Causal Connectivity Transition from Action Observation to Mentalizing Network for Understanding Other's Action Intention

Li Zhang[1], Jing Wang[2], and Yanmei Zhu[1,3,4](✉)

[1] School of Early-Childhood Education, Nanjing Xiaozhuang University, Nanjing 211171, Jiangsu, People's Republic of China
{li_zhang,zhuyanmei}@njxzc.edu.cn

[2] School of Computer Science and Information Technology, Xinyang Normal University, Xinyang 464000, People's Republic of China
wangjing@xynu.edu.cn

[3] Key Laboratory of Child Development and Learning Science, Southeast University, Ministry of Education, Nanjing, People's Republic of China

[4] Research Center for Learning Science, Southeast University, Nanjing 210096, Jiangsu, People's Republic of China

Abstract. The previous neuroimaging studies have found that two major cognitive sub-processes, action perception and mental inference, participate in understanding others' action intention, but it is unclear that the role of action observation network (AON) for mentalizing network (MZN) of intention inference. To provide direct causal evidence about the relationship between the two systems, this EEG study adopted Granger causality method to detect the circuit of directed information transfer from action perception to intention inference process during a "hand-cup interaction" observation task with two types of actions, i.e., usual intention-oriented action and unintelligible action. The graph-theoretical results of causal connectivity network show that left-lateral posterior parietal-occipital brain area acts as "effect" nodes in AON during action perception period but plays the role of "cause" nodes in MZN, especially for understanding other's unintelligible action that requires higher cognitive function for mentalizing inference. From the evidence, this study suggests that left-lateral parietal-occipital brain area can be viewed as a hub of internodal directed connection transition from AON to MZN, so that the two systems could cooperate with each other by means of temporal reception and transmission of perceptional information to judge other's actual intention.

Keywords: Action intention understanding · action observation network · mentalizing network · ERP-based sources · Granger causality

Supported by Key Laboratory of Child Development and Learning Science (Southeast University), Ministry of Education, PR China.

M. Tanveer et al. (Eds.): ICONIP 2022, CCIS 1792, pp. 350–360, 2023.
https://doi.org/10.1007/978-981-99-1642-9_30

1 Introduction

Understanding others' intention from their actions is an essential ability of the human living in the social world [1]. The past studies have identified two stages are highly involved in action intention understanding process of the brain. The first stage is direct perception, which maps the visual information of actions onto memory representation in one's memory system. This process is considered to activate an action observation network (AON) composed of mirror neuron areas. The subsequent stage is intention inference, which judges others' mental state or motivation from observed actions. The process is considered to rely on a mentalizing network (MZN) consisting of frontoparietal system [2–4]. The previous neuroimaging studies have detected significant activation of the both networks in action intention understanding tasks, but they have not achieved a common conclusion about the functional relevance between the two systems. Specifically, the role of AON is unclear in the stage of intention inference while the MZN is activated. Some studies suggest that the mirror neuron areas should provide sensorimotor information to mentalizing areas for inferring others' intentions correctly. By contrast, some studies think that the mirror and the mentalizing systems are probably independent of each other, because concurrent activation of the two systems was rarely detected in action intention understanding tasks [5–7]. Therefore, more direct causal evidence is needed to reveal how the two systems cooperate in action intention understanding process [8,9]. In this electroencephalogram (EEG) study, we used a “hand-cup interaction” action observation task with two types of intentions to test the interactive relationship between AON and MZN. The timing and localization of mirror responses and mentalization were determined by event-related potential (ERP) and source trace. In the time intervals of task-evoked ERPs, EEG channel-level Granger causality (GC) was computed and directed causal network was constructed to capture the change of directed information flow among key brain regions. Furthermore, graph-theoretical measurements of directed networks were discriminated to discover identifiable EEG channels and features in understanding others' different action intentions.

2 Materials and Methods

2.1 EEG Experiment and Data Prepocessing

The EEG experiment was approved by the Academic Committee of the Research Center for Learning Science, Southeast University, China. EEG data were recorded by a 60-channel Neuroscan 10–20 system with sampling rate at 500 Hz. In the EEG experiment, 30 college students were recruited to perform a “hand-cup interaction” observation task, in which 24 subjects' effective data were retained to be used in further data analysis, including 10 males and 14 females aged 22.4 $\pm$ 2.3 (mean $\pm$ SD).

