Review

Neural correlates of second-order verbal deception: A functional near-infrared spectroscopy (fNIRS) study

Xiao Pan Ding a,1, Liyang Sai b,1, Genyue Fu a,⁎, Jiangang Liu c, Kang Lee a,d,⁎

a Zhejiang Normal University, China
b East China Normal University, China
c Beijing Jiaotong University, China
d University of Toronto, Canada

A B S T R A C T

The present study focused on neural correlates underlying second-order deception. In first-order deception, the recipient of deception is unaware of the deceiver’s deceptive intention. However, during second-order deception, the recipient is fully aware of the deceiver’s deceptive intention and thus the deceiver needs to use both lies and truths to deceive the recipient. Using the functional near-infrared spectroscopy (NIRS) methodology and a naturalistic interactive game, we found that second-order deception elicited significantly greater [oxy-Hb] changes in the prefrontal cortex (the right superior frontal gyrus (SFG), BA6) than the non-deceptive control condition. This finding suggests that second-order deception, like first-order deception, engages specifically the cortical regions associated with the planning of complex actions and goal processing. We also found that lying to deceive produced greater neural activities in the right middle frontal gyrus than truth-telling to deceive. This suggests that although both actions serve deceptive purposes, making a false statement contradicting the true state of affairs still requires more executive control and thus greater neural responses in the cortical regions associated with this function. In addition, we found that the successful deception produced greater neural activities in a broad area of the prefrontal frontal cortex than failure to deceive, indicating the involvement of the cortical reward system during second-order deception. Further, failure of truth-telling to deceive produced greater neural responses in the right SFG than failure of lying to deceive. The present findings taken together suggest that second-order deception engages both the cortical executive and reward systems.

⁎ The first two authors contributed equally.

⁎⁎ The present study was supported in part by grants from the National Science Foundation of China (31371041) and Program for Innovative Research Team in Zhejiang Normal University to Genyue Fu, as well as grants from SSHRC to Kang Lee. We would like to thank Fang Yuan for her assistance with the data collection and analysis.

Corresponding authors at: Zhejiang Normal University, 125 Wener Road, Hangzhou, 310012, China.
E-mail addresses: fugy@zjnu.cn (G. Fu), kang.lee@utoronto.ca (K. Lee).

Article info

Article history:
Accepted 10 October 2013
Available online 23 October 2013

Keywords:
Deception
First-order deception
Second-order deception
Functional near-infrared spectroscopy
Reward system

Contents

Introduction .......................................................... 506
Method ................................................................. 507
Participants .......................................................... 507
Procedure .............................................................. 507
Near-infrared spectroscopy (NIRS) recordings .............................................. 508
Data analysis .......................................................... 508
Spatial correlates of deception .......................................................... 508
Time courses of deception .......................................................... 508
Correlational analyses between behavior and NIRS data ..................................... 509
Results ................................................................. 509
Behavioral results .......................................................... 509
NIRS results ............................................................ 510
Deceptive condition vs. non-deceptive control condition ...................................... 510
Lying to deceive trials vs. truth-telling to deceive trials during the action period in the deceptive condition .......................................................... 510
Lying to deceive vs. truth-telling to deceive trials during the feedback period in the deceptive condition .......................................................... 510
The correlation between behavior and NIRS data .......................................................... 511
Introduction

Verbal deception, or lying, is common in our everyday life (DePaulo et al., 2003). People at all ages have been found to tell lies for self-protection, personal gain, as well as prosocial reasons (e.g., white lies). Despite extensive behavioral research on lying for more than a century (Crossman and Lewis, 2006; Talwar and Lee, 2008; see Lee, 2013 for a review), the neural mechanisms underlying deception have only been explored in the last several decades. Earlier studies have mainly used event-related potential (ERP) methodology (for a review, see Rosenfeld and Greely, 2012). Although this methodology has provided invaluable evidence about neural activities associated with lying at an exquisitely fine temporal scale, due to their inherent poor spatial resolution, the brain regions involved in deception were entirely unclear until the last decade.

With the advent of functional magnetic resonance imaging (fMRI) methodologies, researchers over the last decade have learned a great deal about the cortical neural systems associated with not only participants’ judgments of others’ lies (e.g., Harada et al., 2009) but also the act of lying by participants themselves (for reviews, see Abe, 2011). Most of the existing studies typically asked participants to lie about recently learned knowledge, such as words (Abe et al., 2008; Lee et al., 2002, 2005, 2009), cards (Davatzikos et al., 2005; Gamer et al., 2009; Kozel et al., 2005; Langleben et al., 2002, 2005; Phan et al., 2005; Priori et al., 2008), pictures (Ito et al., 2011; Lee et al., 2010), faces (Bhatt et al., 2009), or numbers (Lee et al., 2002). Participants have also been asked to lie about past autobiographic experiences (Abe et al., 2006, 2007; Ganis et al., 2003; Grezes et al., 2004; Lee et al., 2002; Nuñez et al., 2005; Spence et al., 2008). In contrast to the above studies where participants were told to lie or tell the truth according to the experimenter’s wishes, some researchers have recently begun to examine neural activities associated with truth- or lie-telling originated out of the participants’ own volition (e.g., Baumgartner et al., 2009; Ding et al., 2013; Greene and Paxton, 2009).

