Electric stimulation of the right temporo-parietal junction induces a task-specific effect in deceptive behaviors

Noguchi, Yasuki
Oizumi, Rei

(Citation)
Neuroscience Research,128:33-39

(Issue Date)
2018-03

(Resource Type)
journal article

(Version)
Accepted Manuscript

Electric stimulation of the right temporo-parietal junction induces a task-specific effect in deceptive behaviors

Yasuki Noguchi[a,*], and Rei Oizumi[a]

[a]Department of Psychology, Graduate School of Humanities, Kobe University, 1-1 Rokkodai, Nada, Kobe, 657-8501, Japan.

*Corresponding author. Tel: +81-78-803-5516, Email: ynoguchi@lit.kobe-u.ac.jp (Y. Noguchi)

Number of figures: 2 (no table)

Abstract

How the brain generates a lie is an important and unsolved issue in neuroscience. Previous studies indicated that mentalizing, the ability to understand and manipulate the mental states of others, plays a critical role in successful deception. Accordingly, recent neuroimaging studies reported deception-related activity in the right temporo-parietal junction (rTPJ), a brain region closely related to the mentalizing ability. Detailed functions of rTPJ in deception, however, remain unclear. In the present study, we investigated a causal relationship between rTPJ and deception using transcranial direct-current stimulation (tDCS). Subjects received anodal tDCS to their rTPJ or V1 (control) and then performed three tasks in which they aimed to deceive another participant to get monetary rewards. In one of the three tasks, we found a significant decrease in a rate of successful deception when rTPJ was stimulated, indicating that neural enhancement of rTPJ caused poorer (not better) deceptive performances. Our results suggest that, in some tasks involving selfish (money-motivated) lying, neural processing in rTPJ does *not* contribute to successful deception through the metalizing ability. Rather, it would be related to the self-monitoring of morally-unacceptable behaviors (lying). The neural enhancement of rTPJ therefore increased the psychological resistance to lying, resulting in poorer deceptive performances.

Key words: transcranial direct current stimulation; false opinion, moral judgment, lying, social context

1. Introduction

Deception is generally defined as a psychological process in which one person deliberately intends to mislead another, typically, by distorting truthful information. Because of its importance in legal, moral, and clinical domains (Blandon-Gitlin et al., 2014; Ekman and Osullivan, 1991; Ford et al., 1988; Vrij et al., 2010; Walczyk et al., 2003), the neural processing underlying deceptive behaviors and lying have been extensively investigated (Garrett et al., 2016; Greene and Paxton, 2009; Kireev et al., 2013; Kozel et al., 2005; Langleben et al., 2002; Nunez et al., 2005; Phan et al., 2005; Spence et al., 2004; Sun et al., 2015; Yin et al., 2016). Those studies consistently indicated a close relationship of deception with the prefrontal cortex in the human brain (Abe, 2011; Christ et al., 2009). For example, previous researches using the functional magnetic resonance imaging (fMRI) reported stronger activity in the prefrontal cortex when participants responded falsely to verbal questions (e.g. "where were you born?") than when they answered truthfully (Ganis et al., 2003; Lee et al., 2002). Causal approaches using transcranial magnetic stimulation (TMS) and transcranial direct-current stimulation (tDCS) provided further evidence for an involvement of the prefrontal cortex in deception (Fecteau et al., 2012; Karim et al., 2010; Karton et al., 2014; Mameli et al., 2010; Priori et al., 2008).

Although those studies showed a pivotal role of the prefrontal cortex, deception is a complex cognitive activity and can be classified into many subtypes. Some of them are selfish and anti-social (e.g. financial fraud), while others not (e.g. white lies in social situations). This diversity of deception suggests that different types of lies can arise from different sets of neural systems, therefore implying deception-related brain regions other than the prefrontal cortex. Recent studies indicate that one of such regions lies in the temporo-parietal junction (Abe et al., 2014; Harada et al., 2009; Hayashi et al., 2014; Sowden et al., 2015). In Abe et al. (2014), subjects read scenarios of events that can happen in real-life situations (e.g. You broke a door lock of the restroom in a department by mistake. A cleaning crew asks you if you know something about the broken lock.) and decided whether to tell a lie or not. They found stronger activation in the right temporo-parietal junction (rTPJ) when the subjects made dishonest decisions (anti-social lying) compared with honest ones.

What was a functional role of rTPJ in deception? At least two interpretations are possible. First possibility was that neural activation in this area played a critical role in deception. It is known that TPJ is related to mentalizing (Saxe and Kanwisher, 2003; Vollm et al., 2006), the ability to understand and manipulate the mental states of others, especially their intentions and beliefs (Frith and Frith, 2006). Since deception is a cognitively-demanding process (Blandon-Gitlin et al., 2014; Vrij et al., 2008), it requires a variety of high-level functions such as decision making, response monitoring, and mentalizing (Sip et al., 2008; Spence et al., 2004). The deception-related activity in rTPJ (Abe et al., 2014) thus might reflect the process of mentalizing that would be necessary for successful lying in the real world. Another interpretation of the rTPJ activity, however, was that this region was engaged in detection and evaluation of deception. Previous researches showed an involvement of TPJ not only in deception but also in moral judgments (Harada et al., 2009; Hayashi et al., 2014; Parkinson et al., 2011; Sellaro et al., 2015; Young et al., 2010). Hayashi et al. (2014) reported stronger hemodynamic responses in TPJ when subjects read scenarios describing protagonist's anti-social lying than pro-social lying, indicating that TPJ played an important role in monitoring violations of social norms. The significant activity in TPJ when subjects decided to tell a lie (Abe et al., 2014) thus might reflect detection or evaluation of morally-unacceptable behavior produced by themselves.

