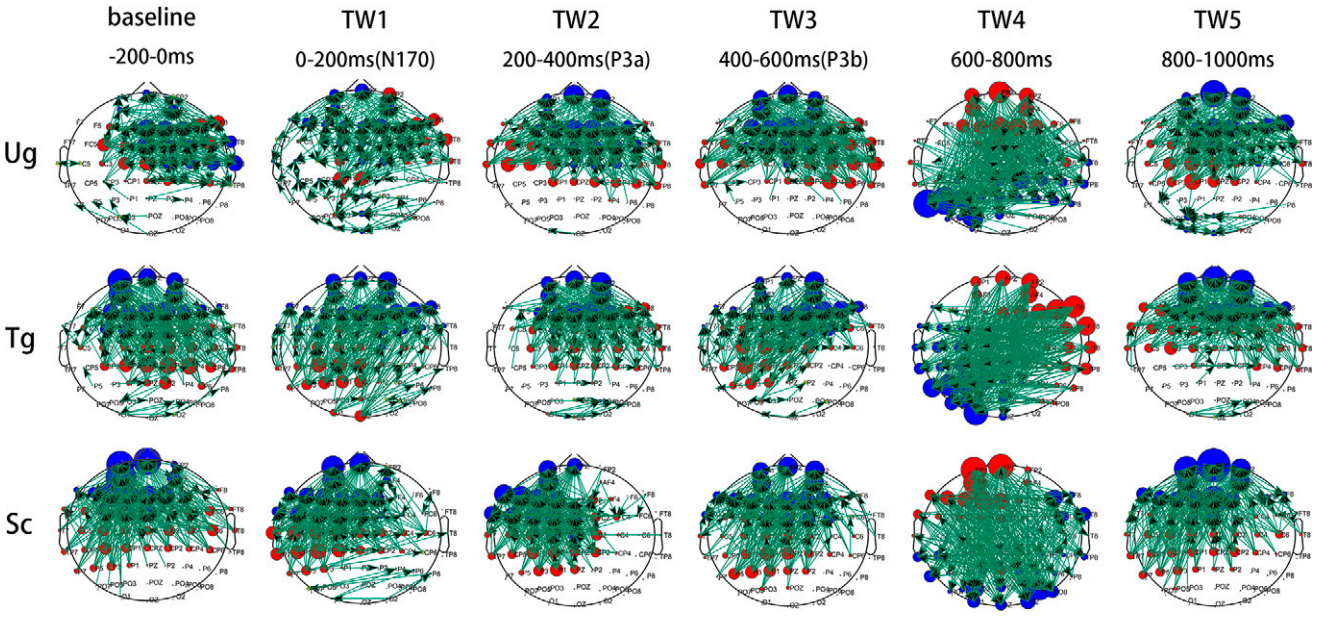


**FIG. 5.** OPDC-based networks formed under Ug, Tg, and Sc conditions in consecutive 200 ms time windows. The directed connections are constructed by setting a fixed connection density for each OPDC matrix. The red color marked in EEG channels represents node with higher outdegree than indegree in a network, whereas the blue channels mean the nodes with higher indegree than outdegree. OPDC, orthogonalized partial directed coherence; EEG, electroencephalogram.

indicate that the "cause" regions with outgoing information flow are more focused, primarily on the frontal cortex (Fig. 5).

Furthermore, in the tests for cosine similarity of indegree (resp. outdegree) vectors of nodes, the OPDC networks formed in baseline and TW1, the networks in TW1 and TW2, and the networks in TW2 and TW3 show high topological similarity in terms of spatial hub distribution (Fig. 7A, B). In the dynamic networks fluctuating along the task course, dramatic hub reorganization can be found in TW4 interval, when the OPDC network topology is greatly different with its former and latter networks in time course, that is, decreased cosine similarity between the indegree

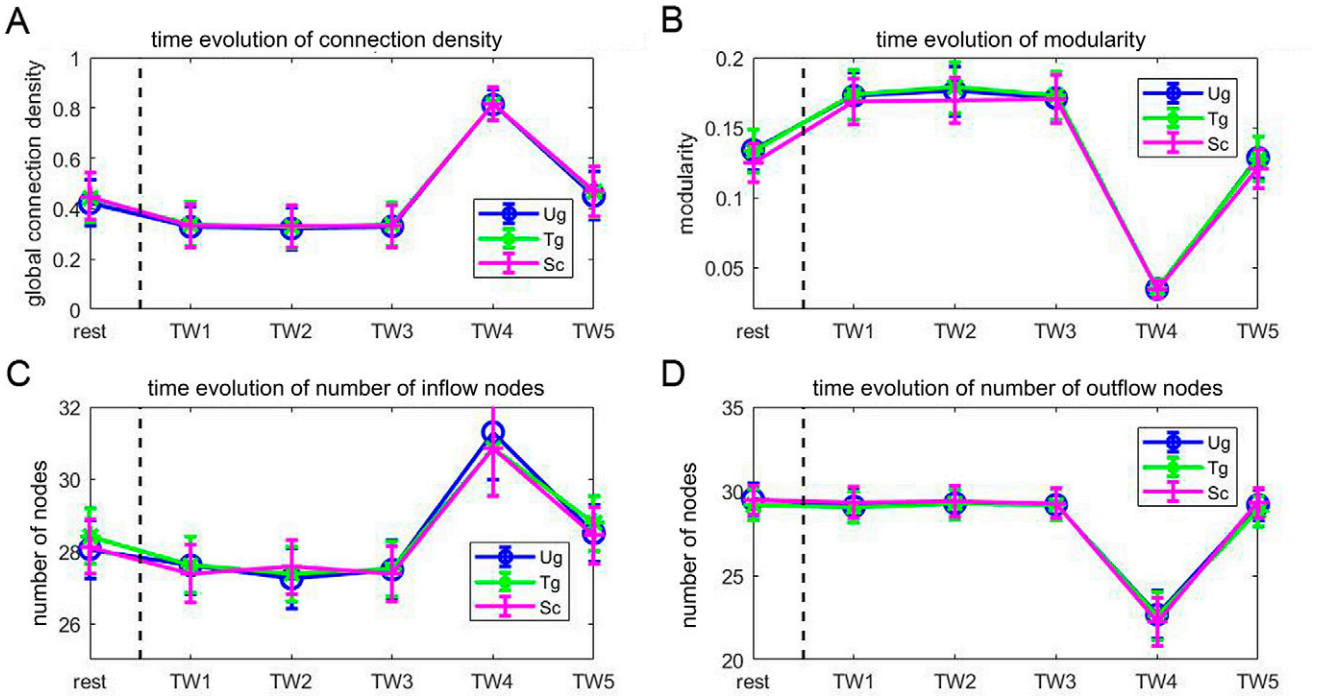


**FIG. 6.** Temporal evolution of global topology of OPDC networks with a fixed threshold in Ug, Tg, and Sc conditions. **(A)** Connection density, **(B)** modularity, **(C)** number of nodes with higher indegree, and **(D)** number of nodes with higher outdegree of the OPDC networks formed in consecutive 200 ms time windows. The blue, green, and purple curves represent Ug, Tg, and Sc conditions, respectively. OPDC, orthogonalized partial directed coherence.

