WILEY

RESEARCH ARTICLE

I lie, why don't you: Neural mechanisms of individual differences in self-serving lying

Lijun Yin[1,2,3] | Bernd Weber[4,5]

[1]Department of Psychology, Sun Yat-sen University, Guangzhou, Guangdong, China

[2]Guangdong Provincial Key Laboratory of Social Cognitive Neuroscience and Mental Health, Sun Yat-sen University, Guangzhou, Guangdong, China

[3]Guangdong Provincial Key Laboratory of Brain Function and Disease, Sun Yat-sen University, Guangzhou, Guangdong, China

[4]Center for Economics and Neuroscience, University of Bonn, Bonn, Germany

[5]Institute of Experimental Epileptology and Cognition Research, University Hospital Bonn, Bonn, Germany

Correspondence
Lijun Yin, Department of Psychology, Sun Yat-sen University, Guangzhou 510006, China
Email: ylj.eagle@gmail.com

Abstract

People tend to lie in varying degrees. To advance our understanding of the underlying neural mechanisms of this heterogeneity, we investigated individual differences in self-serving lying. We performed a functional magnetic resonance imaging study in 37 participants and introduced a color-reporting game where lying about the color would in general lead to higher monetary payoffs but would also be punished if get caught. At the behavioral level, individuals lied to different extents. Besides, individuals who are more dishonest showed shorter lying response time, whereas no significant correlation was found between truth-telling response time and the degree of dishonesty. At the neural level, the left caudate, ventromedial prefrontal cortex (vmPFC), right inferior frontal gyrus (IFG), and left dorsolateral prefrontal cortex (dlPFC) were key regions reflecting individual differences in making dishonest decisions. The dishonesty associated activity in these regions decreased with increased dishonesty. Subsequent generalized psychophysiological interaction analyses showed that individual differences in self-serving lying were associated with the functional connectivity among the caudate, vmPFC, IFG, and dlPFC. More importantly, regardless of the decision types, the neural patterns of the left caudate and vmPFC during the decision-making phase could be used to predict individual degrees of dishonesty. The present study demonstrated that lying decisions differ substantially from person to person in the functional connectivity and neural activation patterns which can be used to predict individual degrees of dishonesty.

KEYWORDS

deception, fMRI, individual differences

1 | INTRODUCTION

People lie for many different reasons and quite often for self-serving goals. One might lie to earn an extra profit, for example, by declaring less income during tax reporting. However, this is not the full picture. Imagine facing a situation where we can lie to benefit ourselves by incurring relatively much fewer costs, would lying be a default behavior in this case for everyone? People make different decisions in the same situation (Ennis, Vrij, & Chance, 2008; Gozna, Vrij, & Bull, 2001; Grolleau, Kocher, & Sutan, 2016). Even when lies lead to higher payoffs and no punishment is expected, individuals do not lie all the time (Grolleau et al., 2016; Yin, Hu, Dynowski, Li, & Weber, 2017; Yin, Reuter, & Weber, 2016). Unconditional liars and honest people exist, while the honesty of some other individuals seems to be influenced

by intrinsic lying costs (Rosenbaum, Billinger, & Stieglitz, 2014). Therefore, people lie to a different extent and individual differences are a vital factor in understanding the deceptive decision-making process.

Among neuroimaging studies about lying, two major types of experimental paradigms were frequently used (Yin et al., 2016). The first type is called "instructed lying." In this paradigm, the experimenters instructed participants to make true or untrue statements mainly by providing lying or truth-telling cues (Abe et al., 2008; Ofen, Whitfield-gabrieli, Chai, Schwarzlose, & Gabrieli, 2017; Spence et al., 2001; Sun, Lee, & Chan, 2015). The major advantage of this type of paradigm is that researchers can easily design the experiment and analyze neuroimaging data. The findings from this paradigm are relatively consistent. The "dishonest" responses generally activated the bilateral dorsolateral and ventrolateral prefrontal cortex (dlPFC and

vlPFC), medial superior frontal cortex, anterior insula, anterior cingulate cortex (ACC), inferior parietal lobule (IPL), and posterior parietal cortex (Christ, Van Essen, Watson, Brubaker, & McDermott, 2009; Farah, Hutchinson, Phelps, & Wagner, 2014). Most of these regions are believed to play an important role in executive control. Despite the consistent findings, the "instructed lying" paradigm suffered from several drawbacks, such as weak ecological validity and failure in capturing individual differences in making dishonest decisions.

To investigate lying in a more ecologically valid manner, a paradigm of "spontaneous lies" has been used lately. In this type of paradigm, participants are allowed to decide whether to lie. Unlike the findings in instructed lying studies, mixed or contradicting findings were found across more ecologically valid studies. For example, the dlPFC is usually found activated when participants were instructed to lie (Christ et al., 2009; Luan Phan et al., 2005; Nunez, Casey, Egner, Hare, & Hirsch, 2005). However, when participants made their own decision, both lying and truth-telling are associated with higher activation in the dlPFC (Abe & Greene, 2014; Baumgartner, Fischbacher, Feierabend, Lutz, & Fehr, 2009; Greene & Paxton, 2009; Sip et al., 2010; Yin et al., 2016; Zhu et al., 2014).

The mixed findings in a more ecologically valid paradigm might be due to individual differences in deceptive decision-making at both behavioral and neural levels. At the behavioral level, individual differences in lying might be reflected in different lying frequencies and response time across individuals. At the neural level, individual differences in lying might lead to different activation patterns in brain regions, especially those associated with cognitive control, and functional connectivity among them. Potential regions which might be influenced by individual differences in lying frequencies could be identified by the comparison between lying and truth-telling, especially in dishonest individuals who make their own decisions. Previous studies have found that lying activated the vlPFC, dlPFC, and IPL, as compared to truth-telling (Greene & Paxton, 2009; Yin et al., 2016). These regions belong to the frontoparietal network that is believed to have a strong association with cognitive control (Christ et al., 2009; Farah et al., 2014). However, without considering different lying frequencies across participants, a direct comparison between lying and truth-telling might miss key information. For example, when comparing the neural response between honest and dishonest individuals, the involvement of different neural mechanisms involved can be observed. Extra cognitive resources (higher involvement of the dlPFC, vlPFC, and IPL) would be spent on making honest decisions in more dishonest individuals, whereas honest responses are closer to the default option for more honest individuals (Yin et al., 2016). If correlations between neural patterns and traits or personal values of honesty were investigated, different neural patterns were found. Participants with higher psychopathic traits showed neural activity and functional connectivity changes in the vlPFC-insula and cerebellum networks, which could be used to predict their improved performance of lying after practice (Shao and Lee, 2017). For participants who valued honesty highly, a stronger coupling among cognitive control associated regions (the dlPFC, dorsomedial prefrontal cortex, and inferior frontal cortex) was observed when economic costs for lying were high (vs. low) (Dogan et al., 2016). From these findings, we know that variations in lying behaviors can be associated with variations in neural

patterns, and these can be experimentally observed and studied. Without taking individual differences into consideration, group-level analyses might result in misidentification of key regions or misinterpretation of observed findings.

In addition to cognitive control regions, activity in reward and value-coding related brain regions can show variations across individuals. When lying decisions accompany monetary consequences, the process of monetary rewards and estimation of subjective value might recruit value coding associated brain regions, such as the striatum (the nucleus accumbens [NAcc], caudate, and putamen) and the ventromedial prefrontal cortex (vmPFC) (Diekhof, Kaps, Falkai, & Gruber, 2012; Galvan et al., 2005; Kim, Shimojo, & O'Doherty, 2010; O'Doherty, 2004). First, the striatum and vmPFC react to lies and truth with different aims and consequences. Individuals apply different values on lying behaviors out of different aims (Lundquist, Ellingsen, Gribbe, & Johannesson, 2009), and the vmPFC responds differently while making honest decisions for either oneself or a higher social goal, such as a charity organization (Yin et al., 2017). The striatum and vmPFC are also sensitive to gains and losses caused by dishonest or honest acts. When facing lies and the truth, the activity in the NAcc was modulated by beneficial or harmful monetary outcomes (Yin & Weber, 2016) and the ventral striatum responds to the outcome evaluation of lies (Sun, Chan, Hu, Wang, & Lee, 2015). Second, the value coding regions are sensitive to individual differences in lying. Individuals with stronger NAcc responses in anticipating rewards showed a stronger dlPFC response when refraining from lying to get higher monetary gains (Abe & Greene, 2014). High self-serving-biased individuals had stronger striatum activity when they had a chance to lie for themselves compared to lie for others (Pornpattananangkul, Zhen, & Yu, 2018).