In the present study, we used tDCS to examine those two possibilities. Anodal tDCS was applied to rTPJ just before subjects (actors) performed deceptive behaviors in social (inter-personal) situations. Those behaviors were videotaped and then presented to another groups of subjects (observers) who judged veracity (truthful/deceptive) of those behaviors. If rTPJ plays a critical role in successful deception, the anodal stimulation would facilitate neural processing in rTPJ, which enables better (more believable and convincing) deceptive performances by actors. This would raise a difficulty of the veracity judgment task by observers, resulting in reduced accuracy of that task. In contrast, if rTPJ is related to moral judgments (lie detection), the anodal tDCS would not improve deceptive performances but would enhance sensitivity of this region to immoral behaviors (lying). This might increase a psychological resistance of subjects to telling a lie, resulting in poorer performances of deception. Our causal approach using tDCS thus would reveal a detailed relationship between rTPJ and deception that has been difficult to be elucidated by neuroimaging methods.

2. Materials and Methods

2.1. Subjects

Forty-one subjects participated in the present study (6 as actors and 35 as observers). No statistical method was used to pre-determine a sample size. They were undergraduate students (age: 18 - 22) in Kobe University, Japan. Four additional naïve volunteers (undergraduate students majoring psychology) participated in a tDCS experiment to judge the veracity of the actors' behaviors (inspectors, see below). All had normal or corrected-to-normal vision. Informed consent was received from each subject after the nature of the study had been explained. All experiments were carried out in accordance with guidelines and regulations approved by the ethics committee of Kobe University.

2.2. Basic procedures of a tDCS experiment

Six subjects (3 females) participated in a tDCS experiment as actors. All subjects were healthy and had no contraindication to tDCS. Each subject visited a laboratory on separate two days, approximately a week apart. Three subjects underwent anodal tDCS of rTPJ on the first day, while their mid-occipital region (MO) was stimulated on the second day. An order of the rTPJ and MO sessions was reversed in the remaining three subjects. The MO, corresponding to the primary visual cortex (V1), has been used as a control region in previous tDCS studies on rTPJ (Santiesteban et al., 2015; Sowden et al., 2015). The tDCS was delivered with two saline-soaked surface electrodes (size: 5 × 7 cm^2) connected to a constant-current stimulator (DC-STIMULATOR Plus, neuroConn GmbH, Germany). In the rTPJ session, an anodal electrode was placed over central parietal 6 (CP6), according to the international EEG 10/20 system (Santiesteban et al., 2012; Sellaro et al., 2015; Sowden et al., 2015; Ye et al., 2015). A cathodal electrode was positioned over the vertex (Cz) of each participant as a reference. In the MO session, anodal and cathodal electrodes were placed over Oz and Cz, respectively. We delivered a weak electrical current of 1 mA for 20 min, which was followed by behavioral tasks involving deception (see below). An effect of this offline (preceding the task) stimulation was reported to be more robust than online stimulation (Pirulli et al., 2013; Santiesteban et al., 2015; Sowden et al., 2015), lasting for 90 min beyond an offset of the stimulation period (Nitsche and Paulus, 2001).

Subjects performed three behavioral tasks after tDCS; Shock, Beverage, and Opinion. The Shock and Beverage tasks had two conditions (Expression and Suppression conditions). An order of the three tasks was counter-balanced across sessions (rTPJ/MO) and subjects. As shown in **Figure 1A**, each subject performed those tasks with two persons; an experimenter (second author, R.O.) and an inspector (undergraduate student majoring psychology at Kobe University). We will describe details of the three tasks in the following sections.

2.3. Shock task

In the Shock task, the experimenter delivered an electric shock to the subject through stimulating electrodes attached to right index and middle fingers (**Fig. 1B**). Each shock was given for 200 ms with a frequency of 50 pulses per second (50 Hz). An intensity of the shock was adjusted (before experiment) by the subject to a level that he/she reported was uncomfortable but not painful. In the Expression condition of Shock task (two trials, upper row in **Fig. 1C**), the subject was instructed to behave as if he/she received a shock. In one trial, the experimenter gave a shock (real shock) to the subject by pressing a button on a stimulator (SEN-5201, Nihon Kohden, Japan), and the subject honestly reacted to the shock (truthful trial). In the other trial, however, the experimenter turned off the stimulator (unknown to the subject and inspector) and then pressed a button (sham shock). The subject pretended to receive a shock in response to a click sound of the button, producing deceptive behaviors (deceptive trial). An order of the real-shock and sham-shock trials was randomly determined by the experimenter. The inspector, sitting in front of the subject, observed the subject's behaviors in the two trials and indicated a trial in which the subject behaved deceptively (1st or 2nd). If the inspector answered incorrectly, the subjects obtained an extra reward of 100 yen (about 0.88 dollar). This reward for successful deception was to motivate the subject, prompting his/her convincing performances to deceive the inspector. In the Suppression condition of Shock task (two trials, lower row in **Fig. 1C**), basic procedures were the same as the Expression condition, except for an instruction to the subject. We asked the subject to behave as if he/she received *no* shock. The subject inhibited their reactions to a real shock (deceptive trial) while he/she behaved honestly to a sham shock (truthful trial). The inspector answered a trial in which he thought the subject behaved deceptively (pretending *not* to be stimulated).