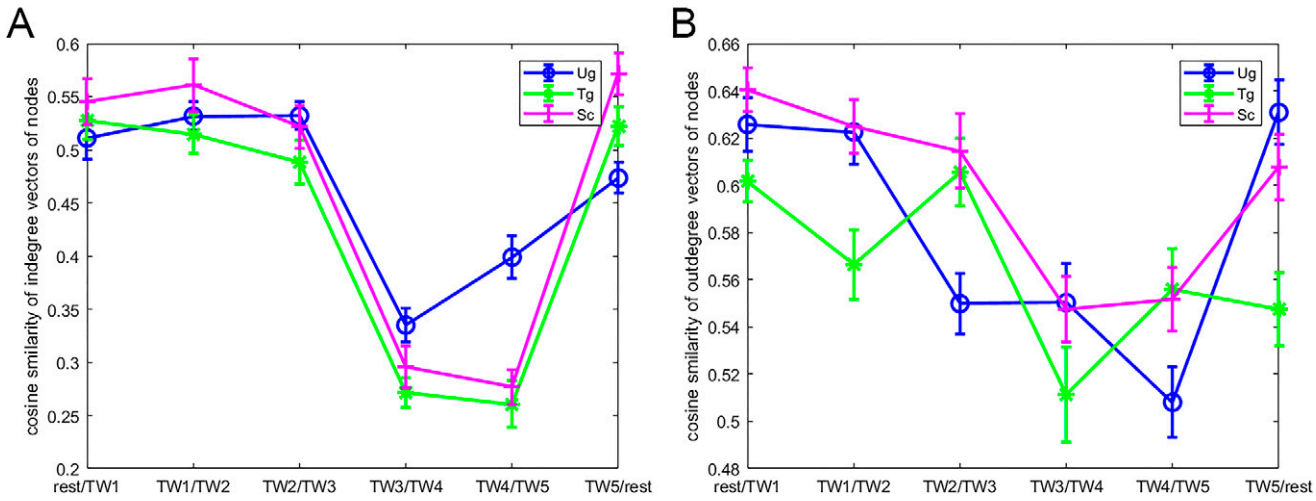


**FIG. 7.** Topological similarity between time-sequential directed networks. **(A)** Cosine similarity between indegree vectors of time-sequential networks. **(B)** Cosine similarity between outdegree vectors of time-sequential network. The blue, green, and purple curves represent Ug, Tg, and Sc conditions, respectively.

(resp. outdegree) vectors of the directed networks in TW3 and TW4 and those in TW4 and TW5.

### *Causality transition from mirror network to mentalizing network*

From the perspective of node dynamics, nodes in a network are processes rather than things (Taichi 2018; Yang et al., 2015). To measure the importance of nodes as input and output hubs in the cognitive process, we constructed the composition of the directed mirror and mentalizing networks, by integrating the adjacent matrices of OPDC networks formed in TW3 and TW4 intervals (Schlesinger et al., 2017; Taichi 2018). According to the interpretations of complex network measurements, nodes with a high number of shortest paths often have high betweenness centrality and, consequently, play an important role of a bridging hub that connects disparate parts of the network (Kaiser 2011; Rubinov and Sporns, 2010). Figure 8A depicts the betweenness score of sequential nodes in the composition networks. Under Ug, Tg, and Sc conditions, the nodes with high betweenness centrality basically focus on anterior frontal area and obey the form of power law, indicating the scale-free