The task was composed of two conditions used for comparing brain activities induced by different action intention types. As shown in Fig. 1A, the actions presented in the experiment include a typical intention-oriented action, i.e., grasping a cup for using it (Ug), and an unintelligible action, i.e., touching a cup without clear purpose (Sc). There were 98 trials for each condition, thus resulting in 196 trials in total. Figure 1B shows the timeline of sequential stimuli of each trial. At first, the symbol "+" at the center of screen was presented for 150 ms. Then a cup was shown for 500 ms. After that, the screen presented a hand interacting with the cup for 2000 ms. Meanwhile, subjects needed to judge the intention in their brains without pressing any button. At the end, the symbol "+" was presented again with a random time length, which was the beginning of next trial.

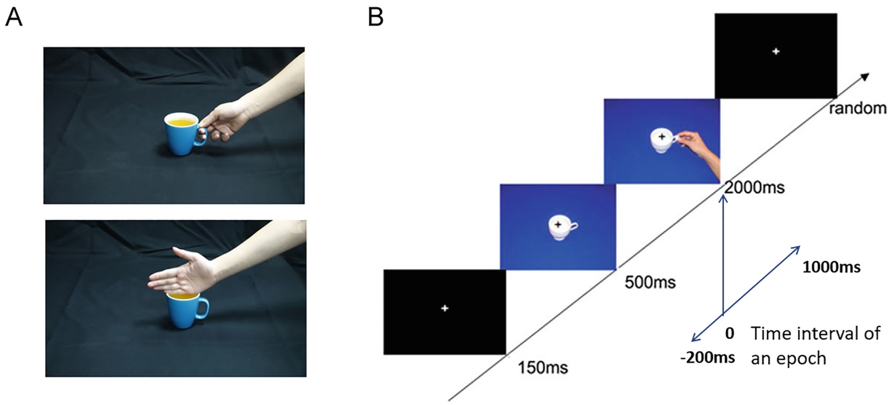

Fig. 1. Experimental paradigm of "hand-cup interaction" observation task. (A) a hand grasping a cup for using it (Ug); a hand touching a cup without any obvious purpose (Sc). (B) Timeline of stimulus presentation and time interval of an epoch of EEG data.

The raw EEG signals were preprocessed by Scan 4.3 software. After extracting the trials with the epoch of 1200 ms (200-ms pre-stimulus and 1000-ms post-stimulus intervals), baseline correction, artifact rejection and low-pass filtering (1–60 Hz) were conducted subsequently. As a result, 1146 and 1139 trials were retained for Ug and Sc task conditions respectively, of which 36–68 trials were retained for each subject under each condition.

2.2 Source Estimation of ERP Difference Wave

According to task-evoked event-related potential (ERP) responses, difference waves between "Ug" and "Sc" conditions were calculated to isolate the brain activities of interest. Based on the topology of difference waveforms with global field potential (GFP) peaks, the cortical sources were estimated by using Brainstorm source estimation procedure (http://neuroimage.usc.edu/brainstorm). In the procedure, a forward model was created by the symmetric Boundary Element Model in OpenMEEG (http://openmeeg.github.io) toolbox [10]. The noise of sensors was removed by the noise covariance matrix of the signals in pre-stimulus interval. After that, an inverse kernel matrix was produced by the forward model and standardized Low Resolution Brain Electromagnetic Tomography (sLORETA) algorithm. As a result, the cortical sources of difference waves between the ERPs of "Ug" and "Sc" conditions were estimated by means of the inverse kernel matrix, which were then mapped onto a distributed cortex source model composed of 15,002 elementary current dipoles.

2.3 Directed Graph Analysis of Effective Connectivity

To identify how action perception and intention inference processes modulate intraregional influence among crucial brain areas, directed connectivity networks were constructed by calculating GC between each pair of EEG signals.