Despite the diversity of research paradigms used to elicit lying-telling, one of the most consistent findings of all the existing studies is that lying engenders greater activities in the prefrontal cortex (PFC) and anterior cingulate cortex (ACC) than truth-telling (Christ et al., 2009; Gamer et al., 2009; Ganis et al., 2009; Greene and Paxton, 2009; Kozel et al., 2004; Langleben et al., 2002; Lee et al., 2002). The greater deception related activations in these areas are not surprising because such findings are consistent with the conceptualization of deception as an executive control intensive task (Langleben, 2008; Langleben et al., 2002; Sip et al., 2008; Vrij et al., 2006). To deceive, one must inhibit the public disclosure of the true state of affairs and present instead a false state of affairs publicly, which requires a host of executive functions such as inhibition, working memory, and planning (Christ et al., 2009; Gombos, 2006; Johnson et al., 2004; Talwar and Lee, 2008). Indeed, the pivotal role of the PFC on deception is also supported causally by a transcranial direct current stimulation (tDCS) study (Karim et al., 2010).

However, all the above studies have only examined first-order deception whereby the deceiver makes an untruthful statement to intentionally deceive an unsuspecting dupe. However, although first-order deception is perhaps most common in everyday social interactions, second-order deception also occurs frequently. Second-order deception refers to the situation whereby the deceivers are fully aware that the recipient is aware of their deceptive intention but carry out intentional deceptive acts anyway to deceive the recipient. Second-order deception, or “reverse psychology” as it is colloquially called, occurs most frequently in highly competitive situations such as political rivalry and elections, sports, warfare, gambling (e.g., poker games), and even business transactions and diplomacy. In second-order deception, not only do untruthful statements (i.e., lies) serve to deceive the recipient, but also truthful ones. This is because knowing that the recipient knows the deceiver to be deceptive, the deceiver can tell the truth to the recipient such that the recipient may mistake it as deceptive and therefore believe the opposite to be true. Thus, in second-order deception, by telling the truth, the deceivers also accomplish their ultimate goal: to deceive the recipient.

To bridge the important gap in the literature, the present study examined the neural correlates of second-order deception. We aimed to elucidate whether lie- and truth-telling during second-order deception, unlike first-order deception, engenders similar levels of neural activations when lying and when telling the “truth” because truth-telling also serves a deceptive purpose. Alternatively, because both first- and second-order lying require stating something that is different from the true state of affairs whereas second-order “truth-telling” only needs to state the true state of affairs, lying during second-order deception would still engender greater levels of neural activations than “truth-telling” even though the “truth” also serves a deceptive purpose. The first goal of the present study was to test specifically these two intriguing possibilities.

The second goal of the present study was to examine the neural correlates of the success versus failure of one’s deceptive attempt. To date, no neuroimaging studies have specifically examined whether neural responses are similar or different when the deceiver succeeds or fails to dupe the recipient. Evidence regarding the neural correlates of success and failure of verbal deception is essential to achieve a comprehensive understanding of neurocognitive mechanisms underlying verbal deception. This is because deception in everyday life is not a frivolous exercise. Rather, individuals engage in deception to achieve specific personal goals (e.g., gaining a competitive advantage). Thus, one should care a great deal about whether their deceit is successful. Success of one’s deception should be gratifying whereas failure of deception should be disappointing. If this is true, one should expect success of deception to engender greater neural responses in the cortical reward network than failure. More specifically, during second-order deception, because lie- and truth-telling both serve the same deceptive purpose, one possibility is that the success of one’s lies and “truths” may be equally rewarding and thus engender similar neural activations relative to failure. Alternatively, because using the truth-telling to deceive is a relatively novel and perhaps more “cunning” strategy, the success of using this strategy may be more rewarding than the typical lie-telling strategy. If the latter is the case, during second-order deception, the success of truth-telling to deceive the recipient should engender greater neural activations in the cortical reward network than the success of lying.

In the present study, we asked participants to play a coin guessing game with another “participant” (a confederate). In the experimental coin guessing game, the participants could hide a coin in one of their hands. They were told that if the confederate guessed the whereabouts
of the coin correctly, they would lose the coin to the confederate; otherwise, they would win the coin to keep for themselves. Before making the guess, the confederate would always ask the participants first about the whereabouts of the coin. The participants were instructed to respond in such a way that would mislead the confederate to guess the wrong location. They were told that the confederate knew their deceptive intentions. Thus, to mislead the confederate who was clearly aware of their deceptive intention, the participants had to make a decision in each trial to tell either the true or false whereabouts of the coin in order to lead the confederate astray. In total, 30 trials were played in this deceptive game condition. In addition, the same participants participated in a non-deceptive control condition where they did not have to mislead the confederate about the location of a coin.