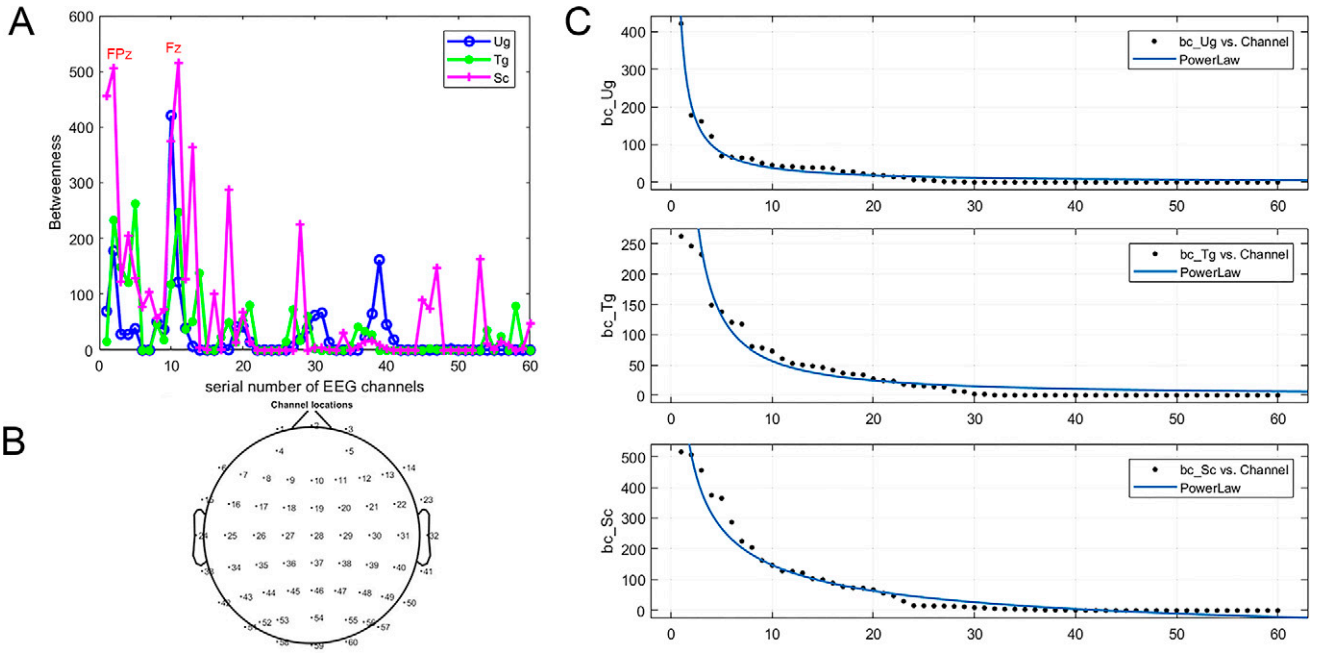


**FIG. 8.** Betweenness centrality of nodes in the composition of OPDC networks in TW3 and TW4. **(A)** Distribution of betweenness scores of nodes in the composition networks under Ug, Tg, and Sc conditions. **(B)** EEG channel locations and sequential numbers corresponding to nodes. **(C)** Power-law fitting of distribution of nodal betweenness scores, with root mean square error of 2.95, 4.79, and 11.41 under Ug, Tg, and Sc conditions, respectively. OPDC, orthogonalized partial directed coherence; TW, time window; EEG, electroencephalogram.

network characteristic with highly-connected local hubs (Fig. 8B, C).

The directed networks were further parsed through the nodes into different communities in TW3 and TW4 intervals using the Louvain method. The composition of communities through input and output hubs contains two important bridging nodes at midline PFC (channel FPz) and midline frontal cortex (channel Fz) with the highest betweenness centrality in the composition network (Fig. 9). It can be found that, in the reorganization of directed networks from mirror response to mentalizing function, the bridging nodes at frontal area experience a transition from input hub to output hub, which glue links together as interface between two time-ordering processes (Taichi 2018).

### *Discriminable graph features from causal flow transition*

As with the change of inflow and outflow strength of the nodes, the differences in internodal OPDC measures related to action intention types are also found in the channels located at frontal area. In TW3 interval, there is almost no difference between task conditions. However, the significant differences are revealed in TW4 interval, in which Sc condition evokes stronger directed connections from right superior frontal to right inferior frontal gyrus, directed links from sensorimotor to right frontotemporal regions, and directed edges from parietal to right frontotemporal regions compared with Ug and Tg conditions (Fig. 10).

Following the causality transition of OPDC networks from TW3 to TW4, the changes of local features are extracted from the nodes at prefrontal (mean value of channels FP1, FPz, FP2, AF3, and AF4), left frontal (mean value of channels F1, F3, F5, and F7), and right frontal (mean value of channels F2, F4, F6, and F8) regions, respectively. The subject-based samples are presented in the scatter diagrams of Figure 11, with a two-dimensional feature distribution composed of the temporal evolution of inflow strength and outflow strength of the channels. Further single-trial discrimination between conditions for each subject shows that, within the OPDC networks, the channels at prefrontal, left frontal, and right frontal areas have relatively high accuracies (around 0.6) in the classifications of "Ug versus Sc" and "Tg versus Sc" task conditions, whereas there is a meaningless accuracy (<0.5) in the binary classification for task conditions "Ug versus Tg" (Fig. 11). The results indicate that the functional transition from input hub to output hub at frontal area in OPDC-based directed networks can be used as effective features in identifying brain responses involved in decoding other's intention-oriented and unintelligible actions.

## Discussion

### *Time intervals of mirror response and mentalizing process*

Early studies using perceptual discrimination tasks suggest that the response time of mirror area is around 250–300 ms,