For two simultaneously measured signals *x(t)* and *y(t)*, if one can predict the first signal better by incorporating the past information from the second signal than using only information from the first one, then the second signal can be called causal to the first one [11]. Clive Granger gave a mathematical formulation of this concept by arguing that when x is influencing y, then if you add past values of *x(t)* to the regression of *y(t)*, and improvement on the prediction will be obtained.

For the univariate autoregressive model (AR),

$$x(n) = \textstyle\sum_{k=1}^{p} a_{x,k} x(n-k) + u_x(n) \tag{1}$$

$$y(n) = \textstyle\sum_{k=1}^{p} a_{y,k} y(n-k) + u_y(n) \tag{2}$$

where $a_{i,j}$ are the model parameters (coefficients usually estimated by least square method), p is the order of the AR model and u_i are the residuals associated to the model. Here, the prediction of each signal (x and y) is performed only by its own past ($\overline{x}$ and $\overline{y}$ respectively). The variances of the residuals are denoted by

$$V_{x|\bar{x}} = var(u_x) \tag{3}$$

$$V_{y|\bar{y}} = var(u_y) \tag{4}$$

For the bivariate AR,

$$x(n) = \textstyle\sum_{k=1}^{p} a_{x|x,k} x(n-k) + \sum_{k=1}^{p} a_{x|y,k} y(n-k) + u_{xy}(n) \tag{5}$$

$$y(n) = \textstyle\sum_{k=1}^{p} a_{y|x,k} x(n-k) + \sum_{k=1}^{p} a_{y|y,k} y(n-k) + u_{yx}(n) \tag{6}$$

The residuals depend on the past value of both signals and their variances are

$$V_{x|\bar{x},\bar{y}} = var(u_{xy}) \tag{7}$$

$$V_{y|\bar{x},\bar{y}} = var(u_{yx}) \tag{8}$$

where *var(.)* is the variance over time and $x \mid x, y$ is the prediction of *x(t)* by the past samples of values of *x(t)* and *y(t)*.

Therefore, GC from y to x (prediction x from y) is

$$GC_{y \to x} = ln(\frac{V_{x|\overline{x}}}{V_{x|\overline{x},\overline{y}}}) \tag{9}$$

The range of $GC_{y \to x}$ is between 0 and ∞. $GC_{y \to x} = 0$ means that the past of *y(t)* does not improve the prediction of *x(t)*, i.e., $V_{x|\bar{x}} \approx V_{x|\bar{x},\bar{y}}$, and $GC_{y \to x} > 0$ denotes that the past of *y(t)* improves the prediction of *x(t)*, i.e., $V_{x|\bar{x}} \gg V_{x|\bar{x},\bar{y}}$ (y G-causes x).

In this study, GCs were calculated in ERP time intervals with statistically significant between-condition differences. Based on GCs of each pair of EEG signals, directed connectivity matrices were generated with asymmetry characteristic. After setting a fixed connection density, the channel-based causal networks were constructed. Then, the local node characteristic was estimated according to graph theory of complex network [12]. In a directed connection network, N is the set of all the nodes in the network, and $(i \to j)$ represents the directed link from nodes i to *j*, $(i, j \in N; i \neq j)$. If there is directed connection status from nodes i to *j*, $a_{i \to j} = 1$; otherwise, $a_{i \to j} = 0$. 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.

$$k_{i(in)} = \sum_{j \in N, i \neq j} a_{j \to i} \tag{10}$$

$$k_{i(out)} = \sum_{j \in N, i \neq j} a_{i \to j} \tag{11}$$

For an individual node, indegree and outdegree were computed to assess the role of a node in a directed network.

2.4 Statistical and Discriminate Analyses

To isolate the brain responses related to action intention types, the group-based ERPs elicited by different task conditions from electrode FZ at frontal midline area were statistically tested by one-way analysis of variance (ANOVA). The internodal GCs of directed networks between "Ug" and "Sc" conditions were statistically compared by the ANOVA to detect differences in links of the Granger Causality networks. A false discovery rate (FDR) procedure was conducted to correct for multiple hypothesis testing, with significance level set to 0.05. The null hypothesis is that the difference between task conditions is zero. Furthermore, local nodal parameters measured in N170-P200 and P400-700 time intervals constitute input features for the discriminant analysis between "Ug" and "Sc" conditions. The subject-based feature samples were recognized by linear discriminant analysis (LDA) with 10-fold cross validation, to reveal the transition of inflow and outflow nodes from AON to MZN and determine distinguishable EEG channels and features of brain states while understanding other's different action intentions.