Taken together, both cognitive control and subjective value coding regions seem to not only differentiate between the process of lying and truth-telling but also reflect individual differences in lying. However, few studies focus on individual differences in lying behaviors and use variations in the neural patterns to predict variations in the lying behaviors, despite the fact that it is important for us to understand lying decision-making. To investigate the individual difference in lying and shed light on the function of potential discrete brain regions found in previous studies, we performed a functional magnetic resonance imaging (fMRI) study. Our study aims to investigate lying decision-making from the neural perspectives, and in particular bridge the connection between lying frequencies and neural patterns, and furthermore predict lying frequencies from neural patterns. In previous studies, differences in neural correlates were found between lying for oneself and lying for others (Cui et al., 2017; Pornpattananangkul et al., 2018; Yin et al., 2017). Given that there are many different types of lies, we would like to restrict our research topic to self-serving lies, which are one of the most common types of lies. We employed a revised version of a color-reporting game (Karton & Bachmann, 2011). In the original game, participants were instructed to report the color shown on a screen either correctly or incorrectly. In the revised game, to induce lying behavior in a more ecologically valid manner, we set the rules such that misreporting the color would lead to higher payoffs compared to the case when the color is reported correctly. We also set punishment rules. In 20% of the trials, the

computer checks participants' responses. Every detected lie would be punished by taking away additional money from participants. The advantages of using the current experimental design are as follows. First, our design can be considered as more ecologically valid, compared to previously used paradigms, including instructed lies and some spontaneous lying paradigms. In real life, people usually make their own decision about lying or telling the truth, which can not be captured by instructed lying paradigms. Moreover, while lies often come with potential benefits, there are also certain associated risks of punishment when lies get caught. The punishment is usually lacking in spontaneous lying paradigms. The current design allows participants to make decisions according to their own will to tell the truth for less profit or choose to lie for greater profits while running the risk of being punished. Second, our design is more appropriate to investigate individual differences. In the design of instructed lying, since (dis)honest response might not be the one that participants prefer to choose if free decisions were allowed, confounding processes might be involved, resulting in mixed neural patterns from different resources. In this case, individual differences are hard to identify. In addition, we combined both traditional univariate analyses and multivariate pattern analysis (MVPA) on neuroimaging data to get a broader view on the underlying data. The current design allows us to distinguish lies from truth-telling at the individual level and to further extract lying frequencies under our experimental context. Based on this, univariate analyses and connectivity analyses could be applied to identify brain regions, in which lying associated activity relates to the probability of lying and functional connectivity reflects individual differences in dishonesty.

MVPA (Haxby, 2012; Norman, Polyn, Detre, & Haxby, 2006) is another useful method to investigate the neural mechanisms of lying and individual difference in lying. The MVPA method makes it possible to decode different cognitive states from patterns of brain activation. Previously, MVPA has been used to decode reward-based behavioral choices (Hampton & O'Doherty, 2007) and social values (Chavez & Heatherton, 2015). More relevant to lying, MVPA was used to decode true thoughts independent of intentions to lie (Yang et al., 2014). The authors found that the superior temporal gyrus, left supra-marginal gyrus, and left middle frontal gyrus (MFG) showed high decoding accuracies. In a mock murder crime study, producing lies recruited cognitive control (bilateral IPL and right vlPFC) and reward-related regions (right striatum) (Cui et al., 2014). The activity of the vlPFC contributed mostly to discern "murderers" from "innocents." However, the aforementioned pattern classification method requires sufficient honest and dishonest decisions to facilitate the learning and classification process. In more ecologically valid studies where individuals are free to make their own decisions, variations in lying frequencies might make the data fail to meet the prerequisite of pattern classification analysis and therefore limit the usage of the method. An alternative is to use pattern regression analysis, which could help us to further relate individual differences in behaviors to neural patterns. It can be used to test whether the neural patterns in the decision-making phase (regardless of the decisions made) predict individual differences in dishonesty. On the one hand, the multivariate nature of the pattern regression analysis allows the detection of subtle effects. On the other hand, it utilizes the whole data set from the phase when individuals make decisions, learn from a subset of participants and predict unobserved individuals. Therefore, by using our design, we obtained lying frequencies for each participant, and used the neural pattern during the whole decision-making phase to predict variations in lying behaviors.

We expected that at the behavioral level, there would be individual differences in the frequencies of lying (degree of dishonesty) and response time. At the neural level, by using both univariate and MVPAs, we would like to investigate: (a) brain regions in which lying and truth-telling associated activity relate to the probability of lying; (b) functional connectivity in regions which reflects individual differences in dishonesty, and (c) brain patterns during the decision-making phase in the regions which can be used to predict individual differences in dishonesty. We expected that the activation patterns and functional connectivity in cognitive control (the dlPFC and inferior frontal gyrus [IFG] in particular) and subjective value coding regions (the striatum and vmPFC in particular), correlate with individual differences in making self-serving lies. Specifically, lying associated activity in the dlPFC, IFG, vmPFC, and caudate would correlate with lying frequencies, since previous studies have found that these regions were not only found activated in the comparison of lying and truth-telling but also showed their vulnerability of being influenced by individual differences. Furthermore, the functional connectivity between cognitive control regions and reward associated regions would also be moderated by the frequencies of lying. Finally, the activity in cognitive control regions and reward associated regions could help to predict individuals' lying frequencies.

2 | METHODS

2.1 | Participants

Forty-three participants were enrolled in the fMRI experiment. Data from six participants were excluded. Three participants were excluded due to excessive head movements (>3 mm or >3°), two failed to pass the test trials during the experiment (accuracy <70%, please see Section 2.2 for more details), and one participant misunderstood the experiment instructions. The remaining 37 participants (26 females) ranged from 19 to 35 years of age (mean $\pm$ SD = 24.27 $\pm$ 3.49). All participants were not colorblind and had normal or corrected-to-normal vision. They all reported that they had no prior history of psychiatric or neurological disorders. All participants provided their informed consent and the study was approved by the Ethics Committee of the University of Bonn.

2.2 | Tasks

In the MRI scanner, participants played a color-reporting game (Figure 1a). They first saw a circle with a random color (blue or yellow) on the screen and they were required to memorize the color of the shown circle. After a few seconds, they saw four circles with different colors and the following question: "which color matches the previous one." They were supposed to choose the one that they saw in the last screen. To create the opportunity of lying to get higher payoffs, we
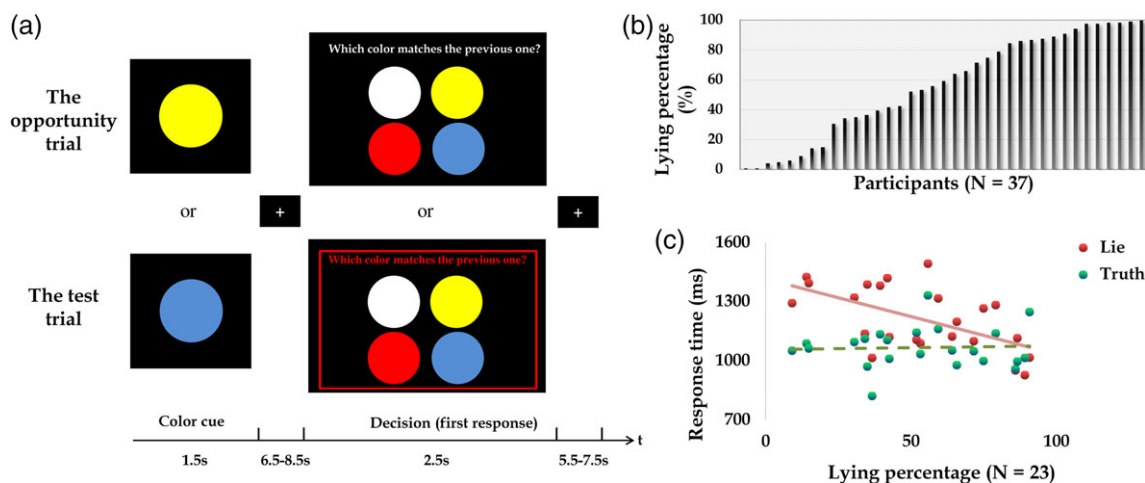
FIGURE 1 (a) Illustration of experimental paradigm. Upper panel: in the opportunity trials, participants first memorized the color of the circle (yellow or blue), and choose the color they previously saw. Misreporting the color earned participants more money than reporting correctly. Lower panel: in the test trials, participants were required to report the color correctly. (b) Distribution of lying percentages across participants (N = 37). (c) Negative correlations were found between lying percentages and reaction time of lying (red dots and solid line; Spearman correlation: $r = -0.45$, $p = .03$). No significant correlation was found between lying percentages and reaction time of truth-telling (green dots and dashed line; Spearman correlation: $r = -0.02$, $p = .92$; N = 23) [Color figure can be viewed at wileyonlinelibrary.com]

told participants that misreporting will earn them more money: if the reported color does not match with the previous color (i.e., lying), participants got 8 points. If the reported color matches with the previous color (i.e., telling the truth), participants got 1 point. After the experiment, the computer randomly chose 20% of all the trials to check if the answers matched with the previously displayed colors. The participants were additionally told that they would be punished if the computer detects incorrect answers. For every detected incorrect answer, the points participants won in that trial would be withdrawn, and additionally, 6 points would be deduced. There were 132 trials in total, 120 trials were opportunity trials described above (Figure 1a, upper panel) and 12 trials were test trials marked with a red frame (Figure 1 (a), lower panel). In the test trials, participants were required to report the color correctly. Before the experiment, they completed a manipulation check to ensure that they fully understood the payment rules. At the end of the experiment, every point would be converted to 0.05 €. After the experiment, participants answered the following question: "how much do you agree with the following statement: misreporting the color during the experiment is a 'lie'," based on a 9-point scale (1: strongly disagree; 5: neural; 9: strongly agree). In the end, we summed up the total points and paid participants accordingly. In addition, participants got extra 5€ as a participation fee. The whole experiment lasted about 45 min.

2.3 | Image acquisition

All images were run on a Siemens Trio 3.0-Tesla scanner with a standard 32-channel head coil. Structural scans included T1-weighted images (TR = 1,660 ms; TE = 2.54 ms; flip angle = 9°; slice thickness = 0.8 mm). One functional session was run which started with a localizer scan and was then followed by the paradigm implemented in presentation (Neurobehavioral Systems; http://www.neurobs.com) during which T2*-weighted echo planar images (EPI) were collected (TR = 2,020 ms; TE = 30 ms; flip angle = 90°; 37 slices with 3 mm

slice thickness; 64 × 64 acquisition matrix; field of view = 480 mm × 480 mm; voxel size = $3 \times 3 \times 3$ mm^3).

