The Truth about Lying: Inhibition of the Anterior Prefrontal Cortex Improves Deceptive Behavior

Ahmed A. Karim[1,2], Markus Schneider[1,2], Martin Lotze[1,3], Ralf Veit[1], Paul Sauseng[1,4], Christoph Braun[1] and Niels Birbaumer[1,5]

[1]Institute of Medical Psychology and Behavioral Neurobiology, University of Tuebingen, 72074 Tuebingen, Germany, [2]International Max Planck Research School of Neural and Behavioral Sciences, 72074 Tübingen, Germany, [3]Department of Functional Imaging, Center for Diagnostic Radiology and Neuroradiology, University of Greifswald, 17489 Greifswald, Germany, [4]Department of Physiological Psychology, University of Salzburg, 5020 Salzburg, Austria and [5]Ospedale San Camillo, IRCCS, Istituto di Ricovero e Cura a Carattere Scientifico, 30126 Venezia, Italy

Recent neuroimaging studies have indicated a predominant role of the anterior prefrontal cortex (aPFC) in deception and moral cognition, yet the functional contribution of the aPFC to deceptive behavior remains unknown. We hypothesized that modulating the excitability of the aPFC by transcranial direct current stimulation (tDCS) could reveal its functional contribution in generating deceitful responses. Forty-four healthy volunteers participated in a thief role-play in which they were supposed to steal money and then to attend an interrogation with the Guilty Knowledge Test. During the interrogation, participants received cathodal, anodal, or sham tDCS. Remarkably, inhibition of the aPFC by cathodal tDCS did not lead to an impairment of deceptive behavior but rather to a significant improvement. This effect manifested in faster reaction times in telling lies, but not in telling the truth, a decrease in sympathetic skin-conductance response and feelings of guilt while deceiving the interrogator and a significantly higher lying quotient reflecting skillful lying. Increasing the excitability of the aPFC by anodal tDCS did not affect deceptive behavior, confirming the specificity of the stimulation polarity. These findings give causal support to recent correlative data obtained by functional magnetic resonance imaging studies indicating a pivotal role of the aPFC in deception.

Keywords: frontal cortex, lie detection, moral cognition, neuroethics, skin-conductance response (SCR), transcranial direct current stimulation (tDCS)

Introduction

Deception is a complex cognitive act, with crucial legal, moral, and social implications. Functional magnetic resonance imaging (fMRI) studies on neural correlates of deception have shown that the prefrontal cortex and the anterior cingulate cortex (ACC) were more strongly activated during lying than during telling the truth (Lee et al. 2002; Ganis et al. 2003). Recent knowledge about characteristic brain activation sites during deception enabled to recognize false statements with a precision between 88% and 99% (Davatzikos et al. 2005). Ganis et al. (2003) demonstrated that the anterior prefrontal cortices (aPFCs; BA 9/10) were engaged during general deception, but that the right aPFC was more involved in lies that were well rehearsed and were part of a coherent story than in spontaneous, noncoherent lies, whereas the ACC was more active during spontaneous generation of nonmemorized lies. In a recent positron emission tomography (PET) study, Abe et al. (2007) differentiated between the process of generating untruthful responses and the social intention to deceive an interrogator. The main effect of generating untruthful responses revealed increased brain activity of the left dorsolat-

eral prefrontal cortex (DLPFC; BA 8) and the right aPFC, whereas the left ventromedial PFC (BA 11) and Amygdala were associated with the process of deceiving the interrogator. Further analysis revealed that only the right aPFC was associated with both factors of deception, indicating that this region has a pivotal role in telling lies. Although these findings are quite remarkable, these neuroimaging studies have at least 3 shortcomings. First, a general problem of neuroimaging techniques like fMRI or PET is that they allow only correlative statements about the brain regions involved in a specific behavior (here deception). Causal relevance can be demonstrated with other methods allowing transient inhibition of cortical excitability such as transcranial magnetic stimulation (TMS) (Karim et al. 2003; Amedi et al. 2004; Karim, Schuler, et al. 2006; Knoch et al. 2006) or transcranial direct current stimulation (tDCS) (Nitsche and Paulus 2000, 2001; Nitsche, Schauenberg, et al. 2003; Knoch et al. 2008; Priori et al. 2008). Second, the functional contribution of the PFC to deception remains elusive. If, for instance, increased activation of the aPFC reflects cognitive processes involved in withholding the truth, suppression of this region should impair deceptive behavior. However, if increased activation of the aPFC rather reflects a moral conflict involved in deceiving the counterpart, then suppressing this area should have exactly the opposite effect and "improve" deceptive behavior through behavioral disinhibition. Neuroimaging studies on psychopaths, classified as pathological liars, have demonstrated that they have significantly less gray matter in the PFC (Yang et al. 2005) and that they do not show moral dilemma like healthy subjects (Anderson et al. 1999). Thirdly, in previous fMRI studies, participants were instructed when to lie and when to say the truth. However, in cognitive processing, there is a crucial difference between a person who decides himself/herself whether to lie or to say the truth, and a person who merely follows the instruction of the investigator to lie for a predefined time in the fMRI scanner and then to say the truth in order to contrast the 2 conditions.

The aim of this study was therefore 1) to realize an experimental setup, in which participants should decide themselves, which questions they would answer truthfully and which ones with a lie and 2) to investigate the causal contribution of the aPFC in deceptive behavior by modulating the excitability of this brain region through tDCS. Three experiments were conducted to test the specificity of the transcranial stimulation effect.