2.4. Beverage task

Structures of the Beverage task were similar to the Shock task. Subjects drank a shot glass of apple-cider vinegar (Ringo-Su, Aichi, Japan) diluted with apple juice, instead of receiving an electric shock. A dilution ratio was adjusted by each subject to a level close to a limit within which he/she could perform deceptively. In one trial of the Expression condition, the subject took a whole glass of the diluted vinegar without interruption and reacted honestly (truthful trial). In the other trial, he/she took a glass of apple juice with no vinegar, behaving as if he/she had taken vinegar (deceptive trial). An order of the vinegar and no-vinegar trials was randomly determined by the experimenter. He prepared a glass for each trial and passed it to the subject. The inspector observed the subject's behaviors in the two trials, indicating deceptive one (1st or 2nd). Since the diluted vinegar was visually identical to apple juice, the inspector had to make a judgment solely based on the subject's behaviors. The subject got an extra reward (100 yen) when he/she successfully deceived the inspector. We applied the same procedures to the Suppression condition of this task, except that the subject was asked to behave as if he/she took *no* vinegar.

2.5. Opinion task

The opinion task in the present study was based on the false-opinion paradigm in previous studies (Mehrabian, 1971; Sowden et al., 2015; Wright et al., 2012, 2013). First, the subject gave his/her opinion on a highly-controversial topic in Japan (e.g. the nuclear power generation or NPG). He/she then prepared two statements on each topic; a truthful statement reflecting his/her own view and a deceptive statement in which he/she expressed an opinion opposite to his/her own. For example, if the subject agreed with the present system of NPG in Japan, he/she made a truthful statement expressing "pro" position on this topic, as follows;

"I am in favor of the present system of nuclear power generation. Compared to thermal and water power generation that can pollute the environment, nuclear power generation is clean in principle, emitting less carbon dioxide. It is also achieving stable supply of power at a lower cost. Stopping nuclear plants will cause serious problems on our daily lives".

He/she also made a deceptive statement expressing "con" position as below.

"I am opposed to the present system of nuclear power generation. As you can see from disasters in Chernobyl and Fukushima, nuclear plants carry a significant risk of environmental destruction. This problem is especially serious here in Japan, a country with a high frequency of earthquake and tsunami. We also have to find a solution to the problem of how to process nuclear waste safely".

The subject gave those two statements, one by one, to an inspector. An order of the statements (pro/con) was randomly determined by an experimenter. The inspector judged which statement was a deceptive one (1st or 2nd). The subject gained an extra reward (100 yen) for successful deception. Since there were two sessions in the tDCS experiment (rTPJ and MO), we prepared two topics for the task; the NPG and the capital-punishment (death-penalty) system (CP). Each subject performed the opinion task of one topic on the first day (e.g. NPG), while he/she performed the task of the other topic on the second day (e.g. CP). An order of the two topics was counter-balanced across subjects.

2.6. Video-Recordings of subjects' behaviors

In all three tasks above, truthful and deceptive behaviors by subjects were videotaped with a HD digital camera (HDR-PJ630V, Sony, Japan). We placed the camera at a location right next to the inspector (**Fig. 1A**), to record the subjects' behaviors from a viewpoint of the inspector. One-day experiment comprised 5 conditions of 3 tasks (Shock Expression, Shock

Suppression, Beverage Expression, Beverage Suppression, and Opinion). Each condition comprised 2 trials (truthful, deceptive), resulting in 10 movies per day. Since we conducted the tDCS experiment on 12 days (6 actors × rTPJ/MO), a total of 120 movies were obtained. Those movies (with sounds) were later edited by the experimenter and used for a veracity judgment task (see below). An average length of one movie after the editing was 6.58 sec for the Shock task and 7.29 sec for the Beverage task. Movies of the Opinion task were longer, around 30 sec. Specifically, a mean (± SE) length of movies of truthful statements made after the rTPJ stimulation (truthful/rTPJ) was 31.0 ± 4.1 sec. The movies of deceptive/rTPJ, truthful/MO, and deceptive/MO were 29.8 ± 5.2, 32.7 ± 5.4, 30.0 ± 4.6 sec, respectively. A two-way ANOVA of statement (truthful/deceptive) × tDCS (rTPJ/MO) indicated no main effect or interaction (main effect of statement: $F_{1,5} = 0.29$, $p = 0.61$, $\eta^2 = 0.06$, main effect of tDCS: $F_{1,5} = 0.04$, $p = 0.84$, $\eta^2 = 0.01$, interaction: $F_{1,5} = 0.19$, $p = 0.68$, $\eta^2 = 0.04$).

2.7. Procedures of veracity judgment task

Thirty-five subjects (22 females) participated in a veracity judgment task as observers. Each observer viewed all 120 movies recorded in the tDCS experiment. The task consisted of six blocks; Shock Expression (12 trials), Shock Suppression (12 trials), Beverage Expression (12 trials), Beverage Suppression (12 trials), Opinion NPG (6 trials), and Opinion CP (6 trials). In all blocks, subjects sequentially viewed two movies, one of truthful and the other of deceptive behaviors by the same actor. They answered a movie in which the actor performed deceptively (1st or 2nd). An instruction of the Shock Expression block, for example, was as follows;

"In each trial, we will show you two movies of one person (actor). In one movie, he/she received an electric shock and reacted honesty, typically, by showing a painful expression with a groan. In the other movie, however, he/she had no shock. Nevertheless, he/she pretended to have a shock because we had asked him/her to do so. Please watch those two movies carefully, and indicate a movie of deceptive behaviors, 1st or 2nd. Note that, when those movies were recorded, the actor performed truthful and deceptive behaviors in a random order. We also randomized an order of the two movies in an editing process. This block comprises 12 trials in total. We will put a short pause of 5 seconds after each trial, so that you can check one of two boxes, 1st or 2nd, on the sheet. You will see 6 actors across the 12 trials".