2.4 | Univariate analysis

For the neuroimaging data analyses, SPM8 (Welcome Department of Cognitive Neurology, London, UK; http://www.fil.ion.ucl.ac.uk/spm/) was used. EPI images were first realigned and resliced. Data sets that exhibited movement of >3 mms or 3° of rotation were not included. The anatomical image was then co-registered with the mean EPI image, and segmented, generating parameters for normalization to MNI space with a $2 \times 2 \times 2$ mm^3 resolution. Finally, the normalized functional images were smoothed by an 8-mm full-width, half-maximum Gaussian filter.

Each participant's brain activation was estimated using general linear models (GLMs 1 and 2). GLM 1 was performed to investigate: (a) the neural difference between lying and truth-telling and (b) the relation of individual differences in lying percentages and the neural activity during lying and truth-telling. Based on participants' decisions of lying and truth-telling in the color-reporting game, data from 14 participants were excluded from the analyses due to insufficient trials (<10) in either lying or truth-telling conditions. Therefore, GLM 1 was performed based on the remaining 23 participants. In GLM 2, we did not isolate lying and truth-telling conditions, instead we focused on the entire decision-making phase regardless of the decisions made by participants. GLM 2 was performed to investigate the relation of individual differences in lying percentages and the neural activity during the decision-making phase by applying MVPA. GLM 2 was performed based on 37 participants.

In GLM 1, four regressors of interest were included: (a) color cue phase in the lying trials, (b) color cue phase in the truth-telling trials, (c) lying decision phase, and (d) truth-telling decision phase. Trials with no response and six sets of motion parameters were also included in the GLM as regressors of no interest. The canonical hemodynamic

response function implemented in SPM8 was used to model the fMRI signal, and a high-pass filter was set at 128 s to reduce low frequency noise. For the second level analysis, we ran paired t tests on the contrast of lying versus truth-telling. To further investigate individual differences in lying percentages, we ran three group-level regression analyses: (a) the neural response of lying with individual lying percentages as the regressor, (b) the neural response of truth-telling with individual lying percentages as the regressor, and (c) the neural response of lying versus truth-telling with individual lying percentages as the regressor. In addition, potential gender differences were also checked by using (b) two sample t tests with gender as the grouping variable and (b) multiple regressions with gender and lying percentages as covariates. Two sample t tests were performed on the conditions of lying, truth-telling, and lying versus truth-telling to compare differences in neural responses between female and male participants. Multiple regressions were performed on three conditions in the regression analyses. Different from previous regression analyses, in addition to lying percentages, gender was entered as one of the covariates. If there is no additional statement, we reported results surviving after a voxel-level height threshold at $p < .001$ and cluster-level family wise error (FWE) correction, $p < .05$. A lenient voxel-level threshold ($p < .001$, uncorrected, $k = 50$) was used to check the underlying relationship between neural activity in the caudate and individual differences in lying decision-making.

In our experimental design, participants were free to make their own decisions to lie or not. Therefore, we have unbalanced trial numbers for lying and truth-telling conditions. This might influence our final results due to the potential different estimations of the conditions. To further support the validity of our results, we downsampled the trial numbers in the conditions with more trials by randomly selecting trials to balance the trial numbers in the condition with fewer trials. The detailed procedures are illustrated here. For each participant, we first identify the condition with less trial number. For example, one participant had 50 truth-telling trials but 70 lying trials. In this case, the truth-telling condition is the one with the lower trial number. We then randomly selected trials from the condition with the higher trial number to match with the trial number in the other. In the example case, 50 trials are randomly selected from the lying condition to enter into the following analyses. After that, we built the GLM again which is similar to GLM 1 as we mentioned above and performed the same analyses.

Finally, to investigate functional connectivity patterns during lying, we conducted generalized psychophysiological interaction (gPPI) analyses (McLaren, Ries, Xu, & Johnson, 2012). The seed regions were defined as a 3-mm sphere in the bilateral dlPFC, left caudate, and right IFG, using the peak voxels from significant clusters to the lying rate regression effects (MNI coordinates of peak voxels: left dlPFC: −36, 54, 14; right dlPFC: 38, 54, 6; left caudate: −12, 12, −2; right IFG: 44, 36, 0). The psychological variables were the contrasts of lying versus baseline. The blood oxygen level-dependent signal served as a physiological variable that was adjusted for confounds using an omnibus F-contrast. We further carried out correlation analyses between gPPI connectivity estimates and participants' lying rates. PPI images for all participants were then entered into a second-level model to determine the change of connectivity estimates with the increase in dishonest degree. Results were voxel-level height thresholded at $p < .001$ and survived after cluster-level FWE correction, $p < .05$.

In GLM 2, two regressors of interest were included. The onsets are as following: (a) color cue phase and (b) lying decision phase. Trials with no response and six sets of motion parameters were also included in the GLM as regressors of no interest. For the first level analysis, the contrasts of decision-making versus baseline (i.e., fixation) were estimated for each participant. The beta images were further used in the MVPA regression analysis (see below). For the second level analysis, individual lying percentage was entered into a group-level regression to investigate its impact on the neural activity
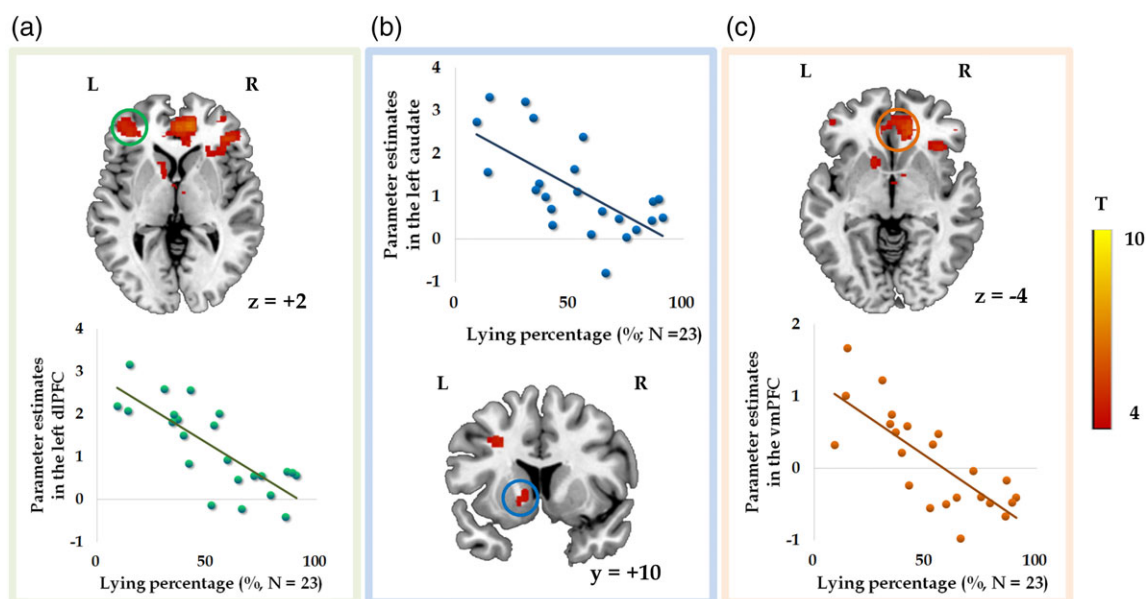
FIGURE 2 In the displayed brain regions, activity during lying negatively correlated with lying percentages (GLM 1; $N = 23$). Parameter estimates were extracted from the activated clusters of left dlPFC (a), left caudate (b; voxel-level $p < .001$ uncorrected, $k = 50$), and vmPFC (c; masked by an orbital medial prefrontal cortex mask from the automated anatomic labeling (AAL) atlas) [Color figure can be viewed at wileyonlinelibrary.com]

during the decision-making phase. In addition, potential gender differences were also investigated by using two sample *t* tests with gender as the grouping variable. Two sample *t* tests were performed on the lying decision phase to compare the differences between female and male participants in neural responses during decision-making. Results were voxel-level height thresholded at $p < .001$ and survived after cluster-level FWE correction, $p < .05$.

2.5 | Multivariate pattern analysis

For the second aim, we used the Pattern Recognition for Neuroimaging Toolbox (PRoNTo) (Schrouff et al., 2013) to analyze data from 37 participants (same as in GLM 2). We first created 90 anatomical regions of interests (ROIs) from the automated anatomic labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002) by using WFU Pickatlas Tool (Maldjian, Laurienti, Kraft, & Burdette, 2003), excluding the regions of cerebellum and vermis. The beta images of the decision-making phase from the first-level analysis in GLM 2 were entered into the analysis. A linear kernel between the voxel intensities within each ROI was created for each participant, thereby generating a 37 × 37 similarity matrix. We used kernel ridge regression (Shawe-Taylor & Cristianini, 2004) as the prediction machine. We adopted the leave-one-subject-out cross-validation with mean-centered features across training data, resulting in 37-fold cross-validation. In each fold, one input image was left out and treated as testing data. The regression machine was trained to associate the lying frequencies with the data in the remaining data from 36 participants and predict lying frequencies in the left-out participant. The predictive accuracy was calculated as the Pearson's correlation coefficient, coefficient of determination and normalized mean squared error between the predicted and actual treatment effectiveness. The significance of the prediction accuracy was evaluated using a permutation test with 10,000 iterations ($p < .01$; a lenient threshold of $p < .05$ was applied to check the prediction accuracy in the vmPFC).