3 Results and Discussions

Under "Ug" and "Sc" conditions of the "hand-cup interaction" action observation task, it can be seen that both the two task conditions evoke significant ERP responses in post-stimulus 170–200 ms, 300 ms and 400–700 ms time intervals (see Fig. 2), which can be represented by N170-P200, P300 (P3a) and P400-700 (P3b) ERP components.

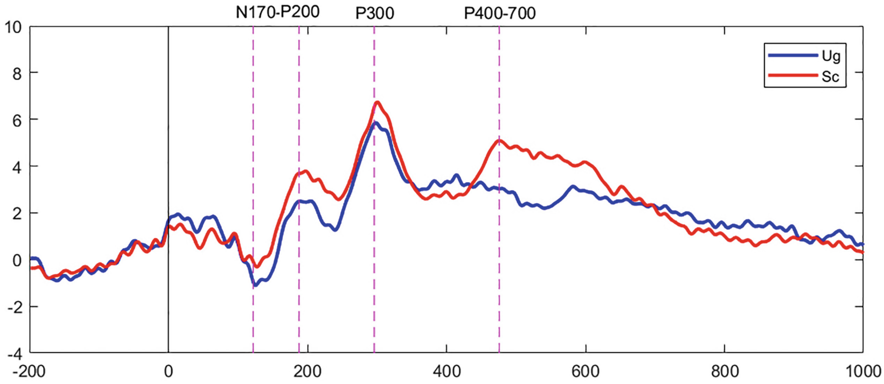

Fig. 2. Grand average of ERPs for "Ug" and "Sc" conditions from EEG channel FZ. Time = 0 corresponds to the onset of "hand-cup interaction" presentation. The figure shows that each condition has elicited significant ERP components marked with vertical dotted lines. The blue and red solid lines represent the "Ug" and "Sc" conditions respectively. (Color figure online)

Further between-condition ANOVA results show that significant difference in ERP responses were generated in N170-P200 and P400-700 time intervals (Table 1), when agent's unintelligible action particularly elicited higher ERP response amplitudes of subjects (Fig. 2).

Table 1. ANOVA results between conditions for the task-evoked ERPs. F is the ratio of between-group mean variance to within-group variance; p value indicates significance level of ANOVA, in which * represents $p < 0.05$ and ** denotes $p < 0.01$.

ERP component	N170-P200	P300	P400-700
Time interval	156-248ms	274-320ms	326-700ms
F	6.01	2.99	7.15
p	0.0143*	0.0837	0.0075**

The source estimation results show that the cortical sources of the difference waves of N170-P200 response are localized at anterior intraparietal sulcus, the premotor cortex and superior temporal sulcus in left cerebral hemisphere, which have been demonstrated as the major brain regions constituting the AON for mirror function. Besides, N170 is a non-specific, motion-related component and P200 is known to be sensitive to physical properties of visual stimuli. P400-700 is generally suggested to indicate central cognitive processing of attended stimulus and related to subsequent memory processing [13]. The sources of the difference waves of P400-700 response are distributed at right temporoparietal junction and the medial prefrontal cortex (see Fig. 3), which are the major components of the MZN for higher-level intention inference.