To make the game as naturalistic as possible while at the same time obtaining data about related cortical hemodynamic responses, we used the functional near-infrared spectroscopy (fNIRS) methodology. Although NIRS does not have as high spatial resolution as the MRI methodology and only measures cortical hemodynamic activities a few centimeters deep in the cortex, it has several advantages, especially, for the purpose of the present study. First, a NIRS machine is quiet, highly mobile, and does not require a specially controlled laboratory environment. Thus, experiments can take place in an environment not as foreign, noisy, dark, and claustrophobic as in an MRI machine (Suda et al., 2010). Second, unlike an fMRI experiment where participants must lie down on a gantry with their head firmly restrained in a head coil, participants in an fNIRS experiment can sit in an upright posture with less head restraining due to NIRS’s relative higher tolerance for motion. Thus, fNIRS is well suited for the present study which required each participant to interact with another person face to face. Third, NIRS has much higher temporal resolution than MRI at 10 Hz. This temporal resolution makes it possible to obtain a recording of the actual time course of one hemodynamic response epoch in response to a specific cognitive task in a specific trial when a slow event-related design is used. In contrast, due to the low temporal resolution, such data are more difficult to obtain with the fMRI methodology. In addition, instead of button-presses, our experimental NIRS setup allows participants to respond verbally. Although verbal deception is very common in our daily life, to date, there are little studies to explore the neural correlates of verbal deception (see Abe et al., 2007; Spence et al., 2008).

We tested the following hypotheses. First, regarding lie- and truth-telling, because the deceptive condition required participants to execute the action of deception whereas the non-deceptive control condition did not, we hypothesized that the deception condition relative to the control condition would produce greater activations in such prefrontal areas as the superior frontal gyrus (SFG) than the control condition (Kozel et al., 2005; Langleben et al., 2002, 2005). Further, because both lying to deceive and truth-telling to deceive require participants to form a deceptious intent and make a deliberative effort to make a statement to deceive the recipient, one possibility is that no difference in neural activities would be found between the trials where the participants told lies and the trials where the participants told the truth. Alternatively, making a false statement to deceive may involve the executive control system more than making a true statement because the former requires participants to inhibit the reporting of the true state of affairs and generate an alternative statement that does not correspond to the truth. Thus, telling a lie to deceive may still be more cognitively taxing than telling the truth to deceive, and thus engender greater neural activations in the prefrontal areas, such as the middle frontal gyrus (MFG), than truth-telling to deceive (Kozel et al., 2005; Mameli et al., 2010).

Regarding success or failure of deceit, all trials would fall into one of the following four categories: lie-success, lie-failure, truth-success, and truth-failure. We hypothesized that the success trials in general would produce greater neural activities in such prefrontal areas as MFG and SFG than the failure trials. This hypothesis was based on the existing findings that the prefrontal regions are the cortical part of the rewards network (McClure et al., 2004; O’Doherty, 2004; Rademacher et al., 2010; Schultz, 2006; Wallis and Miller, 2003). Further, participants might find telling the truth to deceive to be more rewarding, and therefore the truth-success trials would engender greater activations than the lie-success trials.

Method

Participants

Twenty-nine undergraduates participated this study (13 males, with age from 20 to 24 years, M_age = 21.65 years, SD = 1.1 years). Four participants were excluded due to procedural errors. Thus, there were twenty-five valid participants (10 males, 15 females, with age from 20 to 24 years, M_age = 21.72 years, SD = 1.1 years). All participants had normal or corrected to normal vision, and were right-handed. None had a history of any neurological or psychiatric disorders. The research was approved by the university ethics committee. All participants provided informed consent prior to participating in the study and were debriefed after the experiment.

Procedure

Participants were tested individually, seated face-to-face with another “participant” (a confederate) in a quiet room. Each participant completed two sessions: a deceptive game session and a control non-deceptive game session. The deceptive session was always completed first followed by the control session. Before the game started, participants were given 3 practice trials to ensure that they understood the rules of the game.

During the deceptive session, participants were firstly shown a pile of real coins (each valued at 1 Chinese RMB). They were then asked to play a coin guessing game with a confederate and told that the confederate knew their deceptive intentions. Thus, to mislead the confederate who was clearly aware of their deceptive intention, the participants had to make a decision to tell either the true or false whereabouts of the coin to lead the confederate astray. They were also told that if the confederate guessed the whereabouts of the coin correctly, they would lose the coin to the confederate; otherwise, they would win the coin to keep for themselves. Alternatively, they could win all of the coins (30 yuan RMB in total) if they could deceive their confederate on every trial.