In the first experiment, 22 healthy subjects participated in a mock crime, in which they were supposed to steal money and

then to attend an interrogation with a modified version of the Guilty Knowledge Test (GKT). In addition to verbal response (truth vs. lie) reaction time (RT), skin-conductance response (SCR) and feelings of guilt while deceiving the interrogator were assessed. In a double-blind repeated-measures design, subjects received cathodal or sham tDCS of their aPFC during the interrogation of the mock crime. Furthermore, in order to measure skillful lying, we developed a ratio called "lying quotient" (LQ) relating the frequency of lies to critical questions with the frequency of lies to uncritical questions. Skillful lying meant that a person intending to appear innocent should not simply lie on all questions, because this behavior would appear rather suspicious. Instead, as in a real criminal interrogation, the suspects had to decide themselves which questions they would answer truthfully and which ones with a lie. Accordingly, a subject achieved a relatively high LQ if he/she answered all "critical items" (whose correct answer only the interrogator and the thief knew, e.g., the true color of the wallet) with a lie, but all "uncritical items" truthfully. To increase motivation for deceptive behavior, participants were told that they were allowed to keep the stolen money in case they could convince the interrogator that they were not guilty.

To test the specificity of the applied stimulation polarity and stimulation site, we conducted a second experiment with 22 healthy volunteers in which the stimulation polarity was reversed. For "anodal" tDCS of the aPFC, the anodal electrode was placed over FP2 (international EEG 10/20 system), and the cathodal electrode was placed over PO3 (left parieto-occipital cortex) as a control area. In randomized order, anodal or sham tDCS of the aPFC was applied during the interrogation

Further 20 healthy subjects participated in a third experiment, in which the Stroop test (Stroop 1935) was used as a "control task." In experiments 1 and 2, subjects intending to deceive the interrogator had to inhibit the truth as a prepotent response and give instead a deceitful answer. The Stroop task is a widely used index of executive control (MacLeod 1991; Swick and Jovanovic 2002) that tests the ability to inhibit a prepotent response but does not include deceiving the counterpart.

Materials and Methods

Subjects

For experiments 1-3, there were 22, 22, and 20 participants, respectively (13, 9, and 10 men). The mean age standard deviation was 25.6 4.9, 24.8 3.9 and 26.0 4.0. Each subject participated in only 1 of the 3 experiments. All subjects were right handed according to the Edinburgh Handedness Inventory (Oldfield 1971). The study was approved by the ethics committee of the Medical Faculty of the University of Tübingen. Subjects were excluded if information from a standardized medical questionnaire suggested prior neurological, psychiatric, or cardiovascular diseases or consumption of centrally acting medication. Parts of these data were previously presented at the 49th Annual Meeting of the Society of Psychophysiological Research in Vancouver, Canada (Karim, Lotze, et al. 2006).

Experimental Design

Experiments 1 and 2 consisted of a thief role-play, in which money (20 Euros) was stolen and a subsequent interrogation, in which the suspects were asked questions about the course of the mock crime according to the GKT paradigm. The GKT (Lykken 1959, 1960) utilizes a series of multiple-choice questions, each having 1 true alternative and several false alternatives, chosen so that an innocent suspect would not be able to discriminate them from the relevant alternative (e.g., "the color of the stolen wallet was: red? black? brown? blue? gray?"). Thus, if the subject's

physiological responses to the relevant alternative are consistently larger than the control alternatives, knowledge about the crime is inferred (for a meta-analysis on the validity of the GKT see Ben-Shakar and Elaad 2003). The role-play was organized as follows: Two subjects were asked to pick 1 of 2 chits of paper from a cup. The subjects were told that on 1 chit was written "thief" and on the other one "innocent attendee." The subjects were asked to memorize their roles but not to tell the instructor which role they had chosen. After the roles were assigned by drawing lots, the subjects were told to go to an office and wait there for 20 min until the interrogation. This office consisted of a main room and an adjoining room. Both rooms were shown to the subjects before assigning the roles, and they were told that the innocent attendee should wait during the mock crime in the main room, while the thief should go to the adjoining room and search there for money with the intention to steal it. Money could be placed at several locations. Therefore, the thief should not only search for the money thoroughly but also as quickly as possible. The subjects were further told that after the money has been stolen, both subjects will be suspected to be the thief. Each of them will attend independently of each other 2 interrogations with an investigator who will play the role of a police inspector. In the interrogation, the subjects will be asked questions, which they should answer as quickly as possible with a "yes" or a "no." Additionally, the SCR and the RT will be recorded. The subjects were also told that during each of the 2 interrogations, they will receive different types of tDCS. The true "thief" should lie in such a skillful manner that the interrogator would believe he/she is innocent. Skillful lying meant that a person intending to appear innocent should not simply lie on all questions, because this behavior would appear rather suspicious. Instead, as in a real criminal interrogation, the suspects had to decide themselves which questions they would answer truthfully and which ones with a lie. To enhance the motivation of the subjects to identify themselves with their role and to make the role-play as realistic as possible, subjects were told that they were allowed to keep the stolen money in case they could convince the interrogator that they were not guilty. However, in reality, 1 of the 2 subjects was a collaborator of the experiment, a fact unknown to the subject and on both pieces of paper "thief" was written, but the collaborator knew that he had to play the role "innocent attendee." The goal of the investigation was to elucidate, if the subjects would show during cathodal transcranial DC stimulation of the aPFC different deceptive behavior than during anodal or sham stimulation.