3 | RESULTS

3.1 | Behavioral results

In 37 participants, lying percentages (i.e., dishonest degree) in the color-reporting game ranged from 0.83 to 100% (mean $\pm$ SD = 59.29

TABLE 1 Brain activation in the contrast of lying versus truth-telling (GLM 1; N = 23)

Conditions/regions	L/M/R	N	MNI coordinates			T
			x	y	z	
Lying versus truth-telling						
MFG	L	554	−24	−4	56	8.54
SPL	L	3,701	−12	−64	54	6.84
MFG	R	257	28	2	62	5.46
Truth-telling versus lying						
Lingual gyrus	L	531	−10	−90	−8	5.01
Cuneus	L	436	−14	−94	24	4.34

GLM = general linear model; FWE = family wise error; MFG = middle frontal gyrus; SPL = superior parietal lobule.
Voxel-level threshold $p < .001$ uncorrected, cluster-level $p < .05$ FWE correction.

$\pm$ 33.96; Figure 2b). Female and male participants did not show significant differences in lying percentages (mean and *SD* of lying percentages in female and male participants: 56.14% ($\pm$33.25) and 66.73% ($\pm$36.06); t(35) = 0. 86, p = .39). In 23 participants who had more than 10 lying and truth-telling trials (whose data were used in GLM 1), the lying percentages ranged from 9.24 to 91.60% (mean $\pm$ SD = 56.67 $\pm$ 24.08; Figure 2b). Significant negative correlations were found between the lying percentages and response time of lying (Spearman correlation: r = −0.45, p = .03; Figure 2c) but not between the lying percentages and response time of truth-telling (Spearman correlation: r = −0.02, p = .92; Figure 2c). Response time of lying was significantly longer than that of truth-telling (mean and *SD* of lying reaction time: 1,212 $\pm$ 165 ms; mean and *SD* of truth-telling reaction time: 1,067 $\pm$ 104 ms; $t_{[22]}$ = 4.53; p < .001). After the experiment, we asked participants how much they think misreporting in the color-reporting game was a "lie." The mean ratings ($\pm$SD) in 37 and 23 participants were 5.89 ($\pm$2.12) and 6.65 ($\pm$1.61), which were significantly higher than the neutral rating 5 ($t_{(36)}$ = 2.56, p = .015; $t_{(22)}$ = 4.91; p < .001) suggesting that participants tend to agree that misreporting the color in the experiment was a lie.

3.2 | Univariate results

To investigate the differences in neural process between lying and truth-telling, we compared these two decisions in GLM 1. The

TABLE 2 Negative correlations between neural activity and lying percentages (GLM 1; N = 23)

Conditions/regions	L/M/R	N	MNI coordinates			T
			x	y	z	
Lying[a]						
ACC	L	1,931	−12	42	12	10.19[b]
dlPFC	L	417	−36	54	14	6.32[b]
Superior frontal gyrus	R	506	2	16	60	6.14[b]
IFG	R	431	44	36	0	5.52[b]
MFG	L	96	−28	12	38	4.93
Precuneus	L	58	−12	−48	16	4.71
dlPFC	R	198	38	54	6	4.59[b]
Thalamus	L	53	−2	−4	−2	4.57
IPL	R	60	50	−48	50	4.48
Caudate head	L	86	−12	12	−2	4.42
IFG	L	62	−50	20	6	4.34
Cingulate cortex	R	67	8	−34	40	4.34
Lying versus truth-telling[b]						
Insula	R	2,085	34	28	−6	9.58
ACC	R	7,445	6	28	32	9.38
Supramarginal gyrus	L	380	−64	−46	28	6.38
Thalamus	R	283	8	−12	−2	6.35
Culmen	R	227	2	−32	−18	5.72
MFG	R	209	44	2	54	5.19

ACC = anterior cingulate cortex; GLM = general linear model; FWE = family wise error; IPL = inferior parietal lobule; MFG = middle frontal gyrus.
[a] Voxel-level threshold $p < .001$ uncorrected, $k = 50$.
[b] Voxel-level threshold $p < .001$ uncorrected, cluster-level $p < .05$ FWE correction.
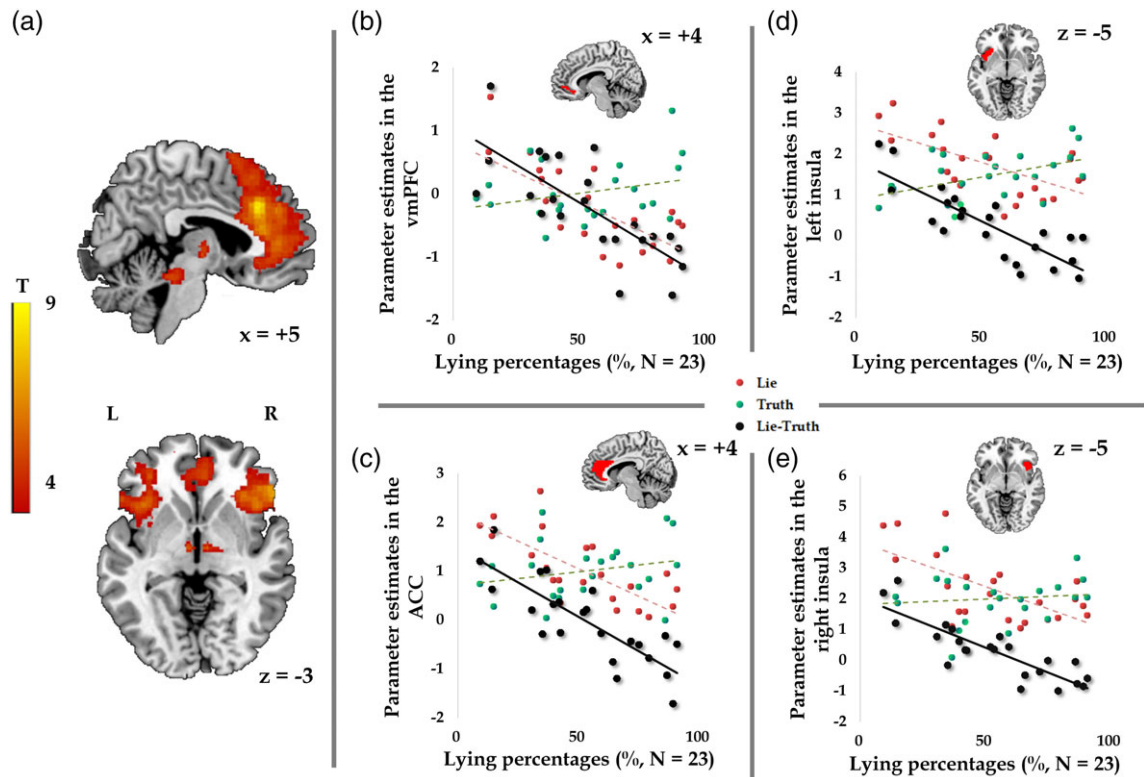
FIGURE 3 (a) In the displayed brain regions, activity in the contrast of lying versus truth-telling negatively correlated with lying percentages (GLM 1; N = 23; voxel-level p < .001 uncorrected, cluster-level p < .05 FWE correction). Negative correlations were found between the lying percentages and activity in the vmPFC (b; masked by orbital medial prefrontal cortex mask from the AAL atlas), ACC (c; masked by anterior cingulate cortex from the AAL atlas), and bilateral insula (d and e; masked by bilateral insula from the AAL atlas) in the contrast of lying versus truth-telling (black dots and lines). Red and green dots and lines represent the relations between lying percentages and the activity elicited by lies and truth, respectively [Color figure can be viewed at wileyonlinelibrary.com]

contrast of lying versus truth-telling significantly activated the bilateral MFG and left superior parietal lobule (SPL) (Table 1). The opposite contrast activated the left lingual gyrus and left cuneus. To further investigate whether the neural activation reflects individual differences in lying percentages (i.e., dishonest degree), we performed regression analyses at the group level. The results showed that lying percentages were negatively correlated with the activity in the left dlPFC (Figure 2a), left caudate (Figure 2b; voxel-level p < .001, uncorrected, 86 voxels), and vmPFC (Figure 2(c)) while participants were lying (Table 2). Negative correlations were also found between the lying percentages and activity in brain regions elicited by the contrast of lying versus truth-telling (Table 2; Figure 3a), including the vmPFC (Figure 3b), ACC (Figure 3c), and bilateral insula (Figure 3d,e). The bilateral caudate was activated by applying a lenient threshold (voxel-level p < .001, uncorrected; left caudate: 11 voxels, peak MNI coordinate is −8, 4, −2; right caudate: 34 voxels, peak MNI coordinate is 10, 8, 2). In GLM 1, no significant positive correlations were found between brain response to lying and lying percentages. Brain activity during truth-telling and lying percentages did not show any significant correlations. Besides, no significant correlations were found between lying percentages and neural activity while making decisions in GLM 2.