Movies in 6 (out of 12) trials were taken from the rTPJ session of the tDCS experiment (when actors underwent anodal tDCS to their rTPJ), while those in the remaining 6 trials were taken from the MO session. An order of those two groups of movies was randomized. A trial was considered as correct when an observer chose a movie of deceptive behaviors. We calculated accuracy of the veracity judgment task, separately for movies of rTPJ and MO sessions, by pooling the data of 35 observers. Lower accuracy of this task indicates a higher rate of successful deception. Those procedures in the Shock Expression block were repeated for the other five blocks. An order of the six blocks was counter-balanced across observers. To motivate observers, we announced before the first block that an observer achieving accuracy of 70 % or higher throughout the six blocks would get an extra reward of 1,000 yen. No observer actually got the reward (range of accuracy: 33.3 - 58.3%).

As supplementary data, we asked all observers about their own position (pro/con) on issues of NPG and CP. At the beginning of Opinion NPG and CP blocks, observers rated a degree to which they agreed or disagreed with the present NPG and CP system, using a four-point scale from 1 (strongly agree) to 4 (strongly disagree). We also checked, at the end of an experiment, whether observers had known actor(s) in person (as friends and roommates, etc.). Presenting pictures of six actors, we asked observers to check boxes on the sheet if they had seen the actors before. No observer claimed personal acquaintance with any actors.

3. Results

Figure 2A shows accuracy of the veracity judgment task by observers in the three task blocks (Shock, Beverage, and Opinion). Data in the Expression and Suppression blocks were pooled within each task because no systematic difference was found between the two blocks. The gray bars show accuracy when observers judged the veracity of truthful/deceptive behaviors by actors who had received the anodal stimulation of their rTPJ. The white bars show accuracy of the veracity judgment on behaviors produced after the stimulation of MO (control region). A two-way ANOVA of task (Shock/Beverage/Opinion) × tDCS (rTPJ/MO) indicated no main effect of task ($F_{2,68}$ = 1.67, p = 0.20, η^2 = 0.05, *power* = 0.34) or tDCS ($F_{1,34}$ = 1.69, p = 0.20, η^2 = 0.05, *power* = 0.24) but yielded a significant interaction ($F_{2,68}$ = 6.68, p = 0.002, η^2 = 0.16, *power* = 0.90). Post-hoc comparisons with the Bonferroni correction revealed a significant difference between rTPJ and MO only in the Opinion block (corrected p = 0.027). The non-significant main effect of task suggested that difficulties of the veracity judgment were balanced across the three task blocks. The selective increase in accuracy in the Opinion block (rTPJ > MO) indicates a decrease in difficulty of the veracity judgment, suggesting that anodal stimulation of rTPJ induced poorer deceptive performances by actors.

Figure 2B shows accuracy of veracity judgment for each topic of the Opinion block (NPG and CP). A two-way ANOVA of topic (NPG/CP) × tDCS (rTPJ/MO) indicated a significant main effect of tDCS ($F_{1,34}$ = 5.32, p = 0.027, η^2 = 0.14, *power* = 0.61). No main effect of topic ($F_{1,34}$ = 2.45, p = 0.13, η^2 = 0.07, *power* = 0.33) or interaction ($F_{1,34}$ = 1.23, p = 0.28, η^2 = 0.04, *power* = 0.19) was observed. The increased accuracy in the rTPJ compared to

MO conditions (**Fig. 2A**) thus did not depend on a specific topic in the Opinion block.

Figure 2C shows an individual plot of accuracy calculated for each actor. In five out of the six actors, accuracy of the Opinion task was higher in the rTPJ than MO conditions. This relationship, however, was reversed (rTPJ < MO) in one actor (shown by triangles).

4. Discussion

We presently investigated an effect of anodal tDCS on three tasks involving deceptive behaviors. Compared to the control (MO) condition, anodal stimulation of rTPJ of actors induced a modest but significant decrease in a rate of successful deception in one of the three tasks (**Fig 2A**. right). Although there have been several TMS and tDCS studies stimulating rTPJ (Mai et al., 2016; Santiesteban et al., 2012, 2015; Sellaro et al., 2015; Sowden et al., 2015; Ye et al., 2015), this is the first study, to our knowledge, reporting an influence of the rTPJ stimulation on performances of deceptive behaviors.

4.1.Successful deception in inter-personal situations

In the present study, we evaluated deceptive performances by actors (senders) based on accuracy of the veracity judgment task by observers (receivers). This sender-receiver paradigm (Wright et al., 2013) makes a contrast with previous tDCS studies in which deceptive performances were evaluated by behavioral measures of senders, such as reaction times in telling lies (Fecteau et al., 2012; Karim et al., 2010; Mameli et al., 2010; Priori et al., 2008). Our use of the sender-receiver paradigm was related to the fact that we presently stimulated rTPJ (not prefrontal cortex as in the previous studies), a brain region playing a critical role in social functions. In many social contexts, a purpose of lying is to deceive others by doing everything that seems effective (using feigned expressions and plausible speech, etc.). A good liar therefore indicates a person who can deceive others with a high rate of success, regardless of whether their reaction times are shorter or longer. In other words, lying in social situations would be better evaluated from a viewpoint of receivers, rather than based on changes in specific behaviors of senders. Accuracy of the veracity judgment task thus can provide a comprehensive and socially-defined measure on how deceptive behaviors by senders looked convincing in inter-personal situations.