Transcranial DC Stimulation

TDCS involves continuous administration of weak currents of ~1 mA through a pair of surface electrodes attached to the scalp (Nitsche and Paulus 2000). Previous studies have demonstrated that cerebral excitability was diminished by cathodal stimulation, which hyperpolarizes neurons (Terzuolo and Bullock 1956; Creutzfeldt et al. 1962; Bindmann et al. 1964; Gartside 1968). Bindmann et al. (1964) have shown that cathodal stimulation in animals can reduce or completely inhibit spontaneous firing of cortical cells. In humans, it has been shown that cathodal stimulation can decrease the excitability of the motor (Nitsche and Paulus 2000; Liebetanz et al. 2002; Nitsche, Nitsche, et al. 2003), visual (Antal et al. 2001, 2004) and somatosensory cortex (Dieckhöfer et al. 2006).

In the first experiment, the cathodal electrode was placed over FP2 and the anodal electrode over PO3 according to the international 10-20 EEG system (Fig. 1a). TDCS polarity refers to the right fronto-polar electrode. PO3 was chosen as a reference for 2 reasons: First, to maximize the distance between the cathodal and the anodal electrode, because current density calculations have shown that increasing the distance between the electrodes decreases the current shunted through the scalp and increases the current density in depth (Rockstroh et al. 1989; Miranda et al. 2006) and second, because previous neuroimaging studies did not show that the parieto-occipital cortex (BA 39) is involved in deception (for a review, see Karim et al. 2009). A constant current flow of 1 mA was applied through wet sponge electrodes (4 × 6 cm), and continuous tDCS was delivered by a battery driven, constant current stimulator (Schneider Electronic, Gleichen, Germany) for 13 min. The interrogation started 3 min after onset of the stimulation and lasted for 8-10 min, so that tDCS was applied through the whole interrogation but had 3 min forerun to reach maximum effects (Nitsche and Paulus 2000).
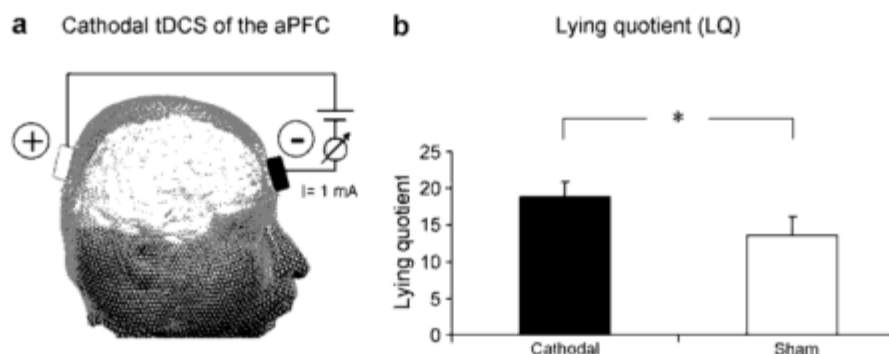
Figure 1. Panel *A* illustrates the technique used for transcranial DC stimulation. Weak direct current (1 mA) was applied between 2 large (24 cm²), wet sponge electrodes placed over FP2 and PO3 according to the international 10–20 EEG system. TDCS polarity refers to the fronto-polar electrode. Panel *B* depicts the effect of cathodal tDCS on skillful lying measured by the LQ. Error bars denote standard error of the mean (SEM). *$P < 0.05$.

The current was always ramped up or down over the first and last 5 s of stimulation, respectively. During tDCS, voltages of more than approximately 10 V can induce a mild tingling sensation in the skin under the scalp electrodes, whereas tDCS at lower voltages is usually not associated with sensory stimulation even in experienced subjects (Hummel et al. 2005). Skin resistance gradually declines after the first few seconds of current application. In consequence, the voltage needed to hold constant current decreases and becomes subthreshold for evoking peripheral sensations. For sham tDCS, placement of the electrodes, current intensity and ramp time was identical to real tDCS; however, the stimulation lasted only for 30 s. The rationale behind this sham procedure was to mimic the transient skin sensation at the beginning of real tDCS without producing any conditioning effects on the brain (Siebner et al. 2004; Hummel et al. 2005). This method of sham stimulation has been shown to be reliable (Gandiga et al. 2006). The interrogator and the subjects were blind to the intervention (tDCS or sham), which was applied by a separate investigator.

In the second experiment, the stimulation polarity was reversed meaning that the anodal electrode was placed over FP2 and the cathodal electrode over PO3 according to the international 10–20 EEG system. Current intensity, ramp time, and duration of stimulation were identical to the first experiment.

In the third experiment, the stimulation parameters and stimulation site were identical to the first experiment. The order of real and sham tDCS was balanced in the 3 experiments.

Measurement of the LQ

In order to measure skillful lying, we developed a ratio called lying quotient (LQ):

$$LQ = \left[\left(\frac{N_{crit}}{N_{tot\_crit}} \right) - \left(\frac{N_{uncrit}}{N_{tot\_uncrit}} \right) \right] \times 100, \quad (1)$$

where N_{crit} = Frequency of lies on critical questions, N_{tot_crit} = Total number of critical questions, N_{uncrit} = Frequency of lies on uncritical questions, and N_{tot_uncrit} = Total number of uncritical questions.

Skillful lying meant that a person intending to appear innocent should not simply lie on all questions, because this behavior would appear rather suspicious. Instead, as in a real criminal interrogation, the suspects had to decide themselves which questions they would answer truthfully and which ones with a lie.

In the interrogation, a modified version of the GKT was applied consisting of 10 critical and 7 uncritical questions, each with 4 choices. An uncritical question was a question, whose answer would be known even by an innocent attendee, who has been in the room but did not steal the money (e.g., "On the chair in the small room there was a jacket. Was the color of the jacket: green? blue? black? brown?"). In contrast, a critical question was a question, whose answer would be known only by the thief (e.g., "In the pocket of the jacket there was wallet. Was the color of the wallet: green? blue? black? brown?").