The participants could hide a coin in one of their hands while the confederate’s eyes were closed. After the coin was completely hidden, the participant put both hands on the table in a fist position with the back of the hand facing down. The experimenter sitting next to the participant gave the first sound cue (“Please open your eyes”) to the confederate. After 5 s, the confederate was cued by another sound message (“Please ask”) and asked the participant where the coin was hidden (“Which hand is the coin in?”). After a sound cue (“Please answer the question”), the participant was prompted to actually carry out their decision, that is, to reveal verbally the true whereabouts of the coin (the truth) or the false one (a lie). After responding verbally, participants were instructed to remain silent until this period was over. This period lasted for 15 s and is henceforth referred to as the action period.

Then, after another sound cue (“Please guess”) was presented, the confederate tapped one of the participant’s hands to indicate the whereabouts of the coin that was thought to contain the coin. The participant opened both hands to reveal whether the confederate guessed correctly. They were instructed to remain silent again for another 15 s. This period is hence referred to as the feedback period because the participant learned whether they had succeeded or failed in deceiving the confederate. The sequence of these periods in a trial is depicted in Figs. 1a & b. At the end of each trial, the winner was awarded...
one Chinese yuan RMB in full view of the loser so as to create a highly competitive situation to motivate the participants.

In the control non-deceptive session, the procedure was identical to the experimental one except that the participants were told to tell the truth all the time. There was no prize for playing this game. The software package E-prime 2.0 was used to present and control the sound cues.

Near-infrared spectroscopy (NIRS) recordings

A 24-channel continuous wave system (ETG-4000, Hitachi Medical Co., Japan) was used in the present study. The probes of the NIRS machine were fixed using one 9 × 9 cm² rubber shell over the frontal areas. The shell was covered with a nylon-net to keep it attached to the head. The shell of 16 probes, consisting of a 4 × 4 arrays with eight light emitters and eight detectors, was capable of measuring the relative concentrations of hemoglobin at 24 points (see Figs. 1c & d). The lowest probes were positioned along the Fp1–Fp2 line in accordance with the international 10–20 system for electroencephalography. The placements of the probes in the dorsal bilateral frontal areas were based on our hypotheses derived from the existing related studies (Christ et al., 2009; Langleben et al., 2002; Tian et al., 2009). The inter-optode distance was 30 mm, which allowed for measuring neural activities approximately 15–25 mm beneath the scalp. Optical data from individual channels were collected at 2 different wavelengths (695 and 830 nm) and analyzed using the modified Beer–Lambert Law for a highly scattering medium (Cope and Delpy, 1988). Changes in oxygenated ([oxy-Hb]) and deoxygenated ([deoxy-Hb]) signals were calculated in units of millimolar-millimeter (mM mm) (Maki et al., 1995). The sampling rate was set to 10 Hz.

We used a 3D-magnetic space digitizer (EZT-DM401, Hitachi Medical Corporation, Japan) to record the 3 dimensional locations of each optode on each participant's scalp. We used the probabilistic registration method (Singh et al., 2005) to obtain an estimation of each channel's corresponding location in the Montreal Neurological Institute (MNI) space.

Data analysis

We analyzed both [oxy-Hb] and [deoxy-Hb] signals. Because changes in [oxy-Hb] are the most sensitive indicators of regional cerebral blood flow (rCBF) activities whereas [deoxy-Hb] signals are noisy and unreliable as already reported by other researchers previously (Homae et al., 2007), we report here statistical results based on [oxy-Hb] signals (for [deoxy-Hb] results, see Table 1). For each participant, we segmented event epochs from the time course data. Each event epoch consisted of a 2 s period prior to the stimulus onset. For the action period, the onset was when the sound cue, “Please answer the question”, was presented. For the feedback period, the onset was when the sound cue, “Please guess”, was presented. Then, for each of the two periods, we used 13 s of the stimulus and recovery period, and a 2 s post-stimulus period.

Spatial correlates of deception

To explore the spatial correlates underlying the different trial types, the [oxy-Hb] data were analyzed using NIRS-SPM (Jang et al., 2009; Tak et al., 2011; Ye et al., 2009). For each participant, data were pre-processed to remove sources of noise and artifact (such as movement and cardiac activities) using the hemodynamic response function (hrf) filter and wavelet-MDL (minimum description length) detrending algorithm. Then, a general linear model incorporating task effects, a mean and a linear trend were used to compute parameter estimates and t-contrasts (the deceptive condition minus the non-deceptive control condition in action period; lying to deceive vs. truth-telling to deceive trials in the action period). F contrasts were calculated for two-way comparisons: response type (lying to deceive vs. truth-telling to deceive) and outcome (success vs. failure in the feedback period) at each channel. To control for false positives, all p values were corrected by false discovery rate (FDR < .05; Singh and Dan, 2006).

In addition, we also conducted additional tests to examine the possibility of “general arousal” or “mental effort” whereby any of the previous significant findings might be due to increased general arousal or mental effort during the experiment. We performed a repeated measures 2 (period: action vs. feedback period) × 2 (response type: lying to deceive vs. truth-telling to deceive) with period and response type as within-subject variables on the changes of [oxy-Hb] signals across periods. This analysis allowed us to ascertain whether a significant response type effect was not general but specific to a period where such effect was supposed to occur.