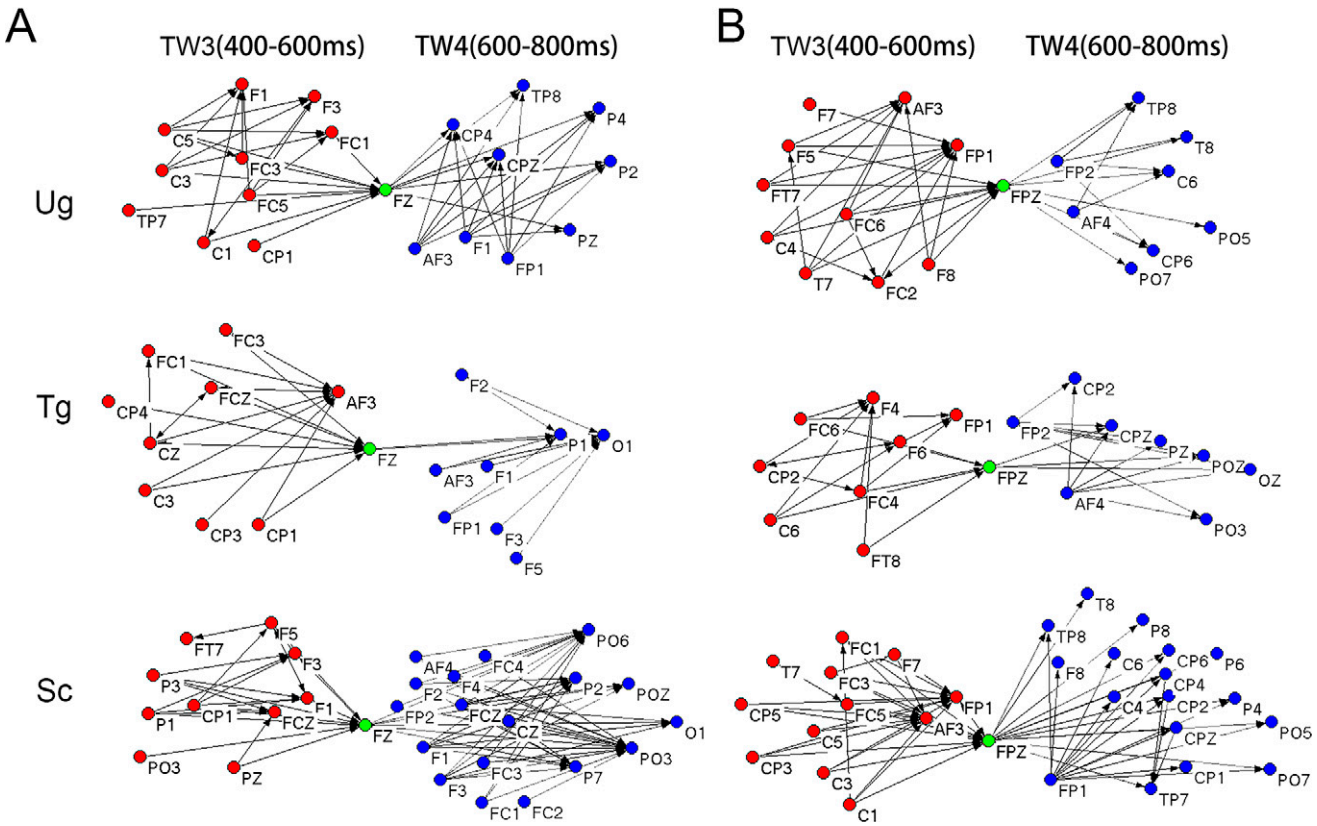


**FIG. 9.** Hub nodes in the composition of communities within OPDC networks in TW3 and TW4 under Ug, Tg, and Sc conditions. Red nodes belong to the community of networks in TW3, and blue nodes lie in the community of networks in TW4. The green nodes are the bridging hubs at midline **(A)** prefrontal and **(B)** frontal regions with high betweenness centrality between the temporally changing communities of the directed networks. OPDC, orthogonalized partial directed coherence; TW, time window.

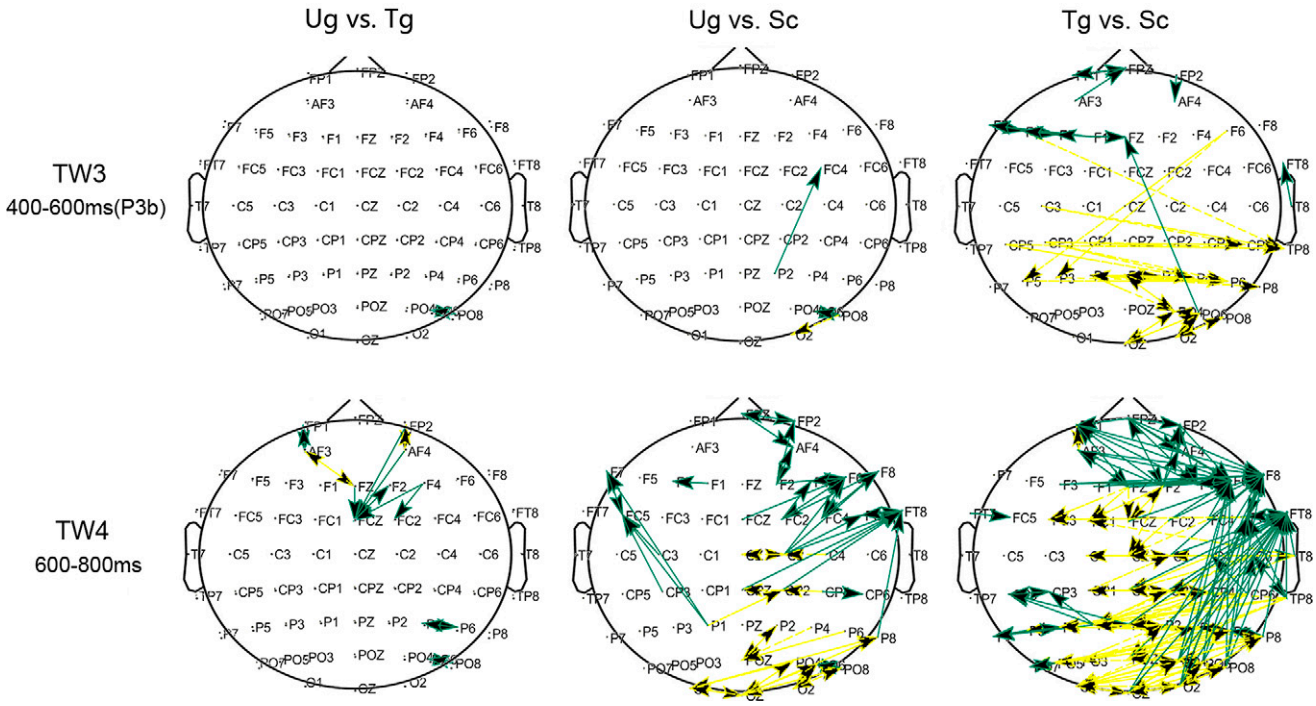


**FIG. 10.** Differences between the OPDC matrices in TW3 and TW4 from different conditions. From left to right: the networks are constructed by the significant difference in internodal OPDC from "Tg minus Ug," "Sc minus Ug," and "Sc minus Tg." The green directed links indicate higher OPDC (i.e., positive difference values), and the yellow directed links represent lower OPDC (i.e., negative difference values) (ANOVAs with a FDR correction: $p < 0.05$). OPDC, orthogonalized partial directed coherence; TW, time window; ANOVA, analysis of variance; FDR, false discovery rate.

whereas response times for intention identification are usually longer and show a wide degree of variability ranging from 600 to 1500 ms (Catmur 2015).

Mirror matching function. The directed networks formed in continuous time intervals, TW1, TW2, and TW3, show similar topological configuration manifested as high cosine similarity in terms of hub node distribution (around 0.6), which reflects sustained perceptual information processing in 0–600 ms interval. In the task-elicited sequential ERPs of the current study, the amplitudes of N170 in TW1 (0–200 ms) are significantly different ($p < 0.05$ in ANOVA test) among the three task conditions (Table 1), reflecting perceptual responses for differential physical properties of visual stimuli (Yang et al., 2011). Subsequently, P3a component in TW2 (200–400 ms) is believed to reflect attention allocation and memory updating involved in cognitive tasks (Berti 2015; Polich, 2007). Meanwhile, there is no significant potential difference ($p > 0.05$ in ANOVA test) among the three task conditions (Table 1), which is probably related with common function in switching focus of attention for novel stimuli within working memory mirrored in P3a component (Berti 2015). As a late positive potential, P3b component in TW3 (400–600 ms) has been suggested to reflect central cognitive processes occurring with the active detection of an attended stimulus and appears related to subsequent memory processing (Polich 2007). Specifically, Sc condition for understanding other's unintelligible action elicits the highest amplitude among the three conditions with significant difference ($p < 0.05$ in ANOVA test). It might be because subjectively improbable events will elicit a P3b and the less probable the event, the larger the P3b amplitude (Polich 2007). Particularly, the source currents of P3b are distributed on central sensorimotor area and extended to posterior parietal cortices, that is, the mirror neuron area revealed by most neuroimaging studies (Van Overwalle and Baetens 2009). ERP study by Möhring et al. (2014) suggests that task-induced N170 component reflects early activation of mirror neuron, whereas N400 potential in 350–550 ms interval can be best interpreted as reflecting sustained, but not initial, mirror neuron activity. Based on this evidence, it is plausible to suppose that the N170 elicited in this study indicates early perceptual initiation for visual stimuli and the P3b across 400–600 ms interval can be viewed as a neural index of mirror matching for perceived actions onto one's memory representation, between which the P3a across 200–400 ms interval reflects memory refresh and retention of visual perception information.

