Title
Resting-state functional connectivity and deception: exploring individualized deceptive propensity with machine learning

Abbreviated title
Resting-state functional connectivity and individualized deception

Author names and affiliations
Honghong Tang¹,², Xiaping Lu²,³, Zaixu Cui⁴, Chunliang Feng², Qixiang Lin²,³, Xuegang Cui¹, Song Su¹*, and Chao Liu¹,²,³*

¹ Business School, Beijing Normal University, Beijing 100875, China
² State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, 100875, China
³ Center for Collaboration and Innovation in Brain and Learning Sciences, Beijing Normal University, Beijing, 100875, China
⁴ Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, 19104, USA

*Correspondence to:
Song Su
Business School, Beijing Normal University, Beijing 100875, China
E-mail: sus@bnu.edu.cn
Or
Chao Liu
State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, 100875, China
E-mail: liuchao@bnu.edu.cn

Number of pages: 27
Number of figures, tables: 5/2
Number of words for Abstract, Introduction, and Discussion: 239/917/1149
Total number of words: 4264
Abstract

Individuals show a great heterogeneity in determining to be honest or deceptive in daily life. A large number of studies have investigated the neural substrates of deception; however, the brain networks contributed to the individual difference in deception still remain unclear. The current study tried to address this issue by employing a machine-learning approach to predict individuals’ deceptive propensity with topological properties of whole-brain resting-state functional connectivity (RSFC). Participants finished the resting-state functional MRI (fMRI) data acquisition, and then played as proposers in a modified ultimatum game in which they spontaneously chose to be honest or deceptive one week later. A linear relevance vector regression (RVR) model was trained and validated to examine the relationship between topological properties of networks of RSFC and actual deceptive behaviors. Results showed that the machine-learning model sufficiently decoded individual differences of deception by using three brain networks based on RSFC, including the executive controlling network (DLPFC, MFC and OFC), the social and mentalizing network (the temporal lobe, TPJ and IPL), and the reward network (putamen and thalamus). These networks have been found to form a signaling cognitive framework of deception by coding mental states of others and the reward or values of deception or honesty, and integrating this information to make final deceptive or honest decisions. These findings suggest the potentiality in using RSFC as a task-independent neural trait to predict deceptive propensity, and shed
light on using machine-learning approaches in deception detection.

**Keywords:** deception, individual difference, neural trait, machine learning, cross validation, resting-state fMRI

**Abbreviations:** DLPFC, dorsolateral prefrontal cortex; OFC, orbitofrontal cortex; MFC, middle frontal cortex; IPL, inferior parietal lobule; TPJ, temporo-parietal junction; RSFC, resting-state functional connectivity; ND, nodal degree centrality; NE, nodal efficiency; NB, nodal betweenness centrality; ROIs, regions of interests; LOOCV, leave-one-out cross-validation; RVR, relevance vector regression.
Introduction

People show great individual difference in deception when they have opportunities to lie. Some tell lies all the time while others never lie (Gibson et al., 2013; Greene and Paxton, 2009; Tang et al., 2017). Explicit behavioral cues of individual difference have been widely investigated in deception detection (DePaulo et al., 2003; Newman et al., 2003; Riggio and Friedman, 1983). Progress in cognitive neuroscience techniques in the past decade leads to considerable interest in examining the underlying neural correlates of deception and deception detection (Abe, 2011; Jenkins et al., 2016; Kozel et al., 2005; Langleben, 2008; Ofen et al., 2017; Sun et al., 2015).

By comparing difference of neural responses between lying and telling the truth, most empirical studies indicate that deception recruits multiple brain networks. That is, the prefrontal cortex (e.g., superior frontal cortex (SFC), orbitofrontal cortex (OFC), and middle frontal cortex (MFC)), the anterior cingulate cortex (ACC), the subcortical (e.g., striatum, thalamus, etc.) and posterior cortical regions (e.g., premotor and motor cortex, precuneus, inferior parietal lobule (IPL)), the temporal lobe and temporo-parietal junction (TPJ) (Lisofsky et al., 2014; Spence et al., 2004). These networks are considered to interact with each other to shape deception as an integral system. The prefrontal cortex and ACC are engaged in the executive controlling and behavior inhibiting processes in deception (Christ et al., 2009); the subcortical regions are engaged in processing the reward or benefits of deception (Abe and Greene, 2014); the regions in the temporal and parietal lobe are involved in
dealing with social context and mentalizing intentions and mental states of others (Molenberghs et al., 2016).

Researchers also tried to predict individualized deception with these brain networks. Task-dependent neuroimaging studies show that activities of dorsolateral prefrontal cortex (DLPFC), MFC, parietal lobe and subcortical regions are correlated with the frequency of deception (Abe and Greene, 2014; Cui et al., 2018; Dogan et al., 2016; Greene and Paxton, 2009). However, these studies used different paradigms and experimental designs, making it hard to obtain the general neural networks contributed to deception propensity of each individual. Moreover, several studies using structural brain measures found that neocortex size and white matter volumes in the brain could predict individual deceptive propensity (Byrne and Corp, 2004; Yang et al., 2007), suggesting the possibility to use task-independent neural measures to capture individual difference in deception.