On the other hand, we admit that our sender-receiver paradigm has several problems. A representative one was the low accuracy of the veracity judgment. It is well known that people are generally poor at detecting deception, achieving an average of 54 % correct lie-truth judgments (Bond and DePaulo, 2006). Consistently, accuracy in the present study was less than 50 % in most conditions (**Fig. 2**). It was thus possible that the low accuracy (high difficulty) of our veracity judgment task produced, for example, a floor effect, resulting in an underestimation of neural enhancement by anodal tDCS. Future studies need to use a better paradigm that can resolve this issue. One promising way would be that veracity judgments are made through a direct interaction between senders and receivers (not through an assessment of videotaped behaviors as in the present study). Such interaction enables receivers to use several techniques (e.g. questioning) to elicit and enhance cues to deception (Vrij and Granhag, 2012), which would make the veracity judgment easier and eliminate a possibility of the floor effect.

4.2. Task-selective effect of tDCS

As shown in **Figure 2A**, an effect of tDCS in the present study was task-selective; an anodal stimulation changed accuracy of the veracity judgment in the Opinion task while no effect was observed in the Shock or Beverage task. This selective effect excludes a possibility that tDCS affected general processes of deception, such as an inhibition of truthful responses (e.g. reactions to an electric shock) and a production of misleading (pretending) behaviors.

Rather, it indicates that tDCS interacted with psychological processes specific to the Opinion task. For example, one characteristic of the Opinion task was that subjects (actors) first clarified their position (for/against) on a controversial topic (NPG and CP). They were then asked to express an opinion opposite to their own view, in order to get monetary rewards. Such unfaithful and money-motivated behaviors would cause some sort of a moral conflict in subjects, which activated a network of brain regions including rTPJ (Hayashi et al., 2014; Parkinson et al., 2011). The task-selective effect of tDCS thus suggests an involvement of psychological factors related to rTPJ (e.g. moral) specifically activated in the Opinion task.

4.3. Disruption of successful deception by activating neural functions of rTPJ

Why did the anodal tDCS of rTPJ decrease (not increase) a rate of successful deception? As described in **Introduction**, we presume this would be related to functions of rTPJ that detects and evaluates lies. Some fMRI studies reported stronger hemodynamic responses of rTPJ to anti-social than pro-social lying (Harada et al., 2009; Hayashi et al., 2014). Consistently, a recent study showed that anodal tDCS of rTPJ improved a detection rate of lies produced by other people (Sowden et al., 2015). In the present study, subjects expressed their opinion to an inspector (and a camera) for tens of seconds (about 30 sec on average). Such a long speech would give them a number of opportunities to monitor their own behaviors during the task. If rTPJ reacts to lies produced not only by others but also by themselves, anodal tDCS would enhance sensitivity of this region, making neural responses to their own lies even stronger. This would increase a psychological resistance to telling a lie, resulting in poorer performances of deception in the rTPJ than MO conditions.

4.4. Limitations and future direction

We finally refer to several limitations of the present study. First, while anodal tDCS induced poorer deceptive performances of the Opinion task in five actors, this effect was *not* observed in one actor (**Fig. 2C**). This inter-individual variability in response to tDCS has been repeatedly reported by recent studies (Laakso et al., 2015; Lopez-Alonso et al., 2014; Wiethoff et al., 2014). One source of this variability was differences in anatomical features of the subject's brain, such as a thickness of a layer of cerebrospinal fluid (Laakso et al., 2015). Although structural brain images of subjects were unavailable in the present study, those images would be of great help in predicting an effect of tDCS and its variability across subjects.

The second limitation was an interpretation of the task-specific effect of tDCS (**Fig. 2**). As described above, this effect indicates an interaction of anodal tDCS with some psychological process in the Opinion task. Although a previous literature on rTPJ suggested an involvement of a moral factor, the Opinion task differed from the Shock and Beverage tasks in many other aspects (a use of language, cognitive demands related to working memory, etc.). Since all of these factors might contribute to the task-specific effect in **Figure 2**, our present data should be interpreted cautiously.

The third limitation was a spatial resolution of tDCS. Simulation studies have shown that tDCS can affect neural activity in a wide region not only under two (anodal and cathodal) electrodes but also between the electrodes (Laakso et al., 2016). It was thus probable that brain regions adjacent to rTPJ were also affected by anodal tDCS in the present study. In addition, we used a large size (5×7 cm^2) of electrodes to stimulate rTPJ. Although this procedure enables a direct comparison of our data with previous studies using the same size of electrodes (Mai et al., 2016; Santiesteban et al., 2015; Sellaro et al., 2015; Sowden et al., 2015), future studies should use smaller electrodes for a focal stimulation of rTPJ.

Conflicts of interest

None of the authors have potential conflicts of interest to be disclosed.

Acknowledgment

This work was supported by KAKENHI Grants (22680022 and 26700011 to Y.N) from the Japan Society for the Promotion of Science (JSPS).

References

Abe, N., 2011. How the brain shapes deception: an integrated review of the literature. Neuroscientist 17, 560-574.

Abe, N., Fujii, T., Ito, A., Ueno, A., Koseki, Y., Hashimoto, R., Hayashi, A., Mugikura, S., Takahashi, S., Mori, E., 2014. The neural basis of dishonest decisions that serve to harm or help the target. Brain Cogn 90, 41-49.

Blandon-Gitlin, I., Fenn, E., Masip, J., Yoo, A.H., 2014. Cognitive-load approaches to detect deception: searching for cognitive mechanisms. Trends Cogn Sci 18, 441-444.

Bond, C.F., Jr., DePaulo, B.M., 2006. Accuracy of deception judgments. Pers Soc Psychol Rev 10, 214-234.

Christ, S.E., Van Essen, D.C., Watson, J.M., Brubaker, L.E., McDermott, K.B., 2009. The contributions of prefrontal cortex and executive control to deception: evidence from activation likelihood estimate meta-analyses. Cereb Cortex 19, 1557-1566.

Ekman, P., Osullivan, M., 1991. Who Can Catch a Liar. Am Psychol 46, 913-920.