Mentalizing inference process. Following the mirror response, the directed network established in TW4 (600–800 ms) interval presents obviously different directional flows in topological structure. In addition, the source currents of the ongoing waveforms in TW4 interval are primarily localized at the TPJ area, and Sc condition for understanding other's unintelligible action induces the strongest activation in the PFC related to executive control for the inferential process. These brain areas have been found to be the fundamental components constituting the mentalizing system (Van Overwalle and Baetens 2009). The evidence suggests that the time segment of 600–800 ms is very likely involved in mentalizing process to respond to higher level judgment for other's action intention. The timing and location of mentalizing process are

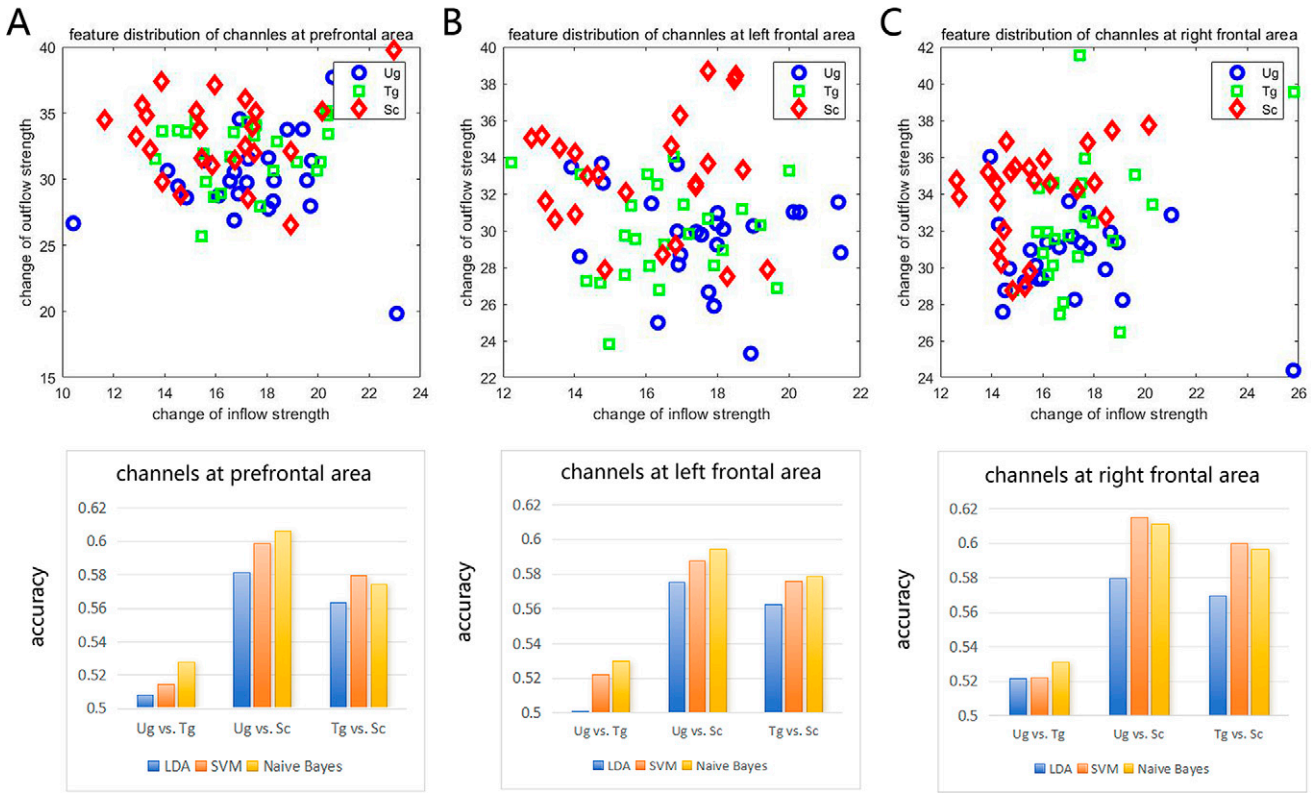


**FIG. 11.** Feature distributions and discrimination results for the change of nodal connection strength from TW3 to TW4. EEG channels **(A)** at the prefrontal cortex, **(B)** at the left frontal area, and **(C)** at the right frontal area. Top: the scattergrams of feature distribution of subject-based samples. The horizontal axis represents the change of inflow strength from TW3 to TW4 and the vertical axis outflow strength. The blue, green, and red markers represent the samples from the Ug, Tg, and Sc conditions, respectively. Bottom: discrimination accuracy averaged from all the subjects for the single-trial samples between each two-task conditions, based on LDA, SVM, and Naive Bayes methods. TW, time window; EEG, electroencephalogram, LDA, linear discriminant analysis; SVM, support vector machine.

partly overlapped with the ERP study by Beudt and Jacobsen' (2015) that identifies esthetic appreciation and judgment of visual abstract to be in the range of 700–1100 ms with significant positive activity at central and parietal areas.

### *Bridging hub in causal flows from mirror network to mentalizing network*

Early research does not completely support the opinion of direct perception and the intention understanding theory of mirror neuron function (Catmur, 2015). A substantial number of studies support the assumption that the mirror and mentalizing systems work together, and mirror neurons provide perceptual information to mentalizing system to complete inferential process (Gardner et al., 2015; Tidoni et al., 2013; Van Overwalle and Baetens, 2009). Therefore, the causal role of mirror and mentalizing areas in identifying other's intentions particularly needs to be validated by more evidence. In this study, the topological structures of mirror and mentalizing networks were analyzed, and crucial nodes were extracted to build the relevance between the two systems.