Time courses of deception

Because we used a slow-event related design, we were able to obtain the time courses of the grand averaged lying to deceive, truth-telling to deceive, and control trials, respectively. Several steps were required using the raw time series data: to remove baseline drifts and pulsation due to the heartbeat, the raw hemoglobin continuous data was filtered by a high-pass filter of 0.01 Hz and a low-pass filter of 0.3 Hz. Then, the baseline was calculated by the least squares model fitting procedure according to the activities in the pre-stimulus and post-stimulus period. For each epoch of each channel, a first-degree baseline fit to the mean of
The 2 s pre-stimulus and 2 s post-stimulus periods was performed. After baseline correction, we computed the mean and the standard deviation of each channel for each participant and converted the raw time course values to Z scores. Then we averaged all the event epochs for the channel of a particular trial type to derive a grand averaged time course waveform of the channel for the trial type.

Correlational analyses between behavior and NIRS data

To identify whether neural activity associated with participants’ deceptive behaviors, we conducted Pearson correlation analyses between NIRS data and deceptive behaviors. The NIRS data were derived from the period just before the action period. They were the beta values of [oxy-Hb] signals in the lying to deceive condition, truth-telling to deceive condition, and the contrast beta values between these two conditions before the action period. Participants’ deceptive behaviors were measured in the following four ways: (1) percent of lying in the action period after participants’ lying to deceive attempt in the previous trial was successful, (2) percent of lying in the action period after participants’ truth-telling to deceive attempt in the previous trial was successful, (3) percent of lying in the action period after participants’ lying to deceive attempt in the previous trial failed, and (3) percent of lying in the action period after participants’ truth-telling to deceive attempt in the previous trial failed.

Results

Behavioral results

Each participant participated in 45 trials in total, including 30 trials of the experimental deceptive game session and 15 trials of the control non-deceptive session. In the experimental session, the mean number of trials (standard deviation) where the participants told the truth was 15 (3.6) times, and 14 (3.4) times when the participants lied. Further, paired t-tests indicated that there was no difference between the number of truth-telling to deceive trials and the number of lying to deceive trials, \( t_{(1, 24)} = 0.52, p > .05 \).

The mean numbers (SD) of the truth-success, truth-failure, lie-success, and lie-failure trials were 7 (1.7), 8 (2.8), 8 (1.9), and 7 (2.6) times, respectively. A 2 (response type: lying to deceive vs. truth-telling to deceive) × 2 (outcome: success vs. failure) repeated measures ANOVA revealed no significant main effect or interaction (ps > .05).

The mean lying rate after the truth-success, truth-failure, lie-success, and lie-failure trials were 47.97% (24.24%), 48.51% (19.01%), 48.46% (20.20%), and 48.46% (19.01%), respectively.
Note.

X.P. Ding et al. / NeuroImage 87 (2014) 505–514

Table 1
The contrast results about beta value of the hemodynamic activities (within second-order condition).

<table>
<thead>
<tr>
<th>Channel</th>
<th>Estimated MNI</th>
<th>Brain region</th>
<th>Probability</th>
<th>Statistic value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>X</td>
<td>Y</td>
<td>Z</td>
<td></td>
</tr>
<tr>
<td>Action period (deceptive condition minus control non-deceptive condition)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ch20</td>
<td>14</td>
<td>15</td>
<td>70</td>
<td>Right SFG (BA6)</td>
</tr>
<tr>
<td>Ch10</td>
<td>24</td>
<td>53</td>
<td>42</td>
<td>Right MFG (BA9)</td>
</tr>
<tr>
<td>Action period (lying to deceive minus truth-telling to deceive)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ch7</td>
<td>38</td>
<td>54</td>
<td>28</td>
<td>Right MFG (BA46)</td>
</tr>
<tr>
<td>Ch20</td>
<td>14</td>
<td>15</td>
<td>70</td>
<td>Right SFG (BA6)</td>
</tr>
<tr>
<td>Feedback period (lying to deceive minus truth-telling to deceive)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ch2</td>
<td>6</td>
<td>69</td>
<td>23</td>
<td>Right SFG (BA10)</td>
</tr>
<tr>
<td>Ch23</td>
<td>3</td>
<td>11</td>
<td>69</td>
<td>Right SFG (BA6)</td>
</tr>
<tr>
<td>Feedback period (success minus failure)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ch2</td>
<td>6</td>
<td>69</td>
<td>23</td>
<td>Right SFG (BA10)</td>
</tr>
<tr>
<td>Ch4</td>
<td>−27</td>
<td>57</td>
<td>30</td>
<td>Left MFG (BA10)</td>
</tr>
<tr>
<td>Feedback period: interaction between 2 (lying to deceive vs. truth-telling to deceive) and 2 (success vs. failure)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ch2</td>
<td>6</td>
<td>69</td>
<td>23</td>
<td>Right SFG (BA10)</td>
</tr>
<tr>
<td>Ch22</td>
<td>−22</td>
<td>10</td>
<td>70</td>
<td>Left SFG (BA10)</td>
</tr>
</tbody>
</table>

Note. The results are corrected by FDR < .05.
* p < .05.
** p < .01.