Fecteau, S., Boggio, P., Fregni, F., Pascual-Leone, A., 2012. Modulation of untruthful responses with non-invasive brain stimulation. Front Psychiatry 3, 97.

Ford, C.V., King, B.H., Hollender, M.H., 1988. Lies and liars: psychiatric aspects of prevarication. Am J Psychiatry 145, 554-562.

Frith, C.D., Frith, U., 2006. How we predict what other people are going to do. Brain Res 1079, 36-46.

Ganis, G., Kosslyn, S.M., Stose, S., Thompson, W.L., Yurgelun-Todd, D.A., 2003. Neural correlates of different types of deception: an fMRI investigation. Cereb Cortex 13, 830-836.

Garrett, N., Lazzaro, S.C., Ariely, D., Sharot, T., 2016. The brain adapts to dishonesty. Nat Neurosci 19, 1727-1732.

Greene, J.D., Paxton, J.M., 2009. Patterns of neural activity associated with honest and dishonest moral decisions. Proc Natl Acad Sci U S A 106, 12506-12511.

Harada, T., Itakura, S., Xu, F., Lee, K., Nakashita, S., Saito, D.N., Sadato, N., 2009. Neural correlates of the judgment of lying: A functional magnetic resonance imaging study. Neurosci Res 63, 24-34.

Hayashi, A., Abe, N., Fujii, T., Ito, A., Ueno, A., Koseki, Y., Mugikura, S., Takahashi, S., Mori, E., 2014. Dissociable neural systems for moral judgment of anti- and pro-social lying. Brain Res 1556, 46-56.

Karim, A.A., Schneider, M., Lotze, M., Veit, R., Sauseng, P., Braun, C., Birbaumer, N., 2010. The truth about lying: inhibition of the anterior prefrontal cortex improves deceptive behavior. Cereb Cortex 20, 205-213.

Karton, I., Rinne, J.M., Bachmann, T., 2014. Facilitating the right but not left DLPFC by TMS decreases truthfulness of object-naming responses. Behav Brain Res 271, 89-93.

Kireev, M., Korotkov, A., Medvedeva, N., Medvedev, S., 2013. Possible role of an error detection mechanism in brain processing of deception: PET-fMRI study. Int J Psychophysiol 90, 291-299.

Kozel, F.A., Johnson, K.A., Mu, Q.W., Grenesko, E.L., Laken, S.J., George, M.S., 2005. Detecting deception using functional magnetic resonance imaging. Biol Psychiat 58, 605-613.

Laakso, I., Tanaka, S., Koyama, S., De Santis, V., Hirata, A., 2015. Inter-subject Variability in Electric Fields of Motor Cortical tDCS. Brain Stimul 8, 906-913.

Laakso, I., Tanaka, S., Mikkonen, M., Koyama, S., Sadato, N., Hirata, A., 2016. Electric fields of motor and frontal tDCS in a standard brain space: A computer simulation study. Neuroimage 137, 140-151.

Langleben, D.D., Schroeder, L., Maldjian, J.A., Gur, R.C., McDonald, S., Ragland, J.D., O'Brien, C.P., Childress, A.R., 2002. Brain activity during simulated deception: an event-related functional magnetic resonance study. Neuroimage 15, 727-732.

Lee, T.M., Liu, H.L., Tan, L.H., Chan, C.C., Mahankali, S., Feng, C.M., Hou, J., Fox, P.T., Gao, J.H., 2002. Lie detection by functional magnetic resonance imaging. Hum Brain Mapp 15, 157-164.

Lopez-Alonso, V., Cheeran, B., Rio-Rodriguez, D., Fernandez-Del-Olmo, M., 2014. Inter-individual variability in response to non-invasive brain stimulation paradigms. Brain Stimul 7, 372-380.

Mai, X., Zhang, W., Hu, X., Zhen, Z., Xu, Z., Zhang, J., Liu, C., 2016. Using tDCS to Explore the Role of the Right Temporo-Parietal Junction in Theory of Mind and Cognitive Empathy. Front Psychol 7, 380.

Mameli, F., Mrakic-Sposta, S., Vergari, M., Fumagalli, M., Macis, M., Ferrucci, R., Nordio, F., Consonni, D., Sartori, G., Priori, A., 2010. Dorsolateral prefrontal cortex specifically processes general - but not personal - knowledge deception: Multiple brain networks for lying. Behav Brain Res 211, 164-168.

Mehrabian, A., 1971. Nonverbal Betrayal of Feeling. J Exp Res Pers 5, 64-73.

Nitsche, M.A., Paulus, W., 2001. Sustained excitability elevations induced by transcranial DC motor cortex stimulation in humans. Neurology 57, 1899-1901.

Nunez, J.M., Casey, B.J., Egner, T., Hare, T., Hirsch, J., 2005. Intentional false responding shares neural substrates with response conflict and cognitive control. Neuroimage 25, 267-277.

Parkinson, C., Sinnott-Armstrong, W., Koralus, P.E., Mendelovici, A., McGeer, V., Wheatley, T., 2011. Is morality unified? Evidence that distinct neural systems underlie moral judgments of harm, dishonesty, and disgust. J Cogn Neurosci 23, 3162-3180.

Phan, K.L., Magalhaes, A., Ziemlewicz, T.J., Fitzgerald, D.A., Green, C., Smith, W., 2005. Neural correlates of telling lies: a functional magnetic resonance imaging study at 4 Tesla. Acad Radiol 12, 164-172.

Pirulli, C., Fertonani, A., Miniussi, C., 2013. The role of timing in the induction of neuromodulation in perceptual learning by transcranial electric stimulation. Brain Stimul 6, 683-689.