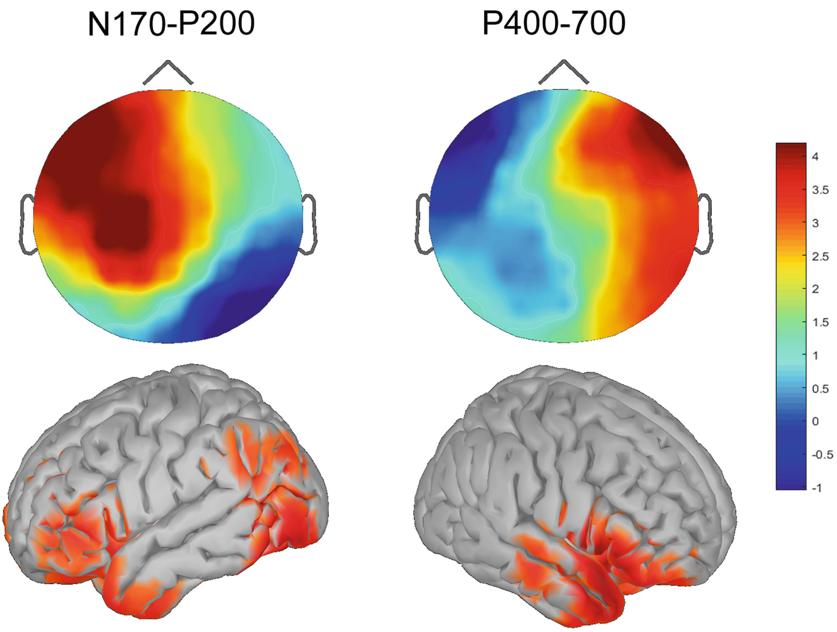

Fig. 3. Source current distribution of difference waves between ERPs evoked by "Ug" and "Sc" task conditions. The top line is the current mapped on the scalp and the bottom is the source current localized on the cortical surface in N170-P200 and P400-700 time intervals.

Based on the results of ERP and source analysis, it can be speculated that N170-P200 is indicative of the mirror mechanism that acquires information from other's action kinematics, i.e., the activation of AON, whereas P400-700 implicates more information of high-order mentalizing process that infers the intentions of other's gestures, i.e., the formation of MZN.

Under the two task conditions, the GC network topologies transformed from AON to MZN are presented in Fig. 4. It can be seen that, during action perception period represented by N170-P200 response, "Ug" and "Sc" conditions basically elicited directed information transmission from dorsolateral frontal regions to midline frontal area, i.e., EEG channels at bilateral dorsolateral frontal and central regions act as "cause" node and EEG channels at midline frontal area can be viewed as "effect" nodes in the AON. During intention inference period, both the two conditions elicited directed information transmission from left frontoparital to left paroetooccipital and right frontal regions.

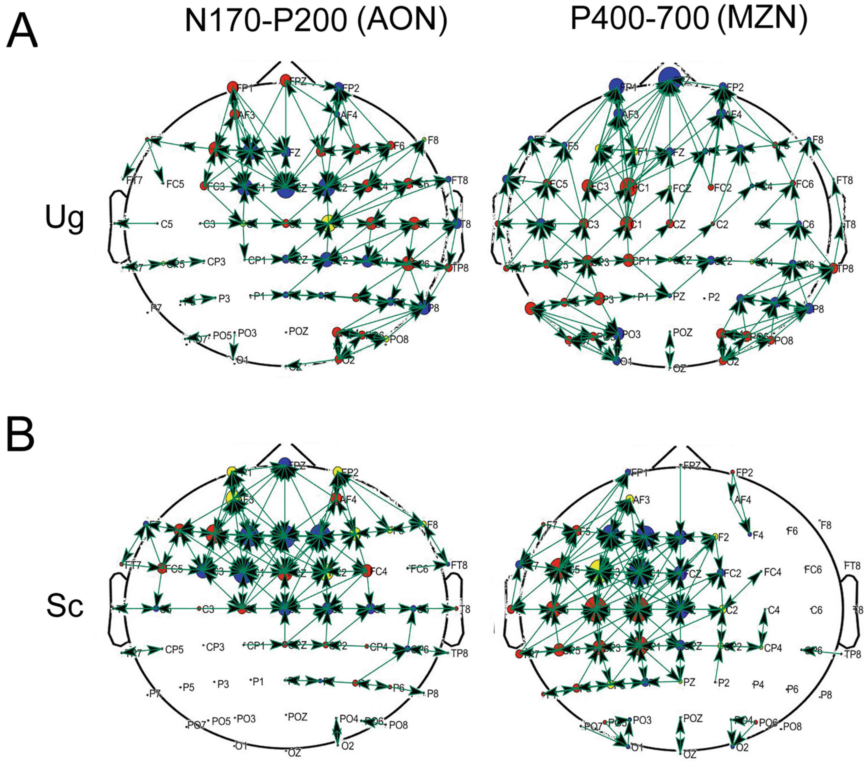

Fig. 4. Channel-based directed networks during N170-P200 and P400-700 time windows of the action intention understanding task. The networks are constructed by setting a fixed threshold for the association matrices of GCs. A red EEG channel represents an outward node with higher outdegree, a blue channel refers to an inward node with higher indegree, and a yellow channel means a node with equal indegree and outdegree in a directed flow network. (Color figure online)