Sensorimotor-frontal causal connectivity involved in mirror response. Effective connectivity networks from TW1 to TW3 intervals show consistent causal directions from central sensorimotor to frontal areas, including prefrontal and superior frontal cortices. The localization of outflow nodes at central sensorimotor area is consistent with the study by Debnath et al. (2019) on desynchronization of Mu Rhythm of EEG (i.e., alpha frequency band recording in central localization) in infants' mirror system activity, which finds evident Mu-based activation over central regions during both execution and observation of movements. The localization of inflow nodes is also coincident with the functional magnetic resonance imaging (fMRI) study by Ge et al. (2023), in which bilateral middle frontal gyrus and bilateral inferior frontal gyrus are identified as important inflow clusters with excellent performance in action perception classification. Combining the current data with previous studies, we suggest that, before the activation of mentalizing system, central sensorimotor areas play the role of transmitting action perception information to the anterior neural system for mirror matching, in which frontal cortex is probably responsive for maintaining and supervising perceptual information within one's working memory system.

Frontal-parietal-temporal causal connectivity related to mentalization. During subsequent mentalizing process, effective connectivity networks were reorganized into a distinct information transfer mode with causal directions originating from frontal area to parietal, occipital, and temporal areas, showing low topological similarity with the networks in early visual perception and mirror matching periods. Specifically,

there are significant differences of the OPDC measures between task conditions (Fig. 10), in which understanding other's unintelligible action (Sc condition) induces higher directed connections from prefrontal to right inferior frontotemporal areas and directed links from parietal to right frontotemporal areas. Compared with usual, rational, or normal actions, intention recognition from irrational, implausible, or unusual actions is cognitively highly demanding and requires extra mental effort (Catmur, 2015; Van Overwalle and Baetens, 2009). Previous dynamic network research has suggested that higher cognitive effort can drive increased inter-regional interactions (Kitzbichler et al., 2011). Therefore, under Sc condition, the stronger right frontotemporal directed connection can be attributed to increased demand for inference function, and right parietal-temporal directed links are crucial for the representation of goals and intentions, to detect real purpose of action executor while observer cannot acquire other's intention from direct perception(Becchio et al., 2012; Brass et al., 2007; Van Overwalle and Baetens, 2009). Besides, right frontal-temporal network is quite special for high-order thinking activity, in which the right frontal cortex can be seen as responsible for imagination and idea generation, and the temporal lobes play an important role in idea editing and evaluation (Flaherty, 2005). Therefore, under Sc condition, the frontal cortex in mentalizing network is very likely to work out why the action is being performed, control the inference process, and make correct judgment, while there is a lack of enough action perception information.

The role of bridging hub of frontal cortex. Van Overwalle and Baetens (2009) suggest that, if the brain areas involved in mirror and mentalizing processes are overlapping, the two systems might share a common functional core, that is, the two systems may cooperate and inform each other. In this study, although the direction and strength of the information flows are significantly distinct between the early mirror network and the later mentalizing network, time ordering of information transfer revealed by the change of "cause-effect" relationship among the participating areas can give evidence about the core areas with shared information exchange.

The graph-theoretical analysis of OPDC networks finds that frontal area comprises the central high-degree nodes participating in many short paths within the directed networks formed in both mirror response and mentalization processes and, consequently, acts as important controls of information flow. In the reorganization of directed networks, frontal cortex plays an important role of "effect" hub containing more inflow nodes in mirror network and "cause" hub comprising more outflow nodes in the mentalizing network. For the nodes at frontal area successively serving as input and output hubs in the cognitive process, further analysis for the composition of mirror and mentalizing networks provides evidence that these nodes have the highest betweenness centrality within the composition network and, thus, act as bridging hubs between the sensorimotor-frontal community in mirror network and frontal-parietal community in mentalizing network (Fig. 9). When nodes are regarded as processes, not static things, the centralized hub organizations can bridge the gap between the parts in the dynamic directed networks (Schlesinger et al., 2017; Taichi 2018; Yang et al., 2015).

### *EEG features of dynamic causality for identifying action intentions*

In the field of human–computer interaction, understanding user's intention according to action is key to build intelligent systems and applications (Xiong et al., 2020; Yu et al., 2015). Based on the dynamic information transmission from recognizing action kinematics to inferring other's intention, feature extraction from OPDC-based directed networks was conducted in the EEG channels at frontal area, according to the special role of the frontal cortex in the hub reorganization of directed network from mirror response to mentalization. The changes in outflow strength and inflow strength of the nodes belonging to frontal community constitute input features for discriminating other's intention-oriented and unintelligible actions. The features identified by different classification algorithms could provide effective EEG time-frequency measurements and channel locations for further action intention recognition across subjects and trials. The major problem is that the classification accuracy of current data is unsatisfactory. Further optimization in experimental task design, classification algorithm, feature extraction, and combination related to rapid brain activity transition is needed to be systematically explored and developed in the future research.

## Conclusions

Using the OPDC method to decode dynamic fluctuations in the brain network structure related to continuous cognitive subprocesses, our study explored the interactive relationship between mirror network and mentalizing network. According to graph-theoretical analysis of directed network, the frontal cortex is identified as the shared core area of the two networks from input hubs transited to output hubs, consequently playing the role of bridging hubs of information transfer between mirror matching and intention inference processes. From the perspective of brain network dynamics, this study helps people deepen the understanding for the neural mechanism of action intention understanding. In the field of human–machine interaction, the EEG features extracted from this study for recognizing different types of action intention have potential application value for medical rehabilitation, such as motor dysfunction caused by stroke, spinal cord injury, and so on, by monitoring and analyzing patients' EEG signals.

## Authors' Contributions

L.Z.: Conceptualization, formal analysis, data curation, writing—original draft, and visualization. L.Z.: Supervision and writing—review and editing. J.W.: Methodology. Y.Z.: Resources, software, and funding acquisition.

## Author Disclosure Statement

The authors declare no conflict of interest.

## Funding Information

This work was supported, in part, by the Natural Science Foundation of Jiangsu Province under Grant BK20221181, the Natural Science Foundation of China under Grant 62077013, and the Fundamental Research Funds for the

Central Universities under Grants 2242022k30036 and 2242022k30037.