Priori, A., Mameli, F., Cogiamanian, F., Marceglia, S., Tiriticco, M., Mrakic-Sposta, S., Ferrucci, R., Zago, S., Polezzi, D., Sartori, G., 2008. Lie-specific involvement of dorsolateral prefrontal cortex in deception. Cereb Cortex 18, 451-455.

Santiesteban, I., Banissy, M.J., Catmur, C., Bird, G., 2012. Enhancing social ability by stimulating right temporoparietal junction. Curr Biol 22, 2274-2277.

Santiesteban, I., Banissy, M.J., Catmur, C., Bird, G., 2015. Functional lateralization of temporoparietal junction - imitation inhibition, visual perspective-taking and theory of mind. Eur J Neurosci 42, 2527-2533.

Saxe, R., Kanwisher, N., 2003. People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind". Neuroimage 19, 1835-1842.

Sellaro, R., Guroglu, B., Nitsche, M.A., van den Wildenberg, W.P., Massaro, V., Durieux, J., Hommel, B., Colzato, L.S., 2015. Increasing the role of belief information in moral judgments by stimulating the right temporoparietal junction. Neuropsychologia 77, 400-408.

Sip, K.E., Roepstorff, A., McGregor, W., Frith, C.D., 2008. Detecting deception: the scope and limits. Trends Cogn Sci 12, 48-53.

Sowden, S., Wright, G.R., Banissy, M.J., Catmur, C., Bird, G., 2015. Transcranial Current Stimulation of the Temporoparietal Junction Improves Lie Detection. Curr Biol 25, 2447-2451.

Spence, S.A., Hunter, M.D., Farrow, T.F., Green, R.D., Leung, D.H., Hughes, C.J., Ganesan, V., 2004. A cognitive neurobiological account of deception: evidence from functional neuroimaging. Philos Trans R Soc Lond B Biol Sci 359, 1755-1762.

Sun, D., Lee, T.M., Chan, C.C., 2015. Unfolding the spatial and temporal neural processing of lying about face familiarity. Cereb Cortex 25, 927-936.

Vollm, B.A., Taylor, A.N., Richardson, P., Corcoran, R., Stirling, J., McKie, S., Deakin, J.F., Elliott, R., 2006. Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. Neuroimage 29, 90-98.

Vrij, A., Granhag, P.A., 2012. Eliciting cues to deception and truth: What matters are the questions asked. J Appl Res Mem Cogn 1, 110-117.

Vrij, A., Granhag, P.A., Porter, S., 2010. Pitfalls and Opportunities in Nonverbal and Verbal Lie Detection. Psychol Sci Public Interest 11, 89-121.

Vrij, A., Mann, S.A., Fisher, R.P., Leal, S., Milne, R., Bull, R., 2008. Increasing cognitive load to facilitate lie detection: The benefit of recalling an event in reverse order. Law Hum Behav 32, 253-265.

Walczyk, J.J., Roper, K.S., Seemann, E., Humphrey, A.M., 2003. Cognitive mechanisms underlying lying to questions: Response time as a cue to deception. Appl Cognitive Psych 17, 755-774.

Wiethoff, S., Hamada, M., Rothwell, J.C., 2014. Variability in response to transcranial direct current stimulation of the motor cortex. Brain Stimul 7, 468-475.

Wright, G.R., Berry, C.J., Bird, G., 2012. "You can't kid a kidder": association between production and detection of deception in an interactive deception task. Front Hum Neurosci 6, 87.

Wright, G.R., Berry, C.J., Bird, G., 2013. Deceptively simple ... The "deception-general" ability and the need to put the liar under the spotlight. Front Neurosci 7, 152.

Ye, H., Chen, S., Huang, D., Zheng, H., Jia, Y., Luo, J., 2015. Modulation of Neural Activity in the Temporoparietal Junction with Transcranial Direct Current Stimulation Changes the Role of Beliefs in Moral Judgment. Front Hum Neurosci 9, 659.

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

Young, L., Camprodon, J.A., Hauser, M., Pascual-Leone, A., Saxe, R., 2010. Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. Proc Natl Acad Sci U S A 107, 6753-6758.