In order to investigate whether gender differences play a potential role in the relation between neural response and individual differences in the aforementioned findings, two sample t tests and multiple

TABLE 3 Results of gPPI functional connectivity analysis during lying (N = 23)

Regions		L/M/R	N	MNI coordinates x	y	z	T
Left dlPFC[a]	IFG	R	852	58	10	22	6.40
	Cingulate gyrus	R	510	18	−44	28	6.30
	Putamen	L	858	−26	−6	−6	5.73
	Precentral gyrus	L	1,207	−46	−4	22	5.63
	IPL	L	747	−34	−40	28	5.55
	Insula	R	850	32	−14	12	5.17
	SPL	L	404	−10	−72	56	5.00
Right dlPFC[a]	Thalamus	R	335	16	−14	10	5.41
	Middle frontal gyrus	L	230	−34	26	18	5.22
Left caudate[a]	Insula	R	2,462	42	−10	−2	6.98
	Superior temporal gyrus	L	1,879	−60	−48	20	6.47
	Superior temporal gyrus	R	544	36	−46	20	5.83
	Precentral gyrus	R	342	42	−24	38	5.58
Right IFG[a]	vmPFC	R	719	6	34	−16	5.52

dlPFC = dorsolateral prefrontal cortex; FWE = family wise error; gPPI = generalized psychophysiological interaction; IFG = inferior frontal gyrus; IPL = inferior parietal lobule; SPL = superior parietal lobule; vmPFC = ventromedial prefrontal cortex.
Voxel-level threshold p < .001 uncorrected, cluster-level p < .05 FWE correction.
[a] Seed regions for gPPI analysis.
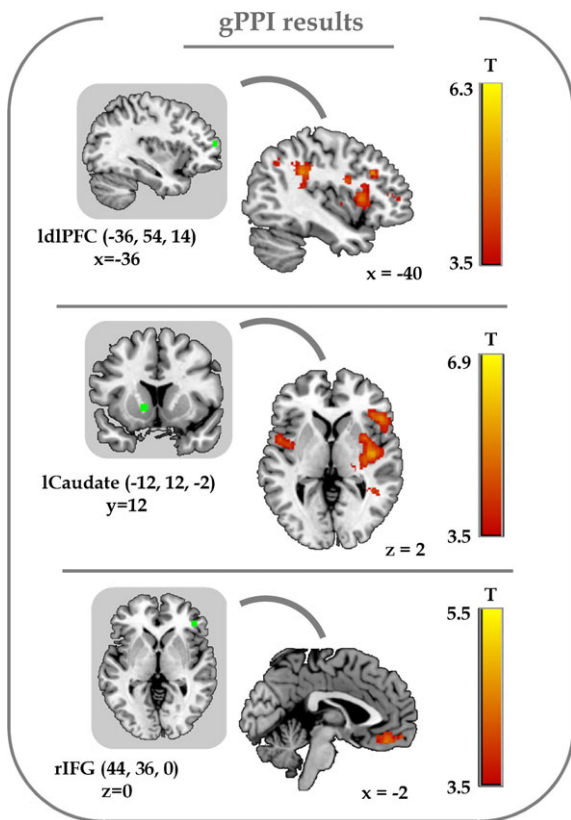
FIGURE 4 The results of generalized psychophysiological interaction (gPPI) analyses. The presented seed regions were the left dlPFC, left caudate and right IFG (MNI coordinates of peak voxels: Left dlPFC: −36, 54, 14; left caudate: −12, 12, −2; right IFG: 44, 36, 0; radius of sphere: 3 mm; in green). The positive functional connectivity between the left dlPFC and insula, the left caudate and insula, as well as the right IFG and vmPFC during lying significantly correlated with the individual dishonest degree (voxel-level $p < .001$ uncorrected, cluster-level $p < .05$ FWE correction) [Color figure can be viewed at wileyonlinelibrary.com]

regression analyses with gender as an additional covariate were performed. The results of two sample t tests showed that only in the condition of lying, females elicited higher activation in the right fusiform gyrus (Figure S1, Supporting Information). In the condition of truth-telling and the contrast of lying and truth-telling, no significant differences were found between female and male participants. With regard

TABLE 4 Results of the pattern regression analysis ($N = 37$)

ROIs	r	$p_{(r)}$	R^2	$p_{(R^2)}$	nMSE	$p_{(nMSE)}$
Left caudate head	0.65	.0012	0.43	.0013	9.71	.0049
Left superior occipital lobule	0.62	.0017	0.39	.0024	7.72	.0016
Left olfactory cortex	0.55	.001	0.3	.0165	8.86	.0026
Right olfactory cortex	0.49	.0068	0.24	.0269	9.74	.0081
Left vmPFC[a]	0.38	.047	0.14	.08	14.58	.059
Right vmPFC[a]	0.41	.03	0.17	.0581	14.59	.0643

nMSE = normalized mean squared error; R^2 = coefficient of determination; r = Pearson's correlation coefficient; ROIs = regions of interests; vmPFC = ventromedial prefrontal cortex.
[a] A lenient threshold $p < .05$ was applied to check the prediction accuracy in the vmPFC.

to the multiple regression results, after controlling for the gender effect, we still found significant negative correlations between lying percentages and neural activity in the left dlPFC, left caudate (voxel-level $p < .001$, uncorrected, 77 voxels), and vmPFC during lying (Figure S2, Supporting Information). Negative correlations were also found between the lying percentages and neural activity in brain regions including the vmPFC, ACC, and bilateral insula when comparing between lying and truth-telling (Figure S3, Supporting Information). To eliminate potential influences of unbalanced trial numbers between lying and truth-telling conditions, we balanced the trial numbers (see Section 2.4 for more details). We still observed the effect in the aforementioned regions (voxel-level $p < .001$, uncorrected).

The individual differences in dishonesty were also found in the functional connectivity among the dlPFC, vmPFC, caudate, and insula. Seed regions for the gPPI analyses were constructed for four ROIs, the activity of which showed significant negative correlations with individuals' dishonest degree during decision-making phases (Figure 4; Table 3). With increase in individuals' degree of dishonesty, the left dlPFC showed more positive functional connectivity with the IFG, cingulate gyrus, putamen, precentral gyrus, IPL, insula, and SPL during lying, while the left caudate showed more positive functional connectivity with the insula, superior temporal gyrus and precentral gyrus, and the right IFG showed higher positive functional connectivity with the vmPFC. To further explore our connectivity results, we applied a lenient threshold ($p < .005$, uncorrected) to see if the connectivity maps of the left dlPFC and caudate overlapped with the seed region of right IFG. We obverse the right IFG overlapped with the connectivity maps of the left dlPFC and caudate (Figure S4, Supporting Information).

3.3 | MVPA results

To explore if there are regions sensitive to predict individuals' degree of dishonesty, we used a pattern regression analysis to examine the accuracy of 90 ROIs in predicting lying percentages. Results showed that the correlation coefficients between actual and predicted lying percentages were significant in the left caudate, left superior occipital lobe and bilateral olfactory cortex (Table 4; Figure 5). By applying a lenient threshold ($p < .05$), patterns in bilateral vmPFC can be used to predict lying percentages as well.

4 | DISCUSSION

In our study, we investigated brain regions in which lying associated activity and functional connectivity could be related to the probability of lying, and brain patterns which can be used to predict individual differences in dishonesty. Our results identify two key aspects of individual differences of lying. First, the lying associated activation in the left caudate, vmPFC, right IFG, and left dlPFC, as well as the functional connectivity among them are sensitive to individual differences in dishonesty. The dishonesty associated activity in these regions decreased with increased dishonesty. Second, the neural patterns in the left caudate and vmPFC can be used to predict individual differences in dishonesty.
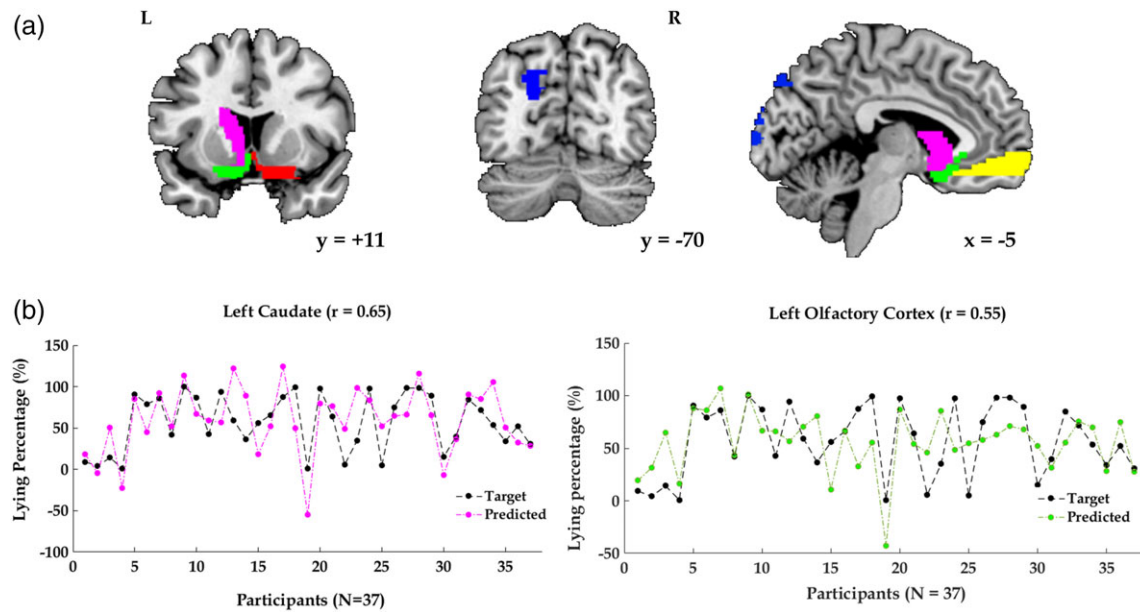
FIGURE 5 (a) In the pattern regression analysis, the correlation coefficients between actual and predicted lying frequencies were significant in the left caudate head (pink), left superior occipital lobe (blue), bilateral olfactory cortex (green and red; $p < .01$), and bilateral vmPFC (yellow; a lenient threshold $p < .05$). (b) Illustration of actual lying percentages (black dots) and the predicted lying percentages in the left caudate head (pink dots) and left olfactory cortex (green dots) [Color figure can be viewed at wileyonlinelibrary.com]