According to formula (1), the LQ can range from −100 to +100. A most skillful liar would have a maximum LQ of 100, if he/she lies on all critical questions, but answers all uncritical questions truthfully.

Subjects who decide simply to lie on all questions independently of their relevance to the criminal act will have an LQ of 0. A quite odd behavior would be, if a subject answers all critical questions truthfully but lies on all uncritical questions. In such a case, that subject would get an LQ of −100. Besides having a direct measure for skillful lying, an important advantage of the LQ is that it enables us to control for the subjects' bias strategies or predisposition to answer almost all questions in an interrogation with a lie or truthfully independently of the fact, if they are critical or not. A subject who decides to lie on all questions would not admit knowing any critical information, but still would appear dishonest, because he/she denies knowing information, which he/she should know even as an innocent attendee. In contrast to this strategy, another subject might prefer to answer almost all questions truthfully. Such a subject would appear very honest; however, he/she would increase the possibility to be detected as the thief, because he/she would admit knowing a lot of information, which only the delinquent could have known.

Measurement of the RT

RT was defined as the time between the end of the question and the onset of the answer. Note that the relevant information in the question was always in the last word (e.g., the color of the wallet was "green." The color of the wallet was "blue," etc.). Subjects answered the questions verbally with a yes or a no. During the interrogation, the investigator and the subjects were wearing headphones with microphones, and the whole interrogation was recorded with Cool Edit Pro (Syntrillium Software Corp., Phoenix, United States). Acoustic information was digitalized at a 16-bit resolution and a sampling rate of 22 kHz. To determine the acoustic onset of the verbal response, an amplitude filter was used that removed all acoustic signals with an amplitude of less than 7.5% of the maximum sound level. The correctness of detecting the onset of each verbal response was checked off-line by making use of the playback function of the program.

Measurement of SCR

SCRs were recorded at 16-Hz sampling rate with a commercial ambulatory device (Varioport, Becker Meditec, Karlsruhe, Germany) using standard Ag/AgCl electrodes filled with unibase electrolyte affixed to the left hand. Data were processed off-line in a Matlab environment (Matlab 6.5, The Mathworks Inc., Natick, MA). Skin-conductance data were smoothed with a 1 s Gaussian kernel. The amplitude of SCR was determined as the largest change in conductance between 1 and 5 s after task onset, relative to the preceding smallest value in the interval. For statistical analysis, SCRs were log transformed (log(SCR + 1)).

Measurement of the Feelings of Guilt While Deceiving the Interrogator

At the end of each interrogation, the subjects were asked to rate their feelings of guilt that they might have experienced while deceiving the

interrogator on a scale from 0 (no feelings of guilt) to 5 (maximum feelings of guilt).

Stroop Task

To test the possible effect of cathodal tDCS on executive prefrontal function (i.e., the ability to inhibit a prepotent response), participants performed the Stroop task during sham and cathodal tDCS of the aPFC, respectively. The task was conducted with a color-coded 4-button keyboard. Participants were presented with color words printed in colored ink and asked to name the color of the ink as quickly as possible. Color words printed in an incongruent color (i.e., "red" printed in blue ink) produces slower RT known as Stroop interference (Stroop 1935). The task consisted of 66 practice trials to minimize the error rate, followed by 66 experimental trials (33 congruent and 33 incongruent in randomized order). The stimulus words were: "red," "green," "blue," and "yellow." Color names appeared on the screen in 1 of the 4 colors. Preceding each trial, a fixation cross was shown for 2 s. The trial interval was constant with a duration of 2 s. After the participants' response, the screen became black for the rest of the trial interval.

Results

Experiment 1

Interestingly, if only the number of lies was compared between cathodal and sham tDCS, no significant difference was found between the 2 conditions ($t = 1.768$, $P = 0.092$). However, concerning the LQ, subjects achieved in the stimulation condition a significantly higher LQ than in the sham condition ($t = 2.254$, $P = 0.035$), meaning that the answers given in the interrogation during cathodal tDCS were less likely to reveal their guilt, than the answers given during sham stimulation (Fig. 1).

A repeated-measures analyses of variance (ANOVA$_{RM}$) with Stimulation Condition$_{(cathodal\ tDCS/sham\ tDCS)}$ and Response$_{(truth/lie)}$ as within-subject factors and Reaction Time as dependent variable revealed no significant main effects (for Stimulation Condition: $F_{1,21} = 2.198$, $P = 0.153$; for Response: $F_{1,21} = 1.156$, $P = 0.294$) but a significant interaction between the 2 factors ($F_{1,21} = 7.037$, $P = 0.020$; Fig. 2a). Posthoc t tests showed that during sham tDCS, the RT for lying was significantly longer than for truthful responding ($t = 2.568$, $P = 0.018$). However, during cathodal tDCS, the RT was significantly shorter for telling lies ($t = 2.447$, $P = 0.02$) but not for telling the truth ($t = 0.611$, $P = 0.548$). To analyze the effect of cathodal tDCS on sympathetic SCR, an ANOVA$_{RM}$ with Stimulation Condition$_{(cathodal\ tDCS/sham\ tDCS)}$ and Response$_{(truth/lie)}$ as within-subject factors was conducted. Again, no significant main effects were found (for Stimulation Condition: $F_{1,21} = 1.908$, $P = 0.191$; for Response: $F_{1,21} = 3.216$, $P = 0.096$) but a significant interaction between the 2 factors ($F_{1,21} = 6.287$, $P = 0.024$). Posthoc t test revealed that in the sham condition, the SCR for lying was significantly higher than for saying the truth ($t = 3.029$, $P = 0.008$). However, in the stimulation condition, this difference in SCR between lies and truthful responses disappeared ($t = 0.626$, $P = 0.539$; Fig. 2b).