(17.30%), and 46.68% (24.59%), respectively. A 2 (response type: lying to deceive vs. truth-telling to deceive) × 2 (outcome: success vs. failure) repeated measures ANOVA revealed no significant main effect or interaction (ps > .05).

NIRS results

Deceptive condition vs. non-deceptive control condition

To investigate the spatial correlates underlying second-order deception, we compared the hemodynamic activities between the deceptive and control conditions during the action period. The GLM results are shown in Table 1 (results were corrected by FDR < .05). Specifically, the deceptive condition elicited significantly larger changes in [oxy-Hb] concentration than the control condition in Channel 20 (BA 6, the right SFG), suggesting that the action to make either a true or false statement with a deceptive intent engendered significantly more [oxy-Hb] activities in the right frontal areas than the action to tell the truth without deceptive intent.

For the time courses of the trials in the deceptive and control conditions in Channel 20, see Fig. 2a.

Lying to deceive trials vs. truth-telling to deceive trials during the action period in the deceptive condition

To investigate the spatial correlates underlying the action period of second-order deception, we compared the hemodynamic activities between the truth-telling to deceive trials and the lying to deceive trials. The GLM results are shown in Table 1 (results were corrected by FDR < .05). Specifically, the lying to deceive trials elicited significantly larger changes in [oxy-Hb] concentration than the truth-telling to deceive trials in Channel 7 (BA 46, the right MFG), suggesting that the action to make false statements engendered significantly more [oxy-Hb] activities in the right frontal areas than the action to make truthful statements although both types of statements served to deceive the recipient. For the time courses of the trials in the lying to deceive condition and the truth-telling to deceive condition in Channel 7, see Fig. 2b.

To test the possibility of “general arousal” or “mental effort”, we found a significant interaction between period and response type in Channel 7 (F (1, 24) = 6.81, p < .05). We performed follow-up post hoc t-tests and found that the changes in [oxy-Hb] signals in the lying to deceive conditions were not different between the action and feedback period. However, in the truth-telling to deceive condition, the changes in [oxy-Hb] signals in the action period were significantly greater than those in the feedback period. This finding further suggests that the differential activations for truth-telling to deceive vs. lying to deceive in the action period were specific to this period.

Lying to deceive vs. truth-telling to deceive trials during the feedback period in the deceptive condition

To investigate the neural correlates during the feedback period of second-order deception, a two-way repeated ANOVA measure was performed on the beta values of hemodynamic activities using two within-subject variables: response type (lying to deceive vs. truth-telling to deceive) and outcome (success vs. failure). The results are shown in Table 1 (results were corrected by FDR < .05). In Channel 2 (BA 10, the right SFG), the truth-telling to deceive trials during the feedback period elicited significantly larger changes in [oxy-Hb] concentration than that of the lying to deceive trials (Table 1) (For the time courses of the trials in the lying to deceive condition and the truth-telling to deceive condition in Channel 2, see Fig. 2c.). In addition, the success trials also elicited greater activations than the failure trials in this channel (For the time courses of the success and failure trials in Channel 2, see Fig. 2d.). Further, a significant interaction between the response type and outcomes was significant. Simple contrast results showed that the truth-failure trials elicited significantly larger [oxy-Hb] changes than the lie-success trials elicited in Channel 2 (BA 10, the right SFG). However, in this channel, no significant difference was found between the lie-success trials and the truth-success trials, suggesting that the feedback about failure after truth-telling engendered significantly more [oxy-Hb] activities in the right SFG than the feedback about failure after lie-telling.

To examine the possibility of “general arousal” in Channel 2, we conducted a repeated measures with period and response type as within-subject variables on the changes of [oxy-Hb] signals across periods, and found that the changes in [oxy-Hb] signals during the action period were significantly higher than that in the feedback period (F (1, 24) = 9.36, p < .01), suggesting that there did not exist a “general arousal” effect that spanned from the action period to the feedback period in Channel 2.
Except for Channel 2, the contrast between the lying to deceive trials and the truth-telling to deceive trials was not significant in any other channels. However, regardless of whether it was the lying to deceive trials or the truth-telling to deceive trials, the success trials elicited significantly larger changes in [oxy-Hb] concentration than the failure trials in all 24 channels. In other words, successful deception engendered a widespread increase in hemodynamic activities in the prefrontal areas than unsuccessful deceptive attempts.