In our results, without considering individual differences in deceptive decision-making, lying elicited higher activation in the MFG and SPL than truth-telling, while truth-telling elicited higher activation in the left lingual gyrus and cuneus. However, if we took individual differences in dishonesty into consideration, we can observe negative correlations between variations in lying percentages and lying associated neural activity in the left caudate, vmPFC, right IFG, and left dlPFC. From here, we can see the dramatic difference between the results of a direct comparison of lying and truth-telling and the results of linking behavioral variations with neural activity. If we only focus on the direct comparison, we can only conclude that in our current experimental setting, lying decisions elicited higher activation in the MFG and SPL. Actually, these two regions were also found in previous studies, which were activated when comparing lying versus truth-telling (Langleben et al., 2005; Lee et al., 2005; Vartanian et al., 2013). These regions might show general effects caused by the process of lying, but we also found some other regions which are sensitive to individual differences in lying. As mentioned in Section 1, the dlPFC and IFG were commonly found activated in studies about lying (Ding et al., 2012; Farah et al., 2014; Pornpattananangkul et al., 2018; Yin et al., 2016). However, the findings from direct comparisons of lying versus truth-telling in these two regions, especially in the dlPFC are not always consistent. Higher activation in the dlPFC was found both when participants were lying and telling the truth (Abe & Greene, 2014; Baumgartner et al., 2009; Greene & Paxton, 2009; Sip et al., 2010; Yin et al., 2016; Zhu et al., 2014). Hence, individual differences might influence activation patterns in the dlPFC and IFG which failed to be detected in the direct comparison.

The reduced activity in the dlPFC and IFG in more dishonest individuals might be associated with the less cognitive resources recruiting to repress self-benefiting motives, and therefore facilitate the decision-making process of lying in these participants. First, behavioral

results showed that with increased individuals' degree of dishonesty, lying reaction time decreased, while truth-telling reaction time remained unchanged. One might expect that the response time of making honest decisions should decrease in those more honest individuals as well, since they made more honest responses and repetition might shorten the response time. However, we found that it is only true for making dishonest decisions. Note that the color-report game is relatively easy and the rules of the game are straightforward, compared to some other economic games (e.g., sender–receiver game) or mock-crime games. The unchanged truth-telling reaction time might be due to the general time needed for recalling the previously shown color and preparing for making the correct response in our experiment. Repetition might not influence the time needed for recalling and responding, regardless of whether the decision makers are more honest or dishonest. However, it is not the case for lying decision-making. Those who made more dishonest responses did lie faster than those who were more honest. A previous study found a gradual escalation of self-serving dishonesty (Garrett, Lazzaro, Ariely, & Sharot, 2016). They also found that the reduction of amygdala activity related to an escalation of self-serving dishonesty. However, we failed to find significant results in the amygdala. We speculated that the differences might be caused by the punishment rules we applied in our experiment. In Garrett et al.'s research, they mentioned that "in our design, no feedback, such as external punishment for dishonesty or praise for honesty, was provided or expected." Punishment is a very important factor which was not captured in their study and should be considered as a limitation. The amygdala is sensitive in detecting a threat (Bishop, Duncan, & Lawrence, 2004; Isenberg et al., 1999; Phelps & LeDoux, 2005) and punishment (Murty, LaBar, & Adcock, 2012). Punishment for lying behaviors in our study might weaken adaptation to dishonesty and therefore we did not observe significant results in the amygdala. Instead of the amygdala, we found that the dlPFC and IFG

contribute to the observed individual differences. Rather than adaptation, the reduced response time together with the reduced activation in the dlPFC and IFG during lying in more dishonest participants might suggest less engagement of executive control. The dlPFC has been found in tasks that require inhibition and cognitive control (Aron, Robbins, & Poldrack, 2004; Spitzer, Fischbacher, Herrnberger, Grön, & Fehr, 2007). Among many brain regions associated with lying, the dlPFC is one of the regions which could well represent individual differences in allocating cognitive resources while making lying decisions. Studies of transcranial magnetic stimulation (TMS) and lesion patient have shown that neural activity or lesions in the dlPFC significantly influences lying behaviors. Inhibitory repetitive TMS (rTMS) effects on the left dlPFC increased lying percentages (Karton & Bachmann, 2011) and enhanced neural excitability in the right dlPFC dramatically reduced cheating especially when cheating benefited the participants themselves (Maréchal, Cohn, Ugazio, & Ruff, 2017). Damage to the dlPFC decreased individuals' honesty concerns and increased their dishonest behavior (Zhu et al., 2014). Consistent with these findings, we found reduced activity in the dlPFC in the participants who lied more often. These results suggest that a low level of cognitive control was engaged when dishonest individuals were lying. With the increase in individuals' dishonest degree, the dlPFC showed higher connectivity with the IPL which also belongs to the frontoparietal network (Yin et al., 2016), conflict resolving associated cingulate gyrus (Shackman et al., 2011; Shenhav, Botvinick, & Cohen, 2013), and motion preparation associated putamen (Balleine & O'doherty, 2010). These regions are critical for resolving conflicts and making appropriate responses. We speculated that these regions might receive signals from the dlPFC and moderate the response time.

The individual differences in dishonesty are not only reflected in lying associated activity and functional connectivity in cognitive control associated regions, but also in the caudate and vmPFC, which might signal the subjective value of dishonest gains. Substantial evidence indicates that the striatum is involved in reward-related tasks. The striatum is thought to be involved in the representation of values of actions (Lau & Glimcher, 2008) and anticipation of gains (Wu, Samanez-Larkin, Katovich, & Knutson, 2014). The ventral striatum together with the vmPFC might represent reward in general and reward magnitude in particular (Diekhof et al., 2012). Note that in our experiment, we applied the punishment rule. The answers in 20% of trials would be checked by the computer. Participants would get 1 point for each correct answer, 8 points for each incorrect answer without being checked, and −6 points for each incorrect answer being checked. Hence, the expected value of lying in each trial is 5.2 (i.e., $8 \times 0.8 + [−6] \times 0.2 = 5.2$), which is larger than the expected value of truth-telling in each trial (i.e., 1). Besides, the total monetary payoffs enhanced with increased lying behaviors. That is, even in 20% of trials where the all-time liars (who lied all the time during the whole experiment) lost additional money, they still earned the most. According to previous findings about the striatum and vmPFC, two findings might be expected. The first one would be that higher activity in these regions should be observed in the comparison between lying and truth-telling, since expected values of lying are larger than that of truth-telling. The second one would be that the activity in the striatum and vmPFC might also increase in participants who lied more often,

since the total payoffs in these participants are larger than those who lied less often.

However, our results showed that first, no significant differences were observed in the caudate and vmPFC in the contrasts between lying versus truth-telling. This implies that at the group level, higher monetary reward gained by lying did not lead to higher activity in the caudate and vmPFC. The individual differences across participants might cancel out the average differences in neural activity between lying and truth-telling. In addition, the subjective value of dishonest gains might be discounted by the means (i.e., lying). In a study about neural responses toward beneficial and harmful lies and truth, only beneficial truth elicited positive activity in the NAcc while beneficial lies and harmful truth exhibited different patterns (low level of activity) in the NAcc (Yin & Weber, 2016). The vmPFC and the caudate code not only the monetary values but also subjective values of decisions (Pornpattananangkul et al., 2018). Therefore, the psychological cost of lying might be weighed against the additional monetary benefits brought by lying.

Second, for those individuals with higher degrees of dishonesty, the lying associated activity in the caudate and vmPFC decreased, whereas the activity in these regions when telling the truth showed no significant correlation with the individual propensity to lie. This finding might also be due to the reduced subject value of dishonest gains and enhanced psychological costs of lying in those participants who behaved more dishonestly. The vmPFC, ventral striatum, and insula also signal aversive value expectations (Basten, Biele, Heekeren, & Fiebach, 2010; Plassmann, O'Doherty, & Rangel, 2010; Shankman et al., 2014; Tom, Fox, Trepel, & Poldrack, 2007). When the gains of a certain decision are physically harmful to others, lower dorsal striatal responses were observed in participants with stronger moral preferences (Crockett, Siegel, Kurth-Nelson, Dayan, & Dolan, 2017). What is more, individuals might have both rewarding feelings from dishonest monetary gains and psychological costs of lying (Yin et al., 2017). Therefore, the subjective value of dishonest gains might be discounted as more dishonest decisions are made. In addition, participants knew that their responses would be recorded and lies would be punished if caught. They might also add additional costs of making dishonest decisions, resulting in reduced subjective value. The caudate and vmPFC might function to code subjective values of dishonest outcomes/consequences and could, therefore, be used to predict individual differences in lying. Indeed, we found that activity in the caudate, vmPFC, and left superior occipital lobule during the decision-making phase predicts individual differences in dishonesty. Out of these, the caudate in particular stood out. It provided the predicted percentage of lying with the highest correlation coefficient with the actual percentage of lying. Besides, dishonest degrees positively correlate with positive functional connectivity between the caudate and IFG/insula, and between the IFG and vmPFC. The vmPFC is important in integrating different sources of value information and valuing choices (Rangel & Hare, 2010; Rushworth, Noonan, Boorman, Walton, & Behrens, 2011) and is a hub in the effective connectivity network identified for the representation of expected value (Minati, Grisoli, Seth, & Critchley, 2012). The dlPFC showed increased functional connectivity with the right IFG, which overlapped with the right IFG which showed increased connectivity with the vmPFC. The seed region of the IFG also overlapped with the right IFG with which the caudate showed increased

functional connectivity. The connection between cognitive control regions like the dlPFC and value coding regions like the vmPFC and caudate might bridge through a third region like the IFG (Hare, Camerer, & Rangel, 2009), a region that is associated with inhibition control (Aron et al., 2004; Aron, Fletcher, Bullmore, Sahakian, & Robbins, 2003; Aron, Robbins, & Poldrack, 2014; Chikazoe, Konishi, Asari, Jimura, & Miyashita, 2007). The reduced activity in the dlPFC, IFG, vmPFC and caudate and increased connectivity among them with the IFG as a hub might complete a full picture of options evaluation, assignment of cognitive resources and facilitation of lying responses in more dishonest individuals.