To further investigate the effect of cathodal tDCS of the aPFC, on the subjective experience of guilt, subjects were asked at the end of the interrogation to rate their feelings of guilt, which they might have experienced during the interrogation, on a scale from 0 (no feelings of guilt) to 5 (maximum feelings of guilt). Wilcoxon signed-rank test revealed that cathodal tDCS of the aPFC led to significantly lower feelings of guilt than in the sham condition ($z = -1.986$, $P = 0.047$; Fig. 2c). Moreover, a Kendall's tau correlation
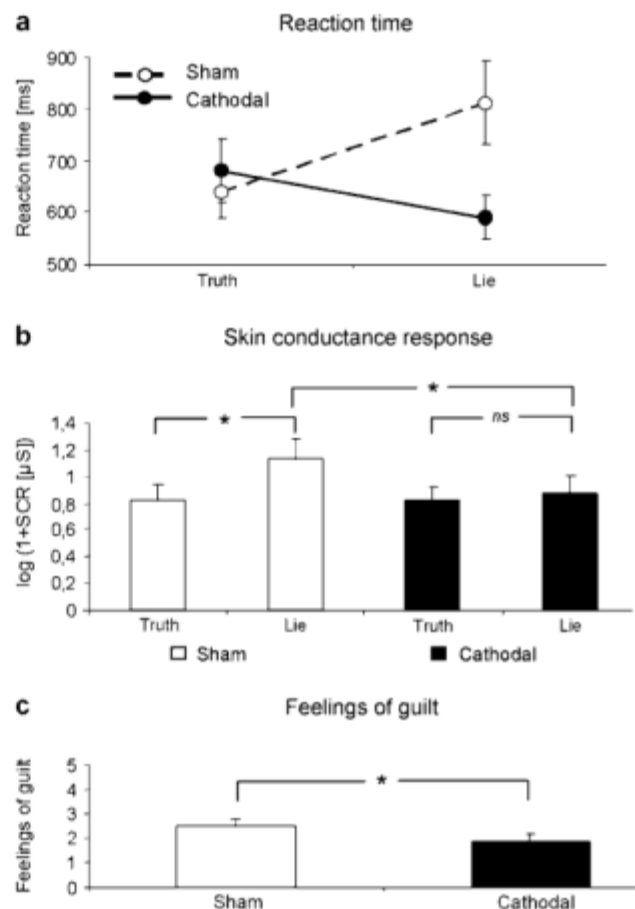
Figure 2. Effects of cathodal transcranial DC stimulation of the aPFC on RT (a), sympathetic SCR (b), and on feelings of guilt (c) when subjects tell lies in an interrogation of a mock crime. Error bars denote SEM. *$P < 0.05$.

analyses revealed a significantly negative correlation between the change of feelings of guilt (cathodal condition minus sham condition) and the change of the LQ ($\tau = -0.386$, $P = 0.023$), indicating that the less feelings of guilt subjects perceived, the better could they deceive during the interrogation.
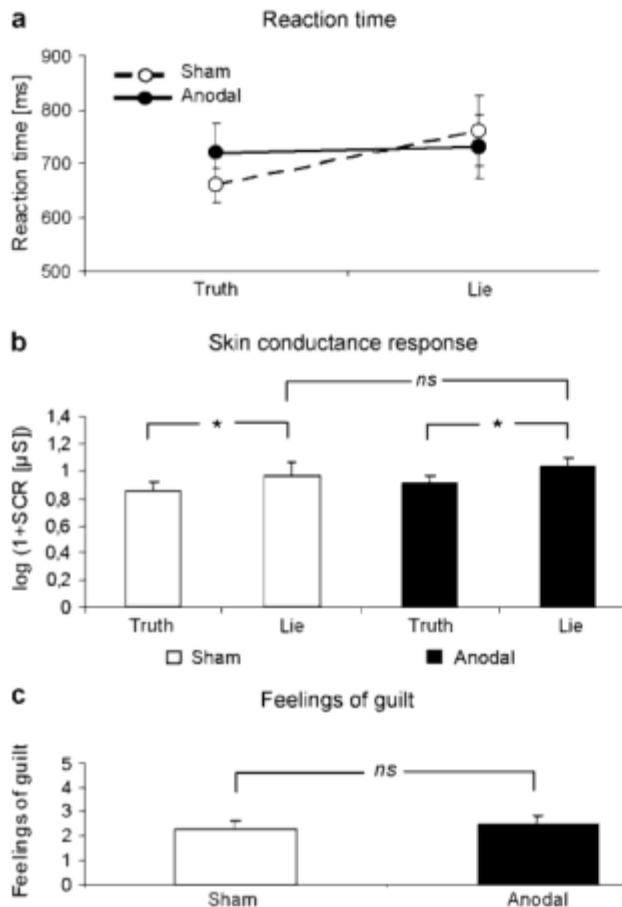
Experiment 2

In order to exclude the possibility that the observed effects were only due to nonspecific effects of the electrical stimulation and not specific to the inhibition of the aPFC by "cathodal" DC stimulation, we conducted a second experiment in which the experimental design was identical to the first experiment but the stimulation polarity was reversed. In contrast to the first experiment, anodal tDCS of the aPFC did not lead to a significant change of the LQ ($t = 0.51$, $P = 0.619$; Fig. 3).

An ANOVA$_{RM}$ with Stimulation Condition$_{(anodal\ tDCS/sham\ tDCS)}$ and Response$_{(truth/lie)}$ as within-subject factors and Reaction Time as dependent variable revealed no significant main effects (for Stimulation Condition: $F_{1,21} = 0.209$, $P = 0.652$; for Response: $F_{1,21} = 2.833$, $P = 0.107$) and no significant interaction between the 2 factors ($F_{1,21} = 2.972$, $P = 0.099$; Fig. 4a).