The current study aims to decode the heterogeneity in deception from intrinsic whole-brain functional connectivity. We used topological properties of resting-state functional connectivity (RSFC) calculated by graph theory analysis to predict heterogeneity in deceptive behaviors across individuals (i.e., to differentiate individuals with high and low propensity of deception) with a machine-learning approach.

RSFC is used to characterize the patterns of connections between brain regions in a task-independent state (Biswal et al., 1995), which has been proposed to be an effective neural trait measure to predict individual difference in behaviors (Gabrieli
et al., 2015; Nash et al., 2015). First, RSFC, especially the whole-brain RSFC, is stable and reliable, which has high test-retest reliability (Cao et al., 2014; Zuo and Xing, 2014). Second, it could depict the uniqueness of each individual’s brain functional organization, severing as a “fingerprint” in identifying individuals (Finn et al., 2015). Furthermore, it is sufficient to predict mental states and social behaviors such as subjective happiness (Luo et al., 2015), impulsivity in economic decision-making (Li et al., 2013), trust behavior (Hahn et al., 2014), reciprocity of a gift (Cáceda et al., 2015), preference to social information (Zhang and Mo, 2016), and costly punishment for unfair behaviors (Feng et al., 2018). Therefore, we hypothesized that individual difference in deception would be predicted by RSFC patterns between brain networks.

To capture the topological properties of brain networks derived from RSFC, we employed a graph theory analysis and put these properties as features in the prediction of deceptive behaviors. As a framework that depicts organization principles of complex networks, the graph theory could provide a quantitative description of brain with graphs composed of nodes (i.e., regions or voxels) and edges (i.e., functional connectivity between nodes) in several networks (Bullmore and Sporns, 2009; He and Evans, 2010). Three network metrics are commonly used to quantify the topological properties of each network node. That is, the nodal degree centrality (ND) which measures the involvement of a node in the network (Boccaletti et al., 2006; He and Evans, 2010); the nodal efficiency (NE) that estimates the efficiency of parallel information processing of a node in the network
(Achard and Bullmore, 2007); and the nodal betweenness centrality (NB) that assesses the degree of influence a node has on spreading information through the network (Freeman, 1978). These metrics of resting-state brain network has been consistently found to be associated with behaviors (Cao, et al., 2014; Feng, et al., 2018; Liu et al., 2017).

With these network metrics, the current study aims to distinguish individuals with different propensity of deception by a machine-learning approach. The machine-learning approach selects metrics of nodes, then trains predictive models of brain-behavior relationships with training samples and assesses the performance of the model with independent-testing samples through cross-validation (Cui and Gong, 2018; Cui et al., 2018; Cui et al., 2016; Shen et al., 2017). Through these procedures, we hope to obtain the discriminative features in the resting-state brain networks as the neuroimaging makers for deceptive behaviors. Specifically, we hypothesized that metrics of brain networks found in previous deception related studies, i.e., the executive controlling and behavior inhibiting network (e.g., DLPFC, MFC, OFC), the social context and mentalizing network (e.g., the temporal lobe, IPL, TPJ), and the reward network (e.g., putamen, thalamus) would be primarily contributive in the prediction.

**Experimental Procedures**

**Participants**

Fifty-two healthy participants from Beijing Normal University (23 females; \( M_{\text{age}} = 23.3, \text{SD} = 2.15 \)). Two participants misunderstood the task; three participants
reported that they did not know they could lie in the task. Their data were not included in the analysis. All participants reported no history of neurological or psychiatric disorders, signed consents and were paid with money. The Institutional Review Board of the State Key Laboratory of Cognitive Neuroscience and Learning of Beijing Normal University approved this study.

**Procedure**

Firstly, all participants were required to keep still and close their eyes to finish a resting-state fMRI scanning session, which lasted for 6 minutes and 46 seconds. They were instructed to not to think about anything systematically or fall asleep during scanning. One week after that, they finished a behavioral test session, during which they played as the proposer in a modified ultimatum game (Güth et al., 1982; Tang et al., 2016; Tang, et al., 2017). They were instructed to divide a total of money units (randomly chosen from 8, 10, 12, or 14 units in each trial) between themselves and anonymous recipients in 12 trials (Figure 1A). The range of the money units would not be shown to them. At the beginning of each trial, they would randomly gain a total amount of money units (True total) from the computer for the division. Then, they needed to report the amount of the total (Reported total) to the recipient who did not know the total amount, which provided a chance for them to tell a lie. Next, they made a division and then waited for the recipient to accept or reject it. If the recipient accepts the division, participants and the recipient would gain the money units according to the division; if the recipient reject it, all of them gained nothing. In each trial, participants’ divisions were paired with choices from
52 real recipients collected previously in another study with the same experimental paradigm and were told that recipients would never know the true total they gained. We aimed to measure participants’ spontaneous tendency of deception, thus we controlled the rejection rate for each participant with 25% based on real recipients’ responses in the previous study. We run at least two participants at the same time to make they believe that they play with real recipients. Each participant sat in a small room to finish the task alone. All participants were debriefed after the whole task. They answered questions including whether they knew they could spontaneously choose to deceive the recipient in the task, whether recipients would know their true totals, and whether recipients’ choices would affect their payments.