The correlation between behavior and NIRS data

Pearson correlation analyses showed that the [oxy-Hb] activity in Channel 7 (the right MFG, BA46) in the truth-telling to deceive condition were significantly negatively correlated with the percent of lying in the action period after participants’ truth-telling to deceive attempt in the previous trial was successful \((r = -0.46, p < .05)\) (Fig. 3a); the [oxy-Hb] activity in Channel 20 (the right SFG, BA6) in the truth-telling to deceive condition was significantly negatively correlated with the percent of lying in the action period after participants’ lying to deceive attempt in the previous trial failed \((r = -0.62, p < .01)\) (Fig. 3b). Thus, the larger the hemodynamic response in the truth-telling to deceive in the right prefrontal areas, the more likely the participants continued to choose the truth-telling to deceive after their truth-telling to deceive attempt in the previous trial succeeded, or the more likely the participants changed their choice to truth-telling to deceive after their lying to deceive attempt in the previous trial failed.

Discussion

The present study examined the neural correlates of second-order verbal deception. With a naturalistic paradigm, we examined whether second-order deception would engage a network of neural regions similar to that found in studies focusing on first-order deception, whether deceptive actions would engender greater neural activations than non-deceptive actions in the prefrontal cortical regions that have been previously implicated in deception, whether lying to deceive and truth-telling to deceive would activate these prefrontal cortical regions similarly or differently, and whether
lying to deceive and truth-telling to deceive would engender the cortical network of the reward system similarly or differentially.

**Second-order deception condition vs. non-deceptive control condition**

Consistent with our hypothesis, the second-order deception condition elicited significantly greater activation in Channel 20 (the right SFG, BA46) than the non-deceptive control condition. The right SFG is believed to play an important role in the planning of complex, coordinated movements (Baker et al., 1996; Hoshi and Tanji, 2000). It has been also found to be involved in higher cognitive functions, particularly the goal-processing operations (Fincham et al., 2002) and executing an action. In the present paradigm, to deceive, participants had to decide whether to use a lying to deceive strategy or a truth-telling to deceive strategy. This decision must take into consideration how frequently the participants had used the lie or truth option to deceive previously and whether they had been successful. This obviously is a highly demanding executive task. In contrast, in the non-deceptive control condition, participants only needed to report the true state of affairs. It was perhaps due to the difference in executive demands between these tasks that produced the present differential findings between deceptive and non-deceptive acts.

**Lying to deceive vs. truth-telling to deceive in the deceptive condition**

We compared neural activations between the lying to deceive and truth-telling to deceive trials in the action period. We found that the lying to deceive trials elicited significantly greater significant changes in [oxy-Hb] signals than the truth-telling to deceive trials in Channel 7 (the right MFG, BA46). The BA 46 is located on the anterior prefrontal cortex and has been found to be involved in executive function, such as response control (Menon et al., 2001; Ridderinkhof et al., 2004). Response control is considered to be one of the higher cognitive functions. It requires an individual to suppress a prepotent response or behavior while generating and completing a different response or behavior. Response control is believed to be central to lying as lying requires one to inhibit the reporting of the true state of affairs while generating and reporting false information (Kozel et al., 2005; Mameli et al., 2010; Priori et al., 2008). During the action period, lying to deceive required the participants to inhibit reporting where the coin really was, and instead claim the coin to be hidden in the other hand. In contrast, truth-telling to deceive, though involving the same intention to deceive the confederate, only required the participants to report where the coin was really hidden with no demands of either inhibiting the reporting of the true state of affairs or generating a report that is inconsistent with the true state of affairs. It was perhaps for the reason of higher executive demands that lying to deceive produced greater changes in [oxy-Hb] signals in the right MFG. We extended these findings by demonstrating that the differences between lying and truth-telling appear not to be due to the fact that a speaker is honest or dishonest. Rather, the difference is likely rooted in the executive demands involved in generating a truthful report versus one containing false information.

**Success vs. failure of deception in the deceptive condition**

During the feedback period, we found that the successful trials elicited significantly greater neural responses in the prefrontal cortex than the failure trials. The significant differences were widespread across all 24 channels over the prefrontal areas. Because the sole difference between the success and failure trials was winning or losing coins, the difference in neural responses could be attributed to the gaining or losing of the monetary reward. Thus, the heightened activity in the frontal area might reflect the involvement of the cortical reward system. Previous fMRI studies have identified the “reward” neural network (Haber and Knutson, 2010). Within this network, the PFC is a major cortical projection region interfacing reward processing with higher-order cognitive and emotional functions, while the ventral striatum (VS) receives its main cortical input from the PFC and projects back to the PFC (Haber and Knutson, 2010). In the present task, success trials led to gaining coins while failure trials led to losing coins; so cortical regions of the “reward” network might have been activated differentially, resulting in the success trials producing significantly greater [oxy-Hb] signals in all the channels over the PFC. Our finding might also be considered to be a support for the idea that the PFC’s involvement in the reward system is highly broad (Haber and Knutson, 2010; Pochon et al., 2002).