To analyze the effect of anodal tDCS on sympathetic SCR, a further ANOVA$_{RM}$ with Stimulation Condition$_{(anodal\ tDCS/sham\ tDCS)}$ and Response$_{(truth/lie)}$ as within-subject factors was conducted.

a Anodal tDCS of the aPFC

I = 1 mA

b Lying quotient (LQ)

ns
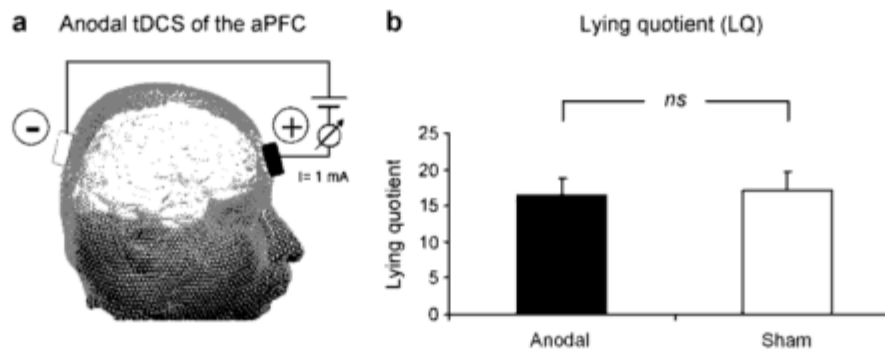
Figure 3. Panel *A* illustrates anodal transcranial DC stimulation of the aPFC. TDCS polarity refers to the fronto-polar electrode. Panel *B* depicts the effect of anodal tDCS on skillful lying measured by the LQ. Error bars denote SEM.

a Reaction time

○ Sham
● Anodal

b Skin conductance response

ns

* *

□ Sham ■ Anodal

c Feelings of guilt

ns

Figure 4. Effects of anodal transcranial DC stimulation of the aPFC on RT (*a*), sympathetic SCR (*b*), and on feelings of guilt (*c*) when subjects tell lies in an interrogation of a mock crime. Error bars denote SEM. *$P < 0.05$.

The Response (lie vs. truth) revealed a significant main effect on SCR ($F_{1,21} = 38.190$, $P < 0.001$); however, the Stimulation Condition (anodal tDCS vs. sham tDCS) had no effect on SCR ($F_{1,21} = 1.164$, $P = 0.298$), and no significant interaction ($F_{1,21} = 0.009$, $P = 0.926$) was found between Stimulation Condition and Response (Fig. 4*b*). Also concerning the feelings of guilt that subjects might have experienced while deceiving the interrogator, in contrast to the first experiment, anodal tDCS did not lead to a significant change of the subjective experience of guilt ($z = -1.89$, $P = 0.059$; Fig. 4*c*).

Experiment 3

We tested a possible impact of cathodal tDCS of the aPFC on general prefrontal executive function by using the Stroop test as a control task. An ANOVA$_{RM}$ with Stimulation Condition$_{(cathodal\ tDCS/sham\ tDCS)}$ and Stroop Condition$_{(congruent/incongruent)}$ as within-subject factors revealed a significant main effect of the Stroop Condition on RT ($F_{1,19} = 46.109$, $P < 0.001$). However, the Stimulation Condition had no effect on RT ($F_{1,19} = 1.050$, $P = 3.18$), and no significant interaction ($F_{1,19} = 1.593$, $P=0.222$) was found between Stimulation Condition and Stroop Condition (see Fig. 5).

Discussion

This study demonstrates for the first time that cathodal transcranial DC stimulation, which has been repeatedly shown to suppress cortical excitability (Nitsche, Nitsche, et al. 2003; Antal et al. 2004; Dieckhöfer et al. 2006) can modulate deceptive behavior. Moreover, our findings give causal support to recent correlative data obtained by neuroimaging studies indicating a predominant role of the aPFC in deceptive behavior (Lee et al. 2002; Ganis et al. 2003; Abe et al. 2007). Whereas in previous studies on neural correlates of deception participants were instructed when to lie and when to say the truth, in the present study, subjects could decide themselves which questions they would answer truthfully and which ones with a lie, taking into account the difference in cognitive processing for cued and uncued lying. Most remarkably, we observed that inhibiting the excitability of the aPFC with cathodal tDCS did not lead to impairment but rather to a significant within-subject improvement of deceptive behavior. This effect was expressed in faster RTs for telling lies, but not for telling the truth, a decrease in sympathetic SCR and feelings of guilt while deceiving the interrogator and a significantly higher LQ, which reflects skillful lying.

In order to exclude the possibility that the observed effects were only due to nonspecific effects of the electrical stimulation and not specific to the inhibition of the aPFC by cathodal DC stimulation, we conducted a control experiment in which the stimulation polarity was reversed. Our data show that shorter RTs in telling lies compared with telling the truth and the absence of increased SCR while deceiving the interrogator were confined to cathodal tDCS of the aPFC and were not detectable during sham tDCS or anodal tDCS. Because subjects were blinded to the stimulation condition and could
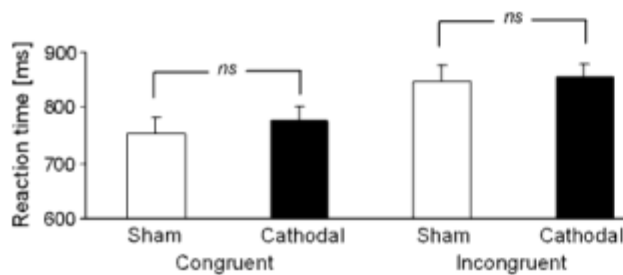
Figure 5. Cathodal transcranial DC stimulation of the aPFC has no effect on RT in the Stroop task. Error bars denote SEM.

not differentiate between the stimulation polarities, nonspecific effects of the stimulation or higher awareness because of stimulation cannot explain the observed effects.