Another important aspect to discuss is the gender differences. The gender difference was found in lying decision-making in both the behavioral (Dreber & Johannesson, 2008; Erat & Gneezy, 2012; Lundquist et al., 2009) and neural domains (Marchewka et al., 2012). Studies showed that men might be more likely to tell self-serving lies than women (Erat & Gneezy, 2012). For smaller stakes, women might show greater aversion to lying, but for larger stakes, no differences in lying by gender were found (Childs, 2012). With regards to the neural domain, in Marchewka et al.'s (2012) research, they found men showed higher activation in the left MFG compared to women only during lying for personal information. However, we did not observe gender differences either in their lying percentages or in the brain regions that were found to be significantly associated with individual differences. We think that the discrepancy between previous studies and our study might be due to the difference in experimental design. In the behavioral studies, a sender–receiver game was used, in which a sender could choose to send a false message to mislead a receiver, resulting in different payoff outcomes for both sides. In this paradigm, liars were not punished for their lying behaviors. But in our design, the application of the punishment rule might influence individuals' decision-making process and cause different results. In Marchewka et al.'s (2012) research, an instructed lying paradigm was used and participants answered questions according to cues provided by the computer. In our experiment, participants could freely choose to lie or not and no instructions about how to make decisions were provided. What is more, they only found gender differences for personal information but not for general information. The question we used is not about personal information. Therefore, this might also be the reason why we did not observe gender differences.

4.1 | Limitation

Although we did not find significant gender differences in lying frequencies and regions reflecting individual differences in lying, we cannot claim that there are no gender differences in making deceptive decisions or that gender does not play a role in individual differences in lying. In our study, we recruited more female participants than male participants. In this sense, the conclusions from the present study might be limited to a certain context and might show differences with that observed in the entire population.

Even though we used a task where individuals were free to make dishonest decisions with additional gains while running the risks of getting caught, another important factor (i.e., interaction between the decision maker and the recipient) was lacking. The interaction might influence the participants' decisions and the underlying neural processes. For example, in the context where participants could interact with others (Sun, Lee, Wang, & Chan, 2016; Sun, Shao, Wang, & Lee, 2018), fewer dishonest choices were made when participants interacted with human than computer investors and distinct neural processing during lying were involved when the decision maker considers the interests of the recipient (Sun et al., 2016). Therefore, the lack of interaction in our current design should be considered as a limitation and the effect of interaction on deception decision-making should be investigated in future studies.

5 | CONCLUSIONS

Taken together, our findings have important implications for understanding the brain regions and their functional connectivity associated with individual differences in dishonesty by using both univariate and MVPA methods. The dramatic differences could be found between the results of the direct comparison of lying and truth-telling and the results of linking behavioral variations with the neural activity. Besides, the present study raises interesting hypotheses about functions of the caudate, vmPFC, IFG, and dlPFC as a network. On the one hand, the reduced activity and the increased functional connectivity in cognitive control associated regions might increase the propensity of lying and reducing lying reaction time. On the other hand, reward-related regions signal the discounted value of dishonest gains. The activity in the dlPFC, IFG, vmPFC, and caudate and the connectivity among them with the IFG as a potential hub might work together to evaluate options, assign cognitive resources and facilitate (dis)honest responses.

CONFLICT OF INTERESTS

None.

ORCID

Lijun Yin https://orcid.org/0000-0003-1748-3187
Bernd Weber https://orcid.org/0000-0002-7811-9605

REFERENCES

Abe, N., & Greene, J. D. (2014). Response to anticipated reward in the nucleus accumbens predicts behavior in an independent test of honesty. *Journal of Neuroscience*, *34*(32), 10564–10572.

Abe, N., Okuda, J., Suzuki, M., Sasaki, H., Matsuda, T., Mori, E., … Fujii, T. (2008). Neural correlates of true memory, false memory, and deception. *Cerebral Cortex*, *18*(12), 2811–2819.

Aron, A. R., Fletcher, P. C., Bullmore, E. T., Sahakian, B. J., & Robbins, T. W. (2003). Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. *Nature Neuroscience*, *6*(2), 115–116.

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences*, *8*(4), 170–177.

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2014). Inhibition and the right inferior frontal cortex: One decade on. *Trends in Cognitive Sciences*, 18(4), 177–185.

Balleine, B. W., & O'doherty, J. P. (2010). Human and rodent homologies in action control: Corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology*, 35(1), 48–69.

Basten, U., Biele, G., Heekeren, H. R., & Fiebach, C. J. (2010). How the brain integrates costs and benefits during decision making. *Proceedings of the National Academy of Sciences of the United States of America*, 107(50), 21767–21772.

Baumgartner, T., Fischbacher, U., Feierabend, A., Lutz, K., & Fehr, E. (2009). The neural circuitry of a broken promise. *Neuron*, 64(5), 756–770.

Bishop, S. J., Duncan, J., & Lawrence, A. D. (2004). State anxiety modulation of the amygdala response to unattended threat-related stimuli. *Journal of Neuroscience*, 24(46), 10364–10368.

Chavez, R. S., & Heatherton, T. F. (2015). Representational similarity of social and valence information in the medial pFC. *Journal of Cognitive Neuroscience*, 27(1), 73–82.

Chikazoe, J., Konishi, S., Asari, T., Jimura, K., & Miyashita, Y. (2007). Activation of right inferior frontal gyrus during response inhibition across response modalities. *Journal of Cognitive Neuroscience*, 19(1), 69–80.

Childs, J. (2012). Gender differences in lying. *Economics Letters*, 114(2), 147–149.

Christ, S., Van Essen, D., Watson, J., Brubaker, L., & McDermott, K. (2009). The contributions of prefrontal cortex and executive control to deception: Evidence from activation likelihood estimate meta-analyses. *Cerebral Cortex*, 19(7), 1557–1566.

Crockett, M. J., Siegel, J. Z., Kurth-Nelson, Z., Dayan, P., & Dolan, R. J. (2017). Moral transgressions corrupt neural representations of value. *Nature Neuroscience*, 20(6), 879–885.

Cui, F., Wu, S., Wu, H., Wang, C., Jiao, C., & Luo, Y. (2017). Altruistic and self-serving goals modulate behavioral and neural responses in deception. *Social Cognitive and Affective Neuroscience*, 13(1), 63–71.

Cui, Q., Vanman, E. J., Wei, D., Yang, W., Jia, L., & Zhang, Q. (2014). Detection of deception based on fMRI activation patterns underlying the production of a deceptive response and receiving feedback about the success of the deception after a mock murder crime. *Social Cognitive and Affective Neuroscience*, 9(10), 1472–1480.

Diekhof, E. K., Kaps, L., Falkai, P., & Gruber, O. (2012). The role of the human ventral striatum and the medial orbitofrontal cortex in the representation of reward magnitude – An activation likelihood estimation meta-analysis of neuroimaging studies of passive reward expectancy and outcome processing. *Neuropsychologia*, 50(7), 1252–1266.

Ding, X. P., Du, X., Lei, D., Hu, C. S., Fu, G., & Chen, G. (2012). The neural correlates of identity faking and concealment: An FMRI study. *PLoS One*, 7(11), e48639.

Dogan, A., Morishima, Y., Heise, F., Tanner, C., Gibson, R., Wagner, A. F., & Tobler, P. N. (2016). Prefrontal connections express individual differences in intrinsic resistance to trading off honesty values against economic benefits. *Scientific Reports*, 6(33263), 1–12.

Dreber, A., & Johannesson, M. (2008). Gender differences in deception. *Economics Letters*, 99(1), 197–199.

Ennis, E., Vrij, A., & Chance, C. (2008). Individual differences and lying in everyday life. *Journal of Social and Personal Relationships*, 25(1), 105–118.

Erat, S., & Gneezy, U. (2012). White lies. *Management Science*, 58(4), 723–733.

Farah, M. J., Hutchinson, J. B., Phelps, E. A., & Wagner, A. D. (2014). Functional MRI-based lie detection: Scientific and societal challenges. *Nature Reviews. Neuroscience*, 15(2), 123–131.

Galvan, A., Hare, T. A., Davidson, M., Spicer, J., Glover, G., & Casey, B. (2005). The role of ventral frontostriatal circuitry in reward-based learning in humans. *The Journal of Neuroscience*, 25(38), 8650–8656.

Garrett, N., Lazzaro, S. C., Ariely, D., & Sharot, T. (2016). The brain adapts to dishonesty. *Nature Neuroscience*, 19(12), 1727–1732.

Gozna, L. F., Vrij, A., & Bull, R. (2001). The impact of individual differences on perceptions of lying in everyday life and in a high stake situation. *Personality and Individual Differences*, 31(7), 1203–1216.

Greene, J., & Paxton, J. (2009). Patterns of neural activity associated with honest and dishonest moral decisions. *Proceedings of the National Academy of Sciences of the United States of America*, 106(30), 12506–12511.

Grolleau, G., Kocher, M. G., & Sutan, A. (2016). Cheating and loss aversion: Do people cheat more to avoid a loss? *Management Science*, 62(12), 3428–3438.