**Image acquisition**

The resting-state fMRI scanning was finished at Beijing Normal University Imaging Center for Brain Research with a Siemens Trio 3 T MRI scanner (Lin et al., 2015; Liu, et al., 2017). Firstly, we obtained participants’ 3D structural images with 144 sagittal slices by a T1-weighted MP-RAGE sequence. The thickness was 1.33 mm; the in-plane resolution was $256 \times 256$; the repeat time (TR) was 2530 ms; the echo time (TE) was 3.39 ms; the inversion time (TI) was 1100 ms; the flip angle was 7°; and the FOV was $256 \times 256$ mm. Next, we sued an echo-planar imaging (EPI) sequence to obtain resting-state MRI data. The axial slices was 33; the thickness was 3.5 mm; the gap was 0.7 mm; the in-plane resolution was $64 \times 64$; the voxel size was $3.1 \times 3.1 \times 3.5$ mm; the repeat time (TR) was 2000 ms; the echo time (TE) was 30 ms; the flip angle was 90°; the field of view (FOV) was $200 \times 200$ mm, and the
volumes was 200.

**Behavioral data analysis**

We calculated the mean deception degree (Deception degree = (True total - Reported total)/True total) as the index of participants’ deceptive tendency. Thus, the higher the deception degree is, the more deceptive the participant.

**Imaging data preprocessing**

SPM 12 (www.fil.ion.ucl.ac.uk/spm) and DPABI software packages (http://rfmri.org/dpabi) (Yan et al., 2016) were used to preprocess fMRI data. Firstly, we removed the first 10 volumes of the functional images to make signals stable, then corrected images with slice timing and realigned them with head motion correction. Next, we co-registered the mean functional image and segmented structure brain image of each participant. After that, we normalized their functional images onto the Montreal Neurological Institute space and resampled images with a voxel size of 3 x 3 x 3 mm. Then the linear trend of the time courses was removed and a band-pass filter (0.01-0.1 Hz) was used to remove low- or high- frequency noise and artifacts (Biswal, et al., 1995; Zuo et al., 2010). A 4 mm FWHM Gaussian kernel was used to smooth images and reduce spatial noises. Finally, the 24 head motion parameters (Fox et al., 2005), the white matter signal, and the cerebrospinal fluid signal (Friston et al., 1996) were regressed out. Since global signal regression might cause ambiguous negative connections, it was not implemented in the preprocessing (Buckner et al., 2009; Xie et al., 2017).

**Network construction**
A graph theoretical analysis was used to obtain the brain network based on resting-state functional connectivity, which depicts the brain as graphs that consisting of nodes (brain regions or voxels) and edges (functional connectivity between nodes) (Bullmore and Sporns, 2009; He and Evans, 2010; Wang et al., 2015; Wang et al., 2010). The GRaph thEoreTical Network Analysis toolbox (GRETNA: http://www.nitrc.org/projects/gretna/) was used to construct and analysis brain networks. Firstly, we used a functional brain atlas with 264 non-overlapping regions of interests (ROIs) to define the whole brain network nodes (Power et al., 2011). This atlas was available in the GRETNA, in which each ROI’s radius is 3 mm. The BOLD signal for all the voxels in each ROI at each time point was averaged as the time course of each node for each participant. To compute the network edges, the Pearson correlation $r$ between ROIs’ time courses in the atlas was calculated, resulting into a $264 \times 264$ functional connectivity map. After transforming the correlation $r$ into Fisher $z$ value, functional brain networks indexed by functional connectivity were binarized to run the network analysis. Positive and negative functional connectivity were analyzed separately in both network construction and network analysis.

**Network analysis**

To characterize the importance of each node in the brain network, three widely used nodal metrics (Nodal degree centrality (ND), Nodal efficiency (NE), Nodal betweenness centrality (NB)) (Figure 2) were calculated based on the binary functional connectivity between nodes in the network (Boccaletti, et al., 2006; He and Evans, 2010). The ND metric refers to the number of edges (connections of a node
connected to other nodes) (Freeman, 1978). The NE metric is the mean of inverted minimum path length between a given node and all other nodes in the network (Achard and Bullmore, 2007). The NB metric is the percentage of all shortest paths of connections that pass through a node (Freeman, 1978). These nodal metrics were used features to predict deception degree in the machine-learning approach as follows.