An alternative explanation for the observed effects in experiment 1 can be stated as follows: Cathodal tDCS of the aPFC did not have an effect on deception per se but on cognitively demanding tasks in general. Because telling lies is cognitively more demanding than telling the truth, one might suspect that this is the main reason why an effect was found. Thus, DC stimulation would have affected any other cognitively demanding task in a similar manner. To exclude this possibility, we conducted a third experiment with the Stroop test as a control task. Our results demonstrate that although the incongruent condition is cognitively more demanding than the congruent one, cathodal tDCS of the aPFC had no effect on performance, suggesting a specific effect on deceptive behavior and not on cognitively demanding tasks in general.

The intriguing question that remains is why did cathodal tDCS lead to "improvement" of deceptive behavior and not to its impairment?

Recent neuroimaging studies have emphasized that the aPFC (BA 9/10) plays a crucial role in moral cognition (Greene et al. 2001; Moll et al. 2002, 2005). Moll et al. (2002, 2005) found increased activation of the aPFC when a moral judgment condition was compared with a nonemotional factual judgment, but not when moral judgments were compared with a social emotional condition, during which a more ventral region was activated. Greene et al. (2001) used a moral judgment task that involved classic moral dilemmas (e.g., should you kill an innocent person in order to save 5 other people?) and found increased activation of the aPFC during emotionally loaded moral judgments. Moreover, neuroimaging studies have also emphasized the importance of the aPFC in social interaction (Stuss et al. 2001; Decety and Sommerville 2003; Amodio and Frith 2006; Heatherton et al. 2006; Raine and Yang 2006). Heatherton et al. (2006) have shown that making judgments about the self relative to an intimate other selectively activates the aPFC. Stuss et al. (2001) have demonstrated on patients with limited focal frontal and nonfrontal lesions that the frontal lobes are necessary for "theory of mind," which includes inferences about feelings of others and empathy for those feelings. The anterior PFC, the ventral PFC, and the amygdala are regions that have been shown to be involved in both antisocial behavior and moral decision making (Raine and Yang 2006). Taking these findings into account, the aPFC seems to be crucially involved in socioemotional judgments. Suppressing the excitability of this region or focal lesions should therefore show an impact on

antisocial and moral behavior. In respect to our study, deceiving another person in order to obtain personal profit seems to create a moral conflict, and if a person is relieved from this moral conflict, he/she might be able to deceive unhinderedly with faster RT, less feelings of guilt and less sympathetic arousal as demonstrated here. Suppressing cortical excitability by cathodal tDCS or low-frequency repetitive transcranial magnetic stimulation (rTMS) has previously been shown to induce so-called paradoxical improvement of performance through "disinhibition" processes (Hilgetag et al. 2001; Kobayashi et al. 2004; Fecteau et al. 2007). Kobayashi et al. (2004) have, for example, demonstrated that suppression of the primary motor cortex by low-frequency rTMS enhances motor performance with the ipsilateral hand by releasing the contralateral motor cortex from transcallosal inhibition. Using tDCS, Fecteau et al. (2007) have recently shown that enhancing DLPFC activity diminished risk-taking behavior, but only when coupled with inhibitory modulation over the contralateral DLPFC. Intriguingly, Koenigs et al. (2007) have also shown that a lesion of the PFC leads to an increase of utilitarian moral decisions. An increase in antisocial behavior following PFC impairment is supposed to result from a release of limbic areas from PFC executive control (Moll et al. 2005). However, it is not the aim of this study to state that the aPFC is the only cortical region, whose stimulation can modulate deceptive behavior. Neuroimaging studies have indicated that also other cortical areas, especially the DLPFC (Phan et al. 2005; Abe et al. 2006, 2007) and the superior temporal sulcus (Phan et al. 2005) are also involved in deception and that in different types of deception (e.g., lies that are rehearsed and part of a coherent story vs. spontaneous noncoherent lies) different cortical networks are involved (Ganis et al. 2003; Abe et al. 2007). Priori et al. (2008) have recently demonstrated that tDCS of the DLPFC alters RT in deception of experienced events but had no effect on RTs in deception of new events. Thus, future studies will have to investigate the effect of stimulation of different cortical areas in different types of lies and the duration of these effects in relation to the stimulation parameters.

A further interesting question is, why anodal tDCS, which has been shown to increase cortical excitability (Gartside 1968; Nitsche and Paulus 2001; Antal et al. 2004), did not lead to opposite effects compared with cathodal tDCS resulting in an impairment of deceptive behavior and an increase of feelings of guilt while deceiving the interrogator? Although our data show that concerning the LQ and feelings of guilt there is a tendency toward lower LQ and higher feelings of guilt during anodal tDCS compared with sham tDCS (cf. Figs 3b and 4c), these changes did not reach significance. It is plausible to assume that disruption of the PFC can have an effect on social cognition (Anderson et al. 1999), moral reasoning (Koenigs et al. 2007), or even on deception as shown in the present study, however, increasing the excitability in a "normal functioning" PFC does not necessarily have to lead to opposite effects presumably due to ceiling effects. However it is tempting to test in patients with "impaired" PFC if increasing cortical excitability by anodal tDCS can help to remedy functional deficits.