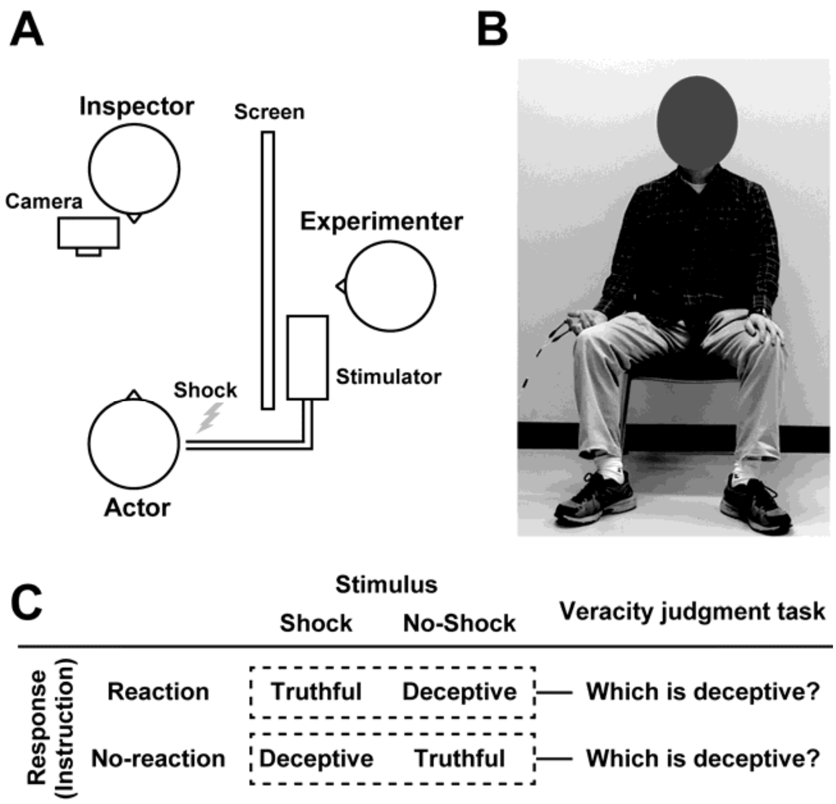

Figure 1. Settings and structures of Shock task in a tDCS experiment. (**A**) Settings. After receiving anodal tDCS to the rTPJ or mid-occipital (MO) cortex (control) for 20 min, a subject (actor) performed three behavioral tasks involving deception; Shock, Beverage, and Opinion. In the Shock task, an experimenter delivered an electric shock (real or sham) to the subject by pressing a button of a stimulator. An inspector observed the subject, judging whether his/her behaviors were truthful or deceptive. A screen was placed in front of the experimenter so that the actor and inspector did not see him directly. (**B**) A sample image of a subject recorded by a camera beside an inspector. All behaviors (reactions) of the subject were videotaped and later used in a veracity judgment task (**Fig. 2**). Although masked in this figure (for a privacy reason), subjects' faces were fully visible in actual veracity judgments. (**C**) Structures of the task. In the Expression condition (upper row), we instructed a subject to behave as if he/she received a shock. In one trial, the subject received a shock and reacted to it honestly (truthful trial). In the other trial, he/she received no shock but pretended to be stimulated (deceptive trial). The inspector indicated which trial was a deceptive one by

observing the subject's behaviors. In the Suppression condition (lower row), we instructed the subject to behave as if he/she received *no* shock. The subject behaved honestly when no shock was given (truthful trial), while he/she behaved dishonestly when stimulated (deceptive trial). The inspector, as in the Expression condition, indicated which trial was deceptive.

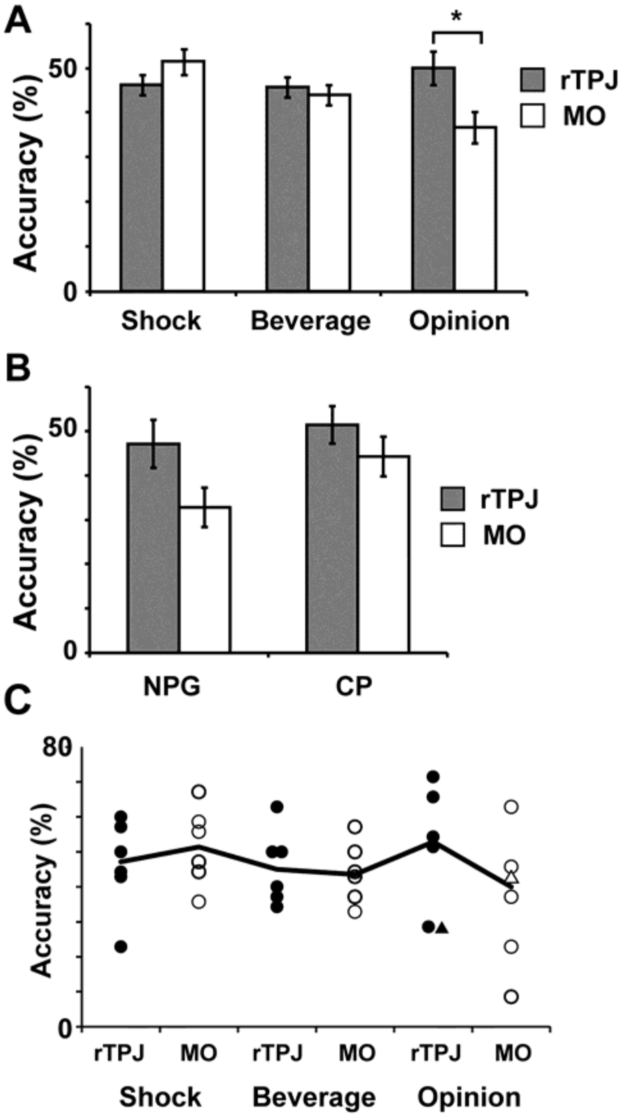

Figure 2. Results of a veracity judgment task. (**A**) Accuracy of veracity judgment in three task blocks. Movies of truthful and deceptive behaviors by actors (**Fig. 1**) were presented to another group of subjects (observers). The observers judged the veracity of those behaviors, as inspectors did in the tDCS experiment. The gray and white bars show accuracy when observers judged the veracity of truthful/deceptive behaviors by actors who had received the anodal stimulation of their rTPJ (gray) and MO (white). We found higher accuracy in the rTPJ than MO conditions in the Opinion block (see **Materials and Methods** for details). (**B**) Accuracy of the veracity judgment for each topic of the Opinion block. Actors made truthful and deceptive statements on two controversial issues in Japan; nuclear power generation (NPG) and the capital punishment (CP) system. In both topics, accuracy of the veracity

judgment was higher in the rTPJ than MO conditions (as shown by a significant main effect of ANOVA, see **Results**). Those increases in accuracy indicate decreases in difficulty of the veracity judgment by the observers, suggesting that anodal tDCS to rTPJ induced poorer deceptive performances by actors. Error bars denote SE across subjects. *corrected $p < 0.05$ (**C**) Individual analyses on accuracy of the veracity judgment task. Each point denotes the data of each actor, with medians across the six actors shown by black lines. In the Opinion task, accuracy was higher in rTPJ than MO conditions in five actors but was lower in one actor (shown by triangles).