**Relevance vector regression (RVR)**

Nodal features ND, NE, NB in the brain network were put into a linear relevance vector regression (RVR) algorithm, which was implemented by the function of PRoNTo toolbox (http://www.mlnl.cs.ucl.ac.uk/pronto/) (Schrouff et al., 2013). RVR is a sparse kernel machine-learning algorithm that has been used to predict behavioral measures with resting-state brain data (Gong et al., 2013). Based on a fully probabilistic Bayesian framework (Tipping, 2001), RVR introduces a mean-zero Gaussian prior over the model weights and each weight corresponds to one training sample. These weights are governed by a set of hyperparameters (one for each model weight) (Tipping, 2001), which are estimated iteratively on the training data. Since the posterior distributions of many of the model weights are sharply peaked around zero, those training samples related with non-zero weights are treated as “relevance vectors”. Then, the regression coefficients of all features are determined as the weighted sum of the feature vector of all “relevance vector” samples. For an unseen testing sample, the predicted behavioral score (e.g., predicted deception degree) was defined as the product of these regression coefficients and the feature vector of this sample. Compared to another widely used machine-learning algorithm, the linear
support vector regression (LSVR), RVR has no free parameter and has comparable predictive power and lower computational cost in behavioral predictions with brain functional connectivity (Cui and Gong, 2018).

In the current study, we put ND, NE, NB features into the RVR both separately and together to find the best predictor of behavioral deception degree. We adopted the leave-one-out cross-validation (LOOCV) to evaluate the generalizability of the model. That is, \(N-1\) (\(N\) equals the sample size of participants) participants' data was used as the training data, and the left one was treated as the testing set. Each feature was linearly scaled to the range of 0-1 across the training dataset, and the same scaling parameters were applied to scale the testing set (Cui and Gong, 2018; Cui, et al., 2018). A prediction model was constructed using all the training samples, and was used to predict the scores of the testing sample. The training and testing procedures were repeated \(N\) times so that each subject was used once as the testing sample. The Pearson correlation coefficient \((r)\) and mean absolute error (MAE) between actual and predicted deception degree across all subjects were computed to quantity the accuracy of the prediction (Cui and Gong, 2018; Cui, et al., 2018).

Then a permutation test was used to determine whether the obtained final accuracy metrics (i.e., coefficient \(r\) and MAE) were significantly better than expected by chance. Specifically, the above prediction procedure was re-applied 1,000 times. For each time, we permuted the behavioral scores across the training samples without replacement. The \(P\) value for the prediction performance (for both the correlation \(r\) and MAE) was calculated by dividing the number of times the permuted value was
greater than (or less than, or equal to) the true value by 1000.

For visualization, we colored all ROIs whose model weight in the prediction was higher than 40% of the maximum predicted weight. This threshold predominantly eliminates noise components whose predicted weight was lower than 40% and could visualize the most predictive brain regions (Ecker et al., 2010; Mourão-Miranda et al., 2005).

**Validation**

To validate our prediction results, a 10-fold cross-validation was applied (Cui et al., 2018; Feng et al., 2018). Similar to the LOOCV, this procedure divided all participants into 10 subsets, and used nine of the subsets as the training data and the remaining one as the testing data. It repeated 10 times to let each subset could be used as the testing data once. To avoid the data division affected the prediction performance, we repeated the 10-fold cross-validation for 50 times and used the average results as the final prediction performance. Finally, the significance of the prediction performance was assessed by a 1000 times permutation test.

**Results**

**Behavioral results**

The mean deception degree was 0.15 ($SD = 0.13$). The distribution of deception degree showed that 19.1% participants ($n = 9$) never deceived, and 4.2% participants misreported their gained totals as less than 60% of the true totals, indicating great individual difference in deceptive propensity (Figure 1B). No gender ($t_{45} = 0.40, p = 0.70$) or age ($r = -0.13, p = 0.40$) difference was found in the deception degree.
Prediction results of RVR

Results of prediction with nodal features derived from positive functional connectivity were shown in Figure 3. The predicted deception degree was highly correlated with the actual behavioral deception degree for ND ($r = 0.38, p = 0.009$) and NE ($r = 0.48, p < 0.001$) features, but not for NB features ($r = -0.05, p = 0.76$). Specifically, combination of ND, NE features also led to significant correlation between predicted and actual deception degree ($r = 0.42, p = 0.003$), yet combination of ND, NE, and NB features did not ($r = 0.29, p = 0.051$). Therefore, we focused on reporting the results of the prediction of ND and NE respectively. The permutation test showed that the correlations for ND (Figure 3D) and NE (Figure 3E) were significantly higher than the chance level (permutation tests, $p_{ND} = 0.021$, $p_{NE} = 0.003$).

No significant correlations between predicted and actual deception degree was found for nodal features derived from negative functional connectivity ($rs < 0.26, ps > 0.08$).