More interestingly, we found that the feedback to the success or failure of truth-telling to deceive elicited significantly greater activations in Channel 2 (the right SFG, BA10) than that to the success or failure of lying to deceive. This effect was also qualified by the success or failure of these trials. When the trials were successful, no significant differences were found between truth-telling to deceive and lying to deceive as though the success of both was of equal rewarding value. However, when both types of trials failed, failure of truth-telling to deceive elicited greater neural responses in the right SFG than failure
of lying to deceive. Thus, although the monetary losses between the lying to deceive and the truth-telling to deceive trials were the same, participants appeared to value the former more. They responded to the failure of truth-telling to deceive more strongly than that of lying to deceive. This is likely due to the fact that truth-telling to deceive might be considered by the participants to be a more “cunning” strategy and thus they may care more about their failure to dupe the confederate, akin to a poker player losing after bluffing as opposed to losing due to having a weaker hand of cards than the opponent.

Further, we found that participants’ differential neural activations between the lying to deceive trials and the truth-telling to deceive trials in the right MFG during the action period were significantly and positively correlated with the differential activations between the two types of trials in the right SFG. Thus, relative to lying to deceive, the greater participants engaged the region associated with executive function for truth-telling to deceive, the greater the activations to truth-telling to deceive in the cortical region associated with the cortical reward system.

It is worth noting that we found the larger the neural responses in the truth-telling to deceive in the right prefrontal areas before the action period, the more the participants choose the truth-telling to deceive after they experienced the success of truth-telling to deceive or the failure of lying to deceive in the previous trial. That is, the neural responses before the action period can predict participants’ actual deceptive behavior. These findings provided evidence that the neural activity difference may mediate individuals’ actual deceptive behaviors (Karim et al., 2010).

Limitations

As mentioned above, the NIRS methodology has its advantages whereby it allows for conducting neuroimaging studies when the participants can perform tasks in a more naturalistic fashion. However, due to the limitation of the cortical depth that the NIRS signals can be detected by the optical detectors on the scalp, we could only assess the cortical regions involved in second-order deception that are closer to the cortical surface. For this reason, we were unable to examine the roles of some important areas such as the ventral striatum, amygdala, and ACC. Specifically, the ventral striatum plays an important part in the reward system (Dreher et al., 2006; Haber and Knutson, 2010; Izuma et al., 2008). However, due to the limitation of NIRS, our study could not assess its role in the feedback stage of deception. In addition, Channel 2 (see Fig. 1) is an inter-hemispheric measurement. It is possible that this channel was sensitive to the vascular changes in the superior sagittal sinus. Additional fMRI studies or fNIRS studies that measure systemic blood flow activities (Boas et al., 2004) are needed to rule out this possibility.

In addition, in the future NIRS studies, one could borrow the design of the fMRI studies that use functional localizers (Poldrack, 2007). For example, one could create a simple deceptive vs. non-deceptive task so as to functionally localize the NIRS channels that are sensitive to the manipulation of deception and a simple success vs. failure task to localize functional channels involved in the reward system. Such approach would allow for more specific functional interpretations of significant findings.

Last, in the present study, we did not compare the first-order and second-order deception directly. Specifically designed studies in the future are needed to compare first-order and second-order deception. For example, two different confederates could be recruited in the study, one who is aware of the participant’s deceptive intent and the other who is unaware of the participant’s deceptive intent, thus creating a second-order and first-order deception conditions, respectively. Comparisons of the neural responses associated with lying to deceive trials between the first-order deception task and the second-order deception task would shed lights on whether second-order deception would engender greater neural activations. Similarly, contrasts between truth-telling in the first-order deception task (a genuine truth-telling act) and truth-telling in the second-order deception task (a deceptive act) can reveal whether second-order deception in the form of truth-telling would result in entirely different neural response patterns from genuine truth-telling.

In summary, the present study focused on neural correlates underlying second-order deception with the use of the near-infrared spectroscopy (NIRS) methodology in a naturalistic interactive game setting. We found that second-order deception elicited significantly greater [oxy-Hb] changes in the prefrontal cortex (the right SFG, BA6) than the non-deceptive control condition. This finding suggests that second-order deception, like first-order deception, engages the cortical regions associated with the planning of complex actions and goal processing. We also found that lying to deceive produced greater neural activities in the right MFG than truth-telling to deceive, suggesting that although both actions serve deceptive purposes, making a false statement contradicting the true state of affairs still requires more executive control and thus greater neural responses in the cortical regions associated with this function. In addition, we found that successful deception produced greater neural activities in a broad area of the prefrontal frontal cortex than failure to deceive, indicating the involvement of the cortical reward system during second-order deception. Further, failure of truth-telling to deceive produced greater neural responses in the right SFG than failure of lying to deceive. The present findings taken together suggest that second-order deception engages both the cortical executive and reward systems.

Conflict of interest

The authors have declared that no competing interests exist.

References