## References


Atique B, Erb M, Gharabaghi A, et al. Task-specific activity and connectivity within the mentalizing network during emotion and intention mentalizing. Neuroimage 2011;55(4):1899–1911; doi: 10.1016/j.neuroimage.2010.12.036

Baker AP, Brookes MJ, Rezek IA, et al. Fast transient networks in spontaneous human brain activity. Elife 2014;3(3):e01867; doi: 10.7554/eLife.01867

Bayazi MJD, Nasrabadi AM, Curran T. Frequency-specific network effective connectivity: ERP analysis of recognition memory process by directed connectivity estimators. Med Biol Eng 2021;1:1–16; doi: 10.1007/s11517-020-02304-8

Becchio C, Cavallo A, Begliomini C, et al. Social grasping: From mirroring to mentalizing. Neuroimage 2012;61(1): 240–248; doi: 10.1016/j.neuroimage.2012.03.013

Berti S. Switching attention within working memory is reflected in the p3a component of the human event-related brain potential. Front Hum Neurosci 2015;9:701–710; doi: 10.3389/fnhum.2015.00701

Beudt S, Jacobsen T. On the role of mentalizing processes in aesthetic appreciation: An ERP study. Front Hum Neurosci 2015;9:600; doi: 10.3389/fnhum.2015.00600

Blakemore SJ, Decety J. From the perception of action to the understanding of intention. Nat Rev Neurosci 2001;2(8): 561–567; doi: 10.1016/j.yhbeh.2004.06.011

Blondel VD, Guillaume JL, Lambiotte R, et al. Fast unfolding of communities in large networks. J Stat Mech 2008;2008(10): P10008; doi: 10.1088/1742-5468/2008/10/p10008

Bola M, Sabel BA. Dynamic reorganization of brain functional networks during cognition. Neuroimage 2015;114:398–413; doi: 10.1016/j.neuroimage.2015.03.057

Bonini L, Rozzi S, Serventi FU, et al. Ventral premotor and inferior parietal cortices make distinct contribution to action organization and intention understanding. Cereb Cortex 2010;20(6):1372–1385; doi: 10.1093/cercor/bhp200

Brandone AC, Stout W. The origins of theory of mind in infant social cognition: Investigating longitudinal pathways from intention understanding and joint attention to preschool theory of mind. J Cogn Dev 2023;24(3):375–396; doi: 10.1080/15248372.2022.2146117

Brass M, Schmitt RM, Spengler S, et al. Investigating action understanding: Inferential processes versus action simulation. Curr Biol 2007;17(24):2117–2121; doi: 10.1016/j.cub.2007.11.057

Buccino G, Baumgaertner A, Colle L, et al. The neural basis for understanding non-intended actions. Neuroimage 2007;36 (Suppl 2):T119–T127; doi: 10.1016/j.neuroimage.2007.03.036

Bullmore E, Sporns O. Complex brain networks: Graph theoretical analysis of structural and functional systems. Nat Rev Neurosci 2009;10(3):186–198; doi: 10.1038/nrn2575

Catmur C. Understanding intentions from actions: Direct perception, inference, and the roles of mirror and mentalizing systems. Conscious Cogn 2015;36:426–433; doi: 10.1016/j.concog.2015.03.012

Centelles L, Assaiante C, Nazarian B, et al. Recruitment of both the mirror and the mentalizing networks when observing social interactions depicted by point-lights: A neuroimaging study. PLoS One 2011;6(1):e15749; doi: 0.1371/journal.pone.0015749

Cheng Y, Meltzoff AN, Decety J. Motivation modulates the activity of the human mirror-neuron system. Cereb Cortex 2007;17(8):1979–1986; doi: 10.1093/cercor/bhl107

Chiavarino C, Apperly IA, Humphreys GW. Understanding intentions. Curr Dir Psychol Sci 2012;21(5):284–289; doi: 10.1177/0963721412452727

De Lange FP, Spronk M, Willems RM, et al. Complementary systems for understanding action intentions. Curr Biol 2008; 18(6):454–457; doi: 10.1016/j.cub.2008.02.057

Debnath R, Salo VC, Buzzell GA, et al. Mu rhythm desynchronization is specific to action execution and observation: Evidence from time-frequency and connectivity analysis. Neuroimage 2019;184:496–507; doi: 10.1016/j.neuroimage.2018.09.053

Erdős P, Rényi A. On the strength of connectedness of a random graph. Acta Math Hungar 1964;12(1–2):261–267; doi: 10.1007/BF02066689

Flaherty AW. Frontotemporal and dopaminergic control of idea generation and creative drive. J Comp Neurol 2005;493(1): 147–153; doi: 10.1002/cne.20768

Friston K, Moran R, Seth AK. Analysing connectivity with Granger causality and dynamic causal modelling. Curr Opin Neurobiol 2013;23(2):172–178; doi: 10.1016/j.conb.2012.11.010

Gaetz W, Liu C, Zhu H, et al. Evidence for a motor gamma-band network governing response interference. Neuroimage 2013;74:245–253; doi: 10.1016/j.neuroimage.2013.02.013

Gardner T, Goulden N, Cross ES. Dynamic modulation of the action observation network by movement familiarity. J Neurosci 2015;35(4):1561–1572; doi: 10.1523/JNEUROSCI.2942-14.2015

Ge S, Ding MY, Zhang Z, et al. Temporal-spatial features of intention understanding based on EEG-fNIRs bimodal measurement. IEEE Access 2017;5:14245–14258; doi: 10.1109/ACCESS.2017.2723428

Ge S, He J, Lin P, et al. Effective connectivity analysis and classification of action observation from different perspectives: An fMRI study. IEEE Trans Biomed Eng 2023;70(2): 723–734; doi: 10.1109/TBME.2022.3201547

Ge S, Liu H, Lin P, et al. Neural basis of action observation and understanding from first- and third-person perspectives: An fMRI study. Front Behav Neurosci 2018;12:283–210; doi: 10.3389/fnbeh.2018.00283

Ge S, Wang P, Liu H, et al. Neural activity and decoding of action observation using combined EEG and fNIRS measurement. Front Hum Neurosci 2019;13:357–315; doi: 10.3389/fnhum.2019.00357

Gramfort A, Papadopoulo T, Olivi E, et al. OpenMEEG: Open-source software for quasistatic bioelectromagnetics. Biomed Eng Online 2010;9(1):45; doi: 10.1186/1475-925X-9-45

Herrmann CS, Frund I, Lenz D. Human gamma-band activity: A review on cognitive and behavioral correlates and network models. Neurosci Biobehav Rev 2010;34(7):981–992; doi: 10.1016/j.neubiorev.2009.09.001

Hipp JF, Hawellek DJ, Corbetta M, et al. Large-scale cortical correlation structure of spontaneous oscillatory activity. Nat Neurosci 2012;15(6):884–890; doi: 10.1038/nn.3101

Huang D, Ren A, Shang J, et al. Combining partial directed coherence and graph theory to analyse effective brain networks of different mental tasks. Front Hum Neurosci 2016; 10:235–211; doi: 10.3389/fnhum.2016.00235