Predictive nodal features

There were 33 ND features contributed to the RVR prediction (Figure 4A and Table 1), which were derived from the temporal lobe (i.e., bilateral inferior and middle temporal gyrus (MTG), fusiform), subcortical regions (i.e., left putamen and thalamus), frontal lobe (i.e., bilateral precentral gyrus, right OFC, bilateral superior and medial frontal gyrus (including DLPFC, MFC)), parietal lobe (i.e., left precuneus, IPL, TPJ), the occipital lobe and cerebellum. Similar to ND features, 25 NE features contributed to the RVR prediction (Figure 4B and Table 2), which were also originated from the bilateral temporal lobe and right fusiform, left putamen, bilateral
thalamus, right OFC, bilateral superior and medial frontal gyrus, left precuneus, bilateral parietal lobe, occipital gyrus and cerebellum.

**Validation**

Results of the 10-fold cross-validation showed that the predicted deception degree remained significantly correlated with the actual deception degree for ND (mean $r = 0.33$, permutation test, $p = 0.031$, Figure 5A) and NE (mean $r = 0.43$, permutation test, $p = 0.007$, Figure 5B).

**Discussion**

Great individual difference has been found in deceptive behaviors, but the neural correlates of it remains unclear. The current study employed a machine-learning approach to predict deceptive propensity at individual level by intrinsic brain network in resting state. We found that graph-theoretical topological properties derived from resting-state brain network were able to predict individual deceptive tendency in an independent experiment. Specifically, individual difference in deception were primarily differentiated by nodal (regional) features across several brain networks, including the executive controlling network (e.g., prefrontal cortex), the social and mentalizing network (e.g. temporal lobe, TPJ, IPL), and the reward network (e.g., putamen, thalamus). These findings suggest that deception recruits interactions between multiple brain networks.

Previous task-dependent deception studies consistently found these brain networks engaged in deception (Baumgartner et al., 2009; Bhatt et al., 2010; Greene and Paxton, 2009; Kozel, et al., 2005; Ofen, et al., 2017; Sun, et al., 2015). The
prefrontal cortex (including the DLPFC, MFC and OFC) is usually regarded as the key network for general deception. It is involved in the executive control processes, including memorizing the truth when generating a lie, suppressing a truthful response, and switching behavioral responses between honesty and deception (Christ, et al., 2009; Greene and Paxton, 2009; Spence, et al., 2004). Specifically, stimulating or damaging the DLPFC could increase or decrease deception (Maréchal et al., 2017; Zhu et al., 2014), indicating its critical role in deciding to be deceptive or honest.

As a contrast, the temporal lobe, TPJ and IPL are found to be more specific to interactive deception. They are recruited in socio-cognitive processes in deception, such as reasoning others’ intentions, beliefs and goals in social interaction (Bhatt, et al., 2010; Lisofsky, et al., 2014; Molenberghs, et al., 2016; Tang, et al., 2016; Tang, et al., 2017). A recent study also find these regions were engaged in consolidating social information in the resting state (Meyer et al., 2018), reflecting their functions in representing abstract social information for behaviors.

However, few studies paid attentions to the function of subcortical regions such as striatum and thalamus in deception. Some studies found negative relationship between dopamine D2-receptor availability in striatum and the “Lie” scale scores of social responsibility, which measures individuals’ tendency to overly behave in a socially desirable way (Cervenka et al., 2010; Egerton et al., 2010; Huang et al., 2006; Reeves et al., 2007). They suggest the function of striatum in differentiating individuals with personality traits related to deception. Recently, neural response of anticipated reward in the dorsal striatum was found to be correlated with response of
dishonest benefits in the DLPFC (Abe, 2011; Abe and Greene, 2014), providing direct evidence for the role of subcortical regions in processing reward seeking for deceptive behaviors.

Consistent with these studies, our results provided the first evidence for the role of these three brain networks in deception with topological properties originated from intrinsic resting-state functional connectivity. These networks interact with each other and then might shape deception in a signaling cognitive framework (Jenkins, et al., 2016). Firstly, the social and mentalizing network represents the characteristics and mental state of both receivers and themselves, the potential actions and potential results related to deception in the deceivers’ mind. Then the reward network decodes the reward or values of different potential actions (e.g., to be honest or deceptive) and forms motivations for each actions. Finally, the two networks send the social and motivational information to the central executive system, which makes the final decision to be deceptive or honest then generate the final behavior (Abe, 2011; Abe and Greene, 2014; Lisofsky, et al., 2014; Spence, et al., 2004). Furthermore, these networks are also active in task-dependent deception studies, implying that neural networks in shaping personality traits related to deceptive propensity in task-independent state may be overlapped rather than being dissociated with networks engaged in making actual decisions to be deceptive or honest (Abe, 2011; Cervenka, et al., 2010; Egerton, et al., 2010; Huang, et al., 2006; Reeves, et al., 2007).

Interestingly, we did not find any topological properties of the anterior cingulate cortex (ACC) in prediction of deceptive propensity. Although both prefrontal cortex
and ACC are greatly active in monitoring cognitive conflict and inhibiting response in deception, their functions are different. The prefrontal cortex is a more general network for different types of deception, whereas the ACC is more specific to deception that requires participants to monitor cognitive conflict to pretend not to know the truth (Abe et al., 2005), or deception that associated with emotional response (Baumgartner, et al., 2009; Kozel, et al., 2005). Therefore, ACC in the task-independent state might not be contributive to deceptive propensity.