The statistical comparison of the GC connectivity matrices further discovers significant difference in internodal causality of brain networks between intention understanding of intention-oriented usual action and unintelligible action. As shown in Fig. 5A, during low-level perceptual input stage, compared to the AON formed in "Sc" condition, stronger Granger causality are distributed from the nodes at frontal brain area to posterior parietal-occipital nodes in the directed network under "Ug" condition. For usual action, the mirror system might result in direct awareness of the goal of a perceived action [4,5]. Therefore, the visual perception of parietal-occiptal cortex in "Ug" condition elicited denser directed information flow.

In the later inferential process, understanding other's unusual action in "Sc" condition induced stronger causal flow from left inferior frontal gyrus to posterior occipital cortex and from parietal regions to right-lateral frontoparietal nodes, but shows less activity from right inferior frontal gyrus to left frontal cortex (Fig. 5B). This is probably because the observation of the unintelligible actions [4,5]. The MZN is strongly recruited to fill in the "missing" information to judge others' mental states.

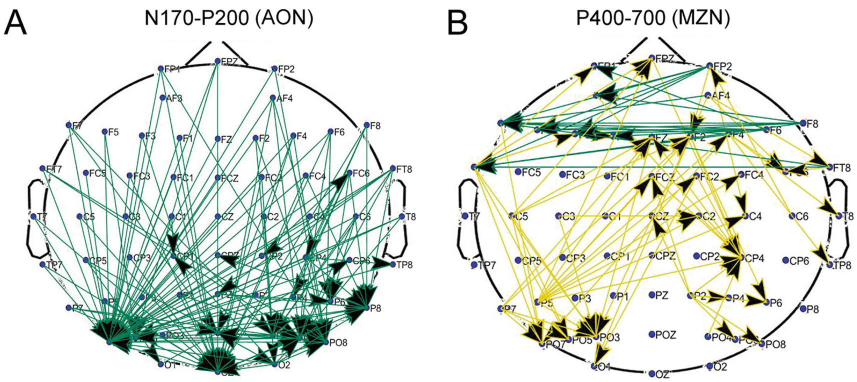

Fig. 5. Topological difference in node pairs of AON and MZN between task conditions. The green edges indicate increased causality and the yellow edges represent decreased causality in "Ug" condition compared to "Sc" condition. The directed links refer to internodal GCs with significant differences in multiple comparisons with a FDR correction ($p < 0.05$). (Color figure online)

According to the significant topological differences in AON and MZN, this study further determine effective EEG channel sites and node parameters for the discrimination between the mental states under "Ug" and "Sc" conditions. The LDA for feature combination of outdegree of channels (F1, FZ, F2) and indegree of EEG channels (PO5, PO7, P6, P8) gets the classification accuracy of 0.7708 in the AON. Additionally, the LDA for feature combination of outdegree of channels (P5, P7) and indegree of channels (FZ, F2, FC4, C4, CP4, P6) acquires the accuracy of 0.6875 (Table 2).

From the discriminant analysis results, we can find that the left-lateral parietal-occipital brain regions act as "effect" nodes with higher inflow connectivity from frontal area during action perception period under "Ug" condition, but play the role of "cause" nodes with higher outflow to right frontoparietal regions under "Sc" condition. Therefore, the brain regions can be viewed as a transition hub from AON to MZN, especially during the intention inference for absence of contextual information of actions or observing unusual actions.

Table 2. EEG feature combination and channel sits in identifying causal connectivity networks for understanding other's usual action (Ug) and unintelligible action (Sc).