Hampton, A. N., & O'Doherty, J. P. (2007). Decoding the neural substrates of reward-related decision making with functional MRI. *Proceedings of the National Academy of Sciences of the United States of America*, 104(4), 1377–1382.

Hare, T. A., Camerer, C. F., & Rangel, A. (2009). Self-control in decision-making involves modulation of the vmPFC valuation system. *Science*, 324(5927), 646–648.

Haxby, J. V. (2012). Multivariate pattern analysis of fMRI: The early beginnings. *NeuroImage*, 62(2), 852–855.

Isenberg, N., Silbersweig, D., Engelien, A., Emmerich, S., Malavade, K., Beattie, B.a., ... Stern, E. (1999). Linguistic threat activates the human amygdala. *Proceedings of the National Academy of Sciences of the United States of America*, 96(18), 10456–10459.

Karton, I., & Bachmann, T. (2011). Effect of prefrontal transcranial magnetic stimulation on spontaneous truth-telling. *Behavioural Brain Research*, 225(1), 209–214.

Kim, H., Shimojo, S., & O'Doherty, J. P. (2010). Overlapping responses for the expectation of juice and money rewards in human ventromedial prefrontal cortex. *Cerebral Cortex*, 21(4), 769–776.

Langleben, D., Loughead, J., Bilker, W., Ruparel, K., Childress, A., Busch, S., & Gur, R. (2005). Telling truth from lie in individual subjects with fast event-related fMRI. *Human Brain Mapping*, 26(4), 262–272.

Lau, B., & Glimcher, P. W. (2008). Value representations in the primate striatum during matching behavior. *Neuron*, 58(3), 451–463.

Lee, T., Liu, H., Chan, C., Ng, Y., Fox, P., & Gao, J. (2005). Neural correlates of feigned memory impairment. *NeuroImage*, 28(2), 305–313.

Luan Phan, K., Magalhaes, A., Ziemlewicz, T., Fitzgerald, D., Green, C., & Smith, W. (2005). Neural correlates of telling lies: A functional magnetic resonance imaging study at 4 Tesla. *Academic Radiology*, 12(2), 164–172.

Lundquist, T., Ellingsen, T., Gribbe, E., & Johannesson, M. (2009). The aversion to lying. *Journal of Economic Behavior and Organization*, 70(1), 81–92.

Maldjian, J. A., Laurienti, P. J., Kraft, R. A., & Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage*, 19(3), 1233–1239.

Marchewka, A., Jednorog, K., Falkiewicz, M., Szeszkowski, W., Grabowska, A., & Szatkowska, I. (2012). Sex, lies and fMRI—Gender differences in neural basis of deception. *PLoS One*, 7(8), e43076.

Maréchal, M. A., Cohn, A., Ugazio, G., & Ruff, C. C. (2017). Increasing honesty in humans with noninvasive brain stimulation. *Proceedings of the National Academy of Sciences of the United States of America*, 114(17), 4360–4364.

McLaren, D. G., Ries, M. L., Xu, G., & Johnson, S. C. (2012). A generalized form of context-dependent psychophysiological interactions (gPPI): A comparison to standard approaches. *NeuroImage*, 61(4), 1277–1286.

Minati, L., Grisoli, M., Seth, A. K., & Critchley, H. D. (2012). Decision-making under risk: A graph-based network analysis using functional MRI. *NeuroImage*, 60(4), 2191–2205.

Murty, V. P., LaBar, K. S., & Adcock, R. A. (2012). Threat of punishment motivates memory encoding via amygdala, not midbrain, interactions with the medial temporal lobe. *Journal of Neuroscience*, 32(26), 8969–8976.

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, 10(9), 424–430.

Nunez, J., Casey, B., Egner, T., Hare, T., & Hirsch, J. (2005). Intentional false responding shares neural substrates with response conflict and cognitive control. *NeuroImage*, 25(1), 267–277.

O'Doherty, J. (2004). Reward representations and reward-related learning in the human brain: Insights from neuroimaging. *Current Opinion in Neurobiology*, 14(6), 769–776.

Ofen, N., Whitfield-gabrieli, S., Chai, X. J., Schwarzlose, R. F., & Gabrieli, J. D. (2017). Neural correlates of deception: Lying about past

events and personal beliefs. *Social Cognitive and Affective Neuroscience*, *12*(1), 116–127.

Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: From animal models to human behavior. *Neuron*, *48*(2), 175–187.

Plassmann, H., O'Doherty, J. P., & Rangel, A. (2010). Appetitive and aversive goal values are encoded in the medial orbitofrontal cortex at the time of decision making. *NeuroImage*, *30*(32), 10799–10808.

Pornpattananangkul, N., Zhen, S., & Yu, R. (2018). Common and distinct neural correlates of self-serving and prosocial dishonesty. *Human Brain Mapping*, *39*, 3086–3103.

Rangel, A., & Hare, T. (2010). Neural computations associated with goal-directed choice. *Current Opinion in Neurobiology*, *20*(2), 262–270.

Rosenbaum, S. M., Billinger, S., & Stieglitz, N. (2014). Let's be honest: A review of experimental evidence of honesty and truth-telling. *Journal of Economic Psychology*, *45*, 181–196.

Rushworth, M. F. S., Noonan, M. A. P., Boorman, E. D., Walton, M. E., & Behrens, T. E. (2011). Frontal cortex and reward-guided learning and decision-making. *Neuron*, *70*(6), 1054–1069.

Schrouff, J., Rosa, M. J., Rondina, J. M., Marquand, A. F., Chu, C., Ashburner, J., ... Mourão-Miranda, J. (2013). PRoNTo: Pattern recognition for neuroimaging toolbox. *Neuroinformatics*, *11*(3), 319–337.

Shackman, A. J., Salomons, T. V., Slagter, H. A., Fox, A. S., Winter, J. J., & Davidson, R. J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nature Reviews. Neuroscience*, *12*(3), 154–167.

Shankman, S. A., Gorka, S. M., Nelson, B. D., Fitzgerald, D. A., Phan, K. L., & O'Daly, O. (2014). Anterior insula responds to temporally unpredictable aversiveness: An fMRI study. *Neuroreport*, *25*(8), 596–600.

Shao, R., & Lee, T. (2017). Are individuals with higher psychopathic traits better learners at lying? Behavioural and neural evidence. *Translational Psychiatry*, *7*, e1175.

Shawe-Taylor, J., & Cristianini, N. (2004). *Kernel methods for pattern analysis*. Cambridge: Cambridge University Press.

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron*, *79*(2), 217–240.

Sip, K., Lynge, M., Wallentin, M., McGregor, W., Frith, C., & Roepstorff, A. (2010). The production and detection of deception in an interactive game. *Neuropsychologia*, *48*(12), 3619–3626.

Spence, S., Farrow, T., Herford, A., Wilkinson, I., Zheng, Y., & Woodruff, P. (2001). Behavioural and functional anatomical correlates of deception in humans. *Neuroreport*, *12*(13), 2849–2853.

Spitzer, M., Fischbacher, U., Herrnberger, B., Grön, G., & Fehr, E. (2007). The neural signature of social norm compliance. *Neuron*, *56*(1), 185–196.

Sun, D., Chan, C. C., Hu, Y., Wang, Z., & Lee, T. M. (2015). Neural correlates of outcome processing post dishonest choice: An fMRI and ERP study. *Neuropsychologia*, *68*, 148–157.

Sun, D., Lee, T. M., & Chan, C. C. (2015). Unfolding the spatial and temporal neural processing of lying about face familiarity. *Cerebral Cortex*, *25*(4), 927–936.

Sun, D., Lee, T. M., Wang, Z., & Chan, C. C. (2016). Unfolding the spatial and temporal neural processing of making dishonest choices. *PLoS One*, *11*(4), 1–19.

Sun, D., Shao, R., Wang, Z., & Lee, T. (2018). Perceived gaze direction modulates neural processing of prosocial decision making. *Frontiers in Human Neuroscience*, *12*(52), 1–11.

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science*, *315*(5811), 515–518.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., ... Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, *15*(1), 273–289.

Vartanian, O., Kwantes, P. J., Mandel, D. R., Bouak, F., Nakashima, A., Smith, I., & Lam, Q. (2013). Right inferior frontal gyrus activation as a neural marker of successful lying. *Frontiers in Human Neuroscience*, *7*(616), 1–6.

Wu, C. C., Samanez-Larkin, G. R., Katovich, K., & Knutson, B. (2014). Affective traits link to reliable neural markers of incentive anticipation. *NeuroImage*, *84*, 279–289.

Yang, Z., Huang, Z., Gonzalez-Castillo, J., Dai, R., Northoff, G., & Bandettini, P. (2014). Using fMRI to decode true thoughts independent of intention to conceal. *NeuroImage*, *99*, 80–92.

Yin, L., Hu, Y., Dynowski, D., Li, J., & Weber, B. (2017). The good lies: Altruistic goals modulate processing of deception in the anterior insula. *Human Brain Mapping*, *38*, 3675–3690.

Yin, L., Reuter, M., & Weber, B. (2016). Let the man choose what to do: Neural correlates of spontaneous lying and truth-telling. *Brain and Cognition*, *102*, 13–25.

Yin, L., & Weber, B. (2016). Can beneficial ends justify lying? Neural responses to the passive reception of lies and truth-telling with beneficial and harmful monetary outcomes. *Social Cognitive and Affective Neuroscience*, *11*, 423–432.

Zhu, L., Jenkins, A. C., Set, E., Scabini, D., Knight, R. T., Chiu, P. H., ... Hsu, M. (2014). Damage to dorsolateral prefrontal cortex affects trade-offs between honesty and self-interest. *Nature Neuroscience*, *17*(10), 1319–1321.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.