In addition, results in the current study demonstrate that RSFC could be used to represent individual difference in social preference and behaviors, which are in line with studies that use RSFC to predict impulsivity (Li, et al., 2013), trust behavior (Hahn, et al., 2014), reciprocity (Cáceda, et al., 2015), and costly punishment (Feng, et al., 2018). These findings suggest the potentiality in using RSFC as a task-independent neural trait to predict deceptive propensity as neocortex size and white matter volumes in the brain (Byrne and Corp, 2004; Nash, et al., 2015; Yang, et al., 2007). Specifically, we employed a machine-learning approach to perform the prediction of brain-behavior relationship at the individual level, which is advantageous in generalizing results into independent new data (Cui and Gong, 2018; Shen, et al., 2017). The usage of this approach not only support the potentiality of predicting social preference and behavior with brain neuroimaging data (Cui, et al., 2018; Feng, et al., 2018), but also shed light on the application of neuroimaging markers in deception detection (Langleben, 2008).

Several limitations should be considered when generalizing and extending results
in this study. First, the behavioral deception degree in the current study captures the deceptive propensity in interactive context, future studies are needed to examine the neural networks underlying both interactive and non-interactive context (Lisofsky, et al., 2014). Second, our results only demonstrate the neural networks associated with actual deception. Future studies combine neural networks contributive to both actual deception and personality traits related to deception would provide more information about how to use neural trait to predict deceptive propensity (Cervenka, et al., 2010; Egerton, et al., 2010; Huang, et al., 2006; Nash, et al., 2015; Reeves, et al., 2007).

Thirdly, future studies that focus on the temporal and directional relationship among these networks would demonstrate how they interact with each other in making decisions about deception. Finally, future studies are encouraged to combine the topological properties and other properties of neural networks, and other neuroimaging measures, such as T1-weighted and diffusion-weighted brain data, to achieve a better and reliable network associated with deceptive propensity.

Taken together, this current study indicate that brain networks based on intrinsic resting-state functional connectivity are sufficient to predict individual difference in deceptive propensity. It sheds light on using the machine-learning approach in deception detection and identifying neural trait associated with social preference and behaviors.

**Acknowledgments**

This work was supported by the National Key R&D Program of China (2017YFC0803402), the National Natural Science Foundation of China (NSFC) (31871094, 31800958, 71572018, 71872016, 71772015), the China Postdoctoral
Science Foundation (2018M630103), the Beijing Municipal Science & Technology Commission (Z151100003915122), the National Program for Support of Top-notch Young Professionals.

Reference


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


Christ SE, Van Essen DC, Watson JM, Brubaker LE, McDermott KB (2009), The contributions of...


DePaulo BM, Lindsay JJ, Malone BE, Muhlenbruck L, Charlton K, Cooper H (2003), Cues to deception. Psychol Bull 129:74. doi:10.1037/0033-2909.129.1.74


doi:10.1016/j.neuropsychologia.2015.01.013


Xie S, Yang J, Zhang Z, Zhao C, Bi Y, Zhao Q, Pan H, Gong G (2017), The effects of the x chromosome...
Zhu L, Jenkins AC, Set E, Scabini D, Knight RT, Chiu PH, King-Casas B, Hsu M (2014), Damage to dorsolateral prefrontal cortex affects tradeoffs between honesty and self-interest. Nat Neurosci 17:1319-1321. doi:10.1038/nn.3798

Figure Legends

Figure 1.  A) Task procedure. All participants played as proposers (P) who could choose to tell the responder (R) the true total of allocation or not in the modified ultimatum game. B) Distribution of participants’ actual deception degree in the task.

Figure 2. The prediction schematic flow using the nodal features extracted from resting-state brain network with graph theoretical analysis. Panels A-D showed the steps of data analysis.

Figure 3. Results of prediction of deception degree for each participant using relevance vector regression (RVR) algorithm and leave-one-out cross-validation (LOOCV). Panels A-C shows the correlation between actual and predicted deception
degree derived from Nodal degree centrality (ND) features (A), Nodal efficiency (NE) features (B) and Nodal betweenness centrality (NB) features (C) respectively. Panels D-E presents the permutation distribution of ND (D) and NE (E) features’ mean absolute error (MAE). The black solid lines indicated the true MAE in the prediction.

**Figure 4.** Contributive nodal features in prediction of individualized deception degree (visualized by BrainNet Viewer (Xia et al., 2013)). A) Primarily contributive Nodal degree centrality (ND) features in RVR prediction. B) Primarily contributive Nodal efficiency (NE) features in RVR prediction.

**Figure 5.** Permutation distribution of prediction using 10-fold CV. A). Permutation distribution of ND features’ mean absolute error (MAE). B) Permutation distribution of NE features’ mean absolute error (MAE). The black solid lines indicated the mean MAE in the prediction using 10-fold CV.
Table 1. MNI coordinates and weights of primarily contributive nodal degree centrality (ND) features in RVR prediction. L, lfet, R; right.

<table>
<thead>
<tr>
<th>ROI</th>
<th>Hemisphere</th>
<th>MNI coordinate</th>
<th>Metric</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Temporal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inferior temporal gyrus</td>
<td>L</td>
<td>-50 -7 -39</td>
<td>ND</td>
<td>0.201</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>46 -47 -17</td>
<td>ND</td>
<td>0.126</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>-68 -41 -5</td>
<td>ND</td>
<td>0.126</td>
</tr>
<tr>
<td>Middle temporal gyrus</td>
<td>L</td>
<td>-56 -13 -10</td>
<td>ND</td>
<td>0.109</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>65 -31 -9</td>
<td>ND</td>
<td>0.085</td>
</tr>
<tr>
<td>Fusiform</td>
<td>L</td>
<td>-31 -10 -36</td>
<td>ND</td>
<td>0.088</td>
</tr>
<tr>
<td><strong>Subcortical</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Putamen</td>
<td>L</td>
<td>-31 -11 0</td>
<td>ND</td>
<td>0.171</td>
</tr>
<tr>
<td>Thalamus</td>
<td>L</td>
<td>-10 -18 7</td>
<td>ND</td>
<td>0.111</td>
</tr>
<tr>
<td><strong>Frontal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Precentral gyrus</td>
<td>L</td>
<td>-38 -15 69</td>
<td>ND</td>
<td>0.159</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>38 -17 45</td>
<td>ND</td>
<td>0.102</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>-32 -1 54</td>
<td>ND</td>
<td>0.092</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>8 48 -15</td>
<td>ND</td>
<td>0.131</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>8 41 -24</td>
<td>ND</td>
<td>0.113</td>
</tr>
<tr>
<td>Orbitofrontal cortex</td>
<td>R</td>
<td>6 67 -4</td>
<td>ND</td>
<td>0.105</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>34 38 -12</td>
<td>ND</td>
<td>0.098</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>13 -1 70</td>
<td>ND</td>
<td>0.122</td>
</tr>
<tr>
<td>Superior frontal gyrus</td>
<td>L</td>
<td>-16 -5 71</td>
<td>ND</td>
<td>0.103</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>10 -17 74</td>
<td>ND</td>
<td>0.090</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>47 10 33</td>
<td>ND</td>
<td>0.101</td>
</tr>
<tr>
<td>Middle frontal gyrus</td>
<td>R</td>
<td>43 49 -2</td>
<td>ND</td>
<td>0.094</td>
</tr>
<tr>
<td>Medial frontal gyrus</td>
<td>L</td>
<td>-8 48 23</td>
<td>ND</td>
<td>0.100</td>
</tr>
<tr>
<td>Dorsolateral prefrontal cortex</td>
<td>R</td>
<td>13 55 38</td>
<td>ND</td>
<td>0.097</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>-42 38 21</td>
<td>ND</td>
<td>0.089</td>
</tr>
<tr>
<td><strong>Parietal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Precuneus</td>
<td>L</td>
<td>-7 -52 61</td>
<td>ND</td>
<td>0.112</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>-16 -77 34</td>
<td>ND</td>
<td>0.089</td>
</tr>
<tr>
<td>Postcentral gyrus</td>
<td>L</td>
<td>-54 -23 43</td>
<td>ND</td>
<td>0.107</td>
</tr>
<tr>
<td>Temporoparietal junction</td>
<td>R</td>
<td>47 -50 29</td>
<td>ND</td>
<td>0.095</td>
</tr>
<tr>
<td>Angular</td>
<td>L</td>
<td>-39 -75 44</td>
<td>ND</td>
<td>0.090</td>
</tr>
<tr>
<td>Inferior parietal lobule</td>
<td>R</td>
<td>33 -53 44</td>
<td>ND</td>
<td>0.084</td>
</tr>
<tr>
<td>Paracentral lobule</td>
<td>L</td>
<td>-7 -33 72</td>
<td>ND</td>
<td>0.081</td>
</tr>
<tr>
<td><strong>Occipital</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle occipital gyrus</td>
<td>R</td>
<td>42 -66 -8</td>
<td>ND</td>
<td>0.093</td>
</tr>
<tr>
<td><strong>Cerebellum</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cerebelum</td>
<td>R</td>
<td>35 -67 -34</td>
<td>ND</td>
<td>0.090</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>-32 -55 -25</td>
<td>ND</td>
<td>0.086</td>
</tr>
</tbody>
</table>
Table 2. MNI coordinates and weights of primarily contributive nodal efficiency (NE) features in RVR prediction. L, left; R, right.