Temporal network	AON	MZN
Input feature of specific EEG channel combination	oudegree of F1, FZ and F2 indegree of PO5,PO7, P6 and P8	outdegree of P5 and P7 indegree of FZ, F2, FC4, C4, CP4 and P6
Directed connectivity	Bilateral frontal → posterior parietal-occiptial regions	left parietal regions → right-lateral frontoparietal regions
Classification accuracy (LDA) between task conditions	0.7708	0.6875

4 Conclusions

By constructing the GC-based directed networks in action observation and intention inference period of the brain, our study reveals the transition of causal relationship among brain regions from the early mirror network to the later mentalizing network. In the brain regions involved in information inflow and outflow of action intention understanding, the left-lateral parietal-occipital cortex can be viewed as a hub of the circuit of dynamic information flow. Based on the information transmission from recognizing action kinematics to inferring intentionality, feature extraction of GC-based network nodes was conducted in EEG channel combinations of AON and MZN for discriminating other's usual and unintelligible actions. The EEG channel sites and nodal parameters identified by our study could provide effective features and brain locations for further guiding individual action intention recognition.

Acknowledgements. This work was supported in part by the Natural Science Foundation of Jiangsu Province under Grant BK20221181, the Natural Science Foundation of China under Grants 62077013 and 31900710, and the Fundamental Research Funds for the Central Universities under Grants 2242022k30036 and 2242022k30037.

References

1. Becchio, C., Cavallo, A., Begliomini, C., Sartori, L., Feltrin, G., Castiello, U.: Social grasping: from mirroring to mentalizing. Neuroimage **61**(1), 240–248 (2016)
2. Cacioppo, S., Weiss, R.M., Runesha, H.B., Cacioppo, J.T.: Dynamic spatiotemporal brain analyses using high performance electrical neuroimaging: theoretical framework and validation. J. Neurosci. Meth. **238**, 11–34 (2014)
3. Cacioppo, S., Cacioppo, J.T.: Dynamic spatiotemporal brain analyses using high-performance electrical neuroimaging, part II: a step-by-step tutorial. J. Neurosci. Meth. **256**, 184–197 (2015)
4. Carter, E.J., Hodgins, J.K., Rakison, D.H.: Exploring the neural correlates of goal-directed action and intention understanding. Neuroimage **54**(2), 1634–1642 (2011)

5. Catmur, C.: Understanding intentions from actions: direct perception, inference, and the roles of mirror and mentalizing systems. Conscious. Cogn. **36**, 426–433 (2015)
6. Ruggiero, M., Catmur, C.: Mirror neurons and intention understanding: dissociating the contribution of object type and intention to mirror responses using electromyography. Psychophysiology e13061 (2018)
7. Catmur, C.: Unconvincing support for role of mirror neurons in "action understanding": Commentary on Michael, et al. Front. Hum. Neurosci. **8**, 553 (2014)
8. Kim, S., Yu, Z., Lee, M.: Understanding human intention by connecting perception and action learning in artificial agents. Neural Netw. **92**, 29–38 (2017)
9. Atique, B., Erb, M., Gharabaghi, A., Grodd, W., Anders, S.: Task-specific activity and connectivity within the mentalizing network during emotion and intention mentalizing. Neuroimage **55**(4), 1899–1911 (2011)
10. Gramfort, A., Papadopoulo, T., Olivi, E., Clerc, M.: OpenMEEG: opensource software for quasistatic bioelectromagnetics. Biomed. Eng. Online **9**(1), 45 (2010)
11. Luo, Q., Lu, W., Cheng, W., Valdes-Sosa, P.A., Wen, X., Ding, M.: Spatio-temporal granger causality: a new framework. Neuroimage **79**, 241–263 (2013)
12. Sporns, O.: Contributions and challenges for network models in cognitive neuroscience. Nat. Neurosci. **17**(5), 652–660 (2014)
13. Zhang, L., Gan, J.Q., Zheng, W., Wang, H.: Spatiotemporal phase synchronization in adaptive reconfiguration from action observation network to mentalizing network for understanding other's action intention. Brain Topogr. **31**(3), 447–467 (2018)