In transcranial stimulation studies, positioning the TMS coil or the tDCS electrodes can provide a great challenge. Although in tDCS studies positioning the relatively large electrodes (about 4 × 6 cm) according to the international 10–20 EEG system is a very common method (s. Knoch et al. 2006; Fecteau et al. 2007; Priori et al. 2008), Herwig et al. (2003) have shown

that for TMS studies, positioning the more focal figure-of-eight TMS coil according to the 10–20 EEG system is reliable when dealing with larger scale cortical areas. Thus, for stimulating a relatively large and well-defined cortical region as the aPFC stereotaxic neuronavigation systems are certainly not necessary. In a PET study, Lang et al. (2005) have placed the tDCS electrodes over the primary motor cortex (identified by inducing motor evoked potentials with TMS) and over the right fronto-polar cortex (directly above the right eyebrow) and found the highest increase in regional cerebral blood flow below the stimulating electrodes in the primary motor cortex and the aPFC. Moreover, Okamoto et al. (2004) established recently for transcranial stimulation studies a correspondence between the 10–20 EEG system and magnetic resonance imaging based stereotaxic space (Talairach coordinates and the standard template of the Montreal Neurological Institute) and expressed the anatomical structures for the 10–20 cortical projection points probabilistically. Their findings show that despite interindividual variance in the structure of the prefrontal cortex, the electrode position over FP2 is with a 100% probability in BA 10. Taking these findings into account, positioning the tDCS electrode over FP2 stimulates mainly BA 10. However, due to the use of relatively large electrodes (4 × 6 cm) to prevent heating artifacts, stimulation of the junction to BA9 has to be considered as well.

Nitsche and Paulus (2000) have shown that a minimum current density of 0.017 mA/cm^2 is necessary to modify cortical excitability by tDCS in humans. The applied current density of 0.04 mA/cm^2 in this study is in accordance with several tDCS studies demonstrating functionally relevant modulating effects on cortical excitability (cf. Hummel et al. 2005; Nitsche et al. 2007). One might further suspect that the 3D pattern of brain sulci and gyri might create an overall change in current polarity in the targeted brain areas. However, current density calculations from our laboratory (Rockstroh et al. 1989) and from other research groups (Rush and Driscoll 1968; Miranda et al. 2006) as well as direct intracellular measurements of DC stimulation (Purpura and McMurtry 1965) revealed an average current flow in the expected direction independent of single sulci and gyri.

The findings of the present study are also particularly interesting in the light of clinical evidence suggesting that psychopaths, who are classified as pathological liars, have significantly less gray matter in their PFC (Yang et al. 2005) and, remarkably, do not show higher SCR when telling lies (Verschuere et al. 2005). We have previously demonstrated that in psychopaths limbic-prefrontal regions (amygdala, orbitofrontal cortex, insula, and the anterior cingulate), and SCR during anticipation of aversive events is pathologically reduced (Veit et al. 2002; Birbaumer et al. 2005). In a social reactive aggression paradigm, Lotze et al. (2007) have shown that during retaliation, subjects with high psychopathic scores had less BA 9/10 activation in comparison to subjects with low psychopathic scores. These findings are in accordance with the results of other research groups reporting decreased prefrontal blood flow (for a review, see Blair 2007) and deficient autonomic responses, for example, SCR, in anticipation of threatening events (Blair et al. 1997; Hare et al. 1978). Moreover, several studies (Anderson et al. 1999; Moll et al. 2005) have also shown that in psychopaths and patients with aPFC lesions, moral cognition is impaired. Thus, our findings support the hypotheses that a dysfunction of the aPFC and its

specific connections may underlie certain psychopathological conditions that are characterized by the absence of sympathetic arousal while performing a wrongful act such as deceiving in a criminal interrogation.

Finally, concerning the current debate on emerging ethical issues in neuroscience (cf. Farah 2002), interdisciplinary research and communication are needed to address the following question: If neuroscientific research can demonstrate that deceptive behavior and moral cognition are not only associated with the activation of specific brain areas, but may even be modulated by noninvasive stimulation of these areas, what implications will such findings have on our concept of personal responsibility and neuroethical applications?

References

Abe N, Suzuki M, Mori E, Itoh M, Fujii T. 2007. Deceiving others: distinct neural responses of the prefrontal cortex and amygdala in simple fabrication and deception with social interactions. J Cogn Neurosci. 19:287–295.

Abe N, Suzuki M, Tsukiura T, Mori E, Yamaguchi K, Itoh M, Fujii T. 2006. Dissociable roles of prefrontal and anterior cingulate cortices in deception. Cereb Cortex. 16:192–199.

Amedi A, Floel A, Knecht S, EZ, Cohen LG. 2004. Transcranial magnetic stimulation of the occipital pole interferes with verbal processing in blind subjects. Nat Neurosci. 7:1266–1270.

Amodio DA, Frith CD. 2006. Meeting of minds: the medial frontal cortex and social cognition. Nat Rev Neurosci. 7:268–277.

Anderson SW, Bechara A, Damasio H, Tranel D, Damasio AR. 1999. Impairment of social and moral behaviour related to early damage in human prefrontal cortex. Nat Neurosci. 2:1032–1037.

Antal A, Kincses TZ, Nitsche MA, Bartfai O, Paulus W. 2004. Excitability changes induced in the human primary visual cortex by transcranial direct current stimulation: direct electrophysiological evidence. Invest Ophthalmol Vis Sci. 45:702–707.

Antal A, Nitsche MA, Paulus W. 2001. External modulation of visual perception in humans. NeuroReport. 12:3553–3555.

Ben-Shakar G, Elaad E. 2003. The validity of psychophysiological detection of information with the guilty knowledge test: a meta-analytic review. J Appl Psychol. 88:131–151.

Bindmann LJ, Lippold OC