<table>
<thead>
<tr>
<th>ROI</th>
<th>Hemisphere</th>
<th>MNI coordinate</th>
<th>Metric</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Temporal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inferior temporal gyrus</td>
<td>L</td>
<td>-50 -7 -39</td>
<td>NE</td>
<td>0.213</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>46 -47 -17</td>
<td>NE</td>
<td>0.135</td>
</tr>
<tr>
<td>Middle temporal gyrus</td>
<td>L</td>
<td>-68 -41 -5</td>
<td>NE</td>
<td>0.108</td>
</tr>
<tr>
<td>Fusiform</td>
<td>R</td>
<td>33 -12 -34</td>
<td>NE</td>
<td>0.125</td>
</tr>
<tr>
<td><strong>Subcortical</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Putamen</td>
<td>L</td>
<td>-31 -11 0</td>
<td>NE</td>
<td>0.149</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>-2 -13 12</td>
<td>NE</td>
<td>0.129</td>
</tr>
<tr>
<td>Thalamus</td>
<td>R</td>
<td>12 -17 8</td>
<td>NE</td>
<td>0.103</td>
</tr>
<tr>
<td><strong>Frontal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medial orbitofrontal cortex</td>
<td>R</td>
<td>6 67 -4</td>
<td>NE</td>
<td>0.142</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>-16 -5 71</td>
<td>NE</td>
<td>0.108</td>
</tr>
<tr>
<td>Superior frontal gyrus</td>
<td>R</td>
<td>10 -17 74</td>
<td>NE</td>
<td>0.105</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>13 -1 70</td>
<td>NE</td>
<td>0.102</td>
</tr>
<tr>
<td>Precentral gyrus</td>
<td>L</td>
<td>-38 -15 69</td>
<td>NE</td>
<td>0.102</td>
</tr>
<tr>
<td>Medial frontal gyrus</td>
<td>L</td>
<td>-3 26 44</td>
<td>NE</td>
<td>0.094</td>
</tr>
<tr>
<td>Dorsolateral prefrontal cortex</td>
<td>R</td>
<td>43 49 -2</td>
<td>NE</td>
<td>0.094</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>-42 38 21</td>
<td>NE</td>
<td>0.099</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>31 33 26</td>
<td>NE</td>
<td>0.087</td>
</tr>
<tr>
<td><strong>Parietal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Superior parietal lobule</td>
<td>L</td>
<td>-16 -46 73</td>
<td>NE</td>
<td>0.131</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>-28 -58 48</td>
<td>NE</td>
<td>0.104</td>
</tr>
<tr>
<td>Inferior parietal lobule</td>
<td>R</td>
<td>33 -53 44</td>
<td>NE</td>
<td>0.107</td>
</tr>
<tr>
<td>Temporoparietal junction</td>
<td>R</td>
<td>47 -50 29</td>
<td>NE</td>
<td>0.086</td>
</tr>
<tr>
<td>Precuneus</td>
<td>L</td>
<td>-16 -77 34</td>
<td>NE</td>
<td>0.088</td>
</tr>
<tr>
<td><strong>Occipital</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lingual gyrus</td>
<td>R</td>
<td>17 -91 -14</td>
<td>NE</td>
<td>0.111</td>
</tr>
<tr>
<td>Middle occipital gyrus</td>
<td>L</td>
<td>-47 -76 -10</td>
<td>NE</td>
<td>0.098</td>
</tr>
<tr>
<td><strong>Cerebellum</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cerebelum</td>
<td>L</td>
<td>-32 -55 -25</td>
<td>NE</td>
<td>0.106</td>
</tr>
</tbody>
</table>
Figure 1

A

| Pairing... | The total amount is 10 | The total amount is | The offer is | Waiting for R's choice | P | R |

B

![Graph showing deception degree against subject number](image)

Subject number 0 to 50

Deception degree

Figure 2

A Network construction

B Network analysis

C Extracting nodal features

D Predicting behaviors

Resting-state brain data

Resting-state functional connectivity (RSFC)

Actual behaviors

Predicted behaviors
Figure 3

A Nodal degree centrality

\[ r_{nd} = 0.38, \ p = 0.009 \]

B Nodal efficiency

\[ r_{ce} = 0.48, \ p < 0.001 \]

C Nodal betweenness centrality

\[ r_{be} = 0.05, \ p = 0.76 \]

D

\[ p_{\text{permutation ND}} = 0.021 \]

E

\[ p_{\text{permutation NE}} = 0.003 \]
Figure 4

A Nodal degree centrality

B Nodal efficiency

Figure 5

A \( P_{\text{permutation 10-fold NO}} = 0.031 \)

B \( P_{\text{permutation 10-fold NE}} = 0.007 \)
Graphical abstract

A Network construction

B Network analysis

C Extracting nodal features

D Predicting deceptive propensity

Resting-state brain data

Nodal degree centrality (ND)

Nodal efficiency (NE)

Atlas

Resting-state functional connectivity (RSFC)
Highlights

1. The neural substrates of individualized deceptive propensity could be captured by resting-state functional connectivity.

2. A relevance vector regression machine-learning approach was used to predict individualized deceptive propensity based on RSFC.

3. Individualized deception was predicted by the executive controlling, the social and mentalizing, and the reward network.

4. These networks have been found to form a signaling cognitive framework of deception in task-dependent fMRI studies.