Kaiser M. A tutorial in connectome analysis: Topological and spatial features of brain networks. Neuroimage 2011;57(3): 892–907; doi: 10.1016/j.neuroimage.2011.05.025

Kitzbichler MG, Henson RN, Smith ML, et al. Cognitive effort drives workspace configuration of human brain functional networks. J Neurosci 2011;31(22):8259–8270; doi: 10.1523/JNEUROSCI.0440-11.2011

Kumar N, Jaiswal A, Roy D, et al. Effective networks mediate right hemispheric dominance of human 40 Hz auditory steady-state response. Neuropsychologia 2023;184:108559; doi: 10.1101/2023.02.02.526849

Liew SL, Han S, Aziz-Zadeh L. Familiarity modulates mirror neuron and mentalizing regions during intention understanding. Hum Brain Mapp 2011;32(11):1986–1997; doi: 10.1002/hbm.21164

Möhring N, Brandt ESL, Mohr B, et al. ERP adaptation provides direct evidence for early mirror neuron activation in the inferior parietal lobule. Int J Psychophysiol 2014;94(1): 76–83; doi: 10.1016/j.ijpsycho.2014.07.001

Oberman LM, Pineda JA, Ramachandran VS. The human mirror neuron system: A link between action observation and social skills. Soc Cogn Affect Neurosci 2007;2(1):62–66; doi: 10.1093/scan/nsl022

Omidvarnia A, Azemi G, Boashash B, et al. Measuring time-varying information flow in scalp EEG signals: Orthogonalized partial directed coherence. IEEE Trans Biomed Eng 2014;61(3):680–693; doi: 10.1109/TBME.2013.2286394

Opsahl T, Agneessens F, Skvoretz J. Node centrality in weighted networks: Generalizing degree and shortest paths. Soc Networks 2010;32(3):245–251; doi: 10.1016/j.socnet.2010.03.006

Polich J. Updating P300: An integrative theory of P3a and P3b. Clin Neurophysiol 2007;118(10):2128–2148; doi: 10.1016/j.clinph.2007.04.019

Rizzolatti G, Fogassi L, Gallese V. Neurophysiological mechanisms underlying the understanding and imitation of action. Nat Rev Neurosci 2001;2(9):661–670; doi: 10.1038/35090060

Rubinov M, Sporns O. Complex network measures of brain connectivity: Uses and interpretations. Neuroimage 2010;52(3): 1059–1069; doi: 10.1016/j.neuroimage.2009.10.003

Schlesinger KJ, Turner BO, Grafton ST, et al. Improving resolution of dynamic communities in human brain networks through targeted node removal. Plos One 2017;12(12): e0187715–e0187728; doi: 10.1371/journal.pone.0187715

Seth AK. A matlab toolbox for granger causal connectivity analysis. J Neurosci Methods 2010;186(2):262–273; doi: 10.1016/j.jneumeth.2009.11.020

Sporns O, Honey CJ, Kötter R. Identification and classification of hubs in brain networks. Plos One 2007;2(10):e1049; doi: 10.1371/journal.pone.0001049

Spunt RP, Satpute AB, Lieberman MD. Identifying the what, why, and how of an observed action: An fMRI study of mentalizing and mechanizing during action observation. J Cogn Neurosci 2011;23(1):63–74; doi: 10.1162/jocn.2010.21446

Tadel F, Baillet S, Mosher JC, et al. Brainstorm: A user-friendly application for MEG/EEG analysis. Comput Intell Neurosci 2011;2011(8):879716; doi: 10.1155/2011/879716

Taichi H. Open networks from within: Input or output betweenness centrality of nodes in directed networks. Appli Netw Sci 2018;3(1):15; doi: 10.1007/s41109-018-0076-1

Tidoni E, Borgomaneri S, di Pellegrino G, et al. Action simulation plays a critical role in deceptive action recognition. J Neurosci 2013;33(2):611–623; doi: 10.1523/JNEUROSCI.2228-11.2013

van den Heuvel MP, Sporns O. Network hubs in the human brain. Trends Cogn Sci 2013;17(12):683–696; doi: 10.1016/j.tics.2013.09.012

Van Overwalle F, Baetens K. Understanding others' actions and goals by mirror and mentalizing systems: A meta-analysis. Neuroimage 2009;48(3):564–584; doi: 10.1016/j.neuroimage.2009.06.009

Vezoli J, Bastos A, Bosman C, et al. Inter-areal causal interactions in the gamma and beta frequency bands define a functional hierarchy in the primate visual system. Perception 2013;42(6):143–143.

Woodward AL, Gerson SA. Mirroring and the development of action understanding. Philos Trans R Soc Lond B Biol Sci 2014;369(1644):20130181; doi: 10.1098/RSTB.2013.0181

Xiong X, Yu Z, Ma T, et al. Classifying action intention understanding EEG signals based on weighted brain network metric features. Biomed Signal Proces 2020;59:101893; doi: 10.1016/j.bspc.2020.101893

Yang Y, Gu G, Guo H, et al. Early event-related potential components in face perception reflect the sequential neural activities. Acta Physiol Sin 2011;63(2):97–105; doi: 10.1007/s12583-011-0153-1

Yang C, Qiu J, He H. Exponential synchronization for a class of complex spatio-temporal networks with space-varying coefficients. Neurocomputing 2015;151:401–407; doi: 10.1016/j.neucom.2014.09.025

Yu Z, Kim S, Mallipeddi R, et al. 2015. Human intention understanding based on object affordance and action classification. In: IEEE international joint conference on neural networks, pp. 1–6; doi: 10.1109/IJCNN.2015.7280587

Zhang L, Gan JQ, Zheng W, et al. Spatiotemporal phase synchronization in adaptive reconfiguration from action observation network to mentalizing network for understanding other's action intention. Brain Topogr 2018;31(3):447–467; doi: 10.1007/s10548-017-0614-7

Zhang Q, Hu Y, Potter T, et al. Establishing functional brain networks using a nonlinear partial directed coherence method to predict epileptic seizures. J Neurosci Methods 2020;329: 108447; doi: 10.1016/j.jneumeth.2019.108447

Address correspondence to:
*Yanmei Zhu*
*Nanjing Xiaozhuang University*
*No. 3601*
*Hongjing Avenue*
*Jiangning District*
*Nanjing*
*PR China*

*E-mail:* zhuyanmei@njxzc.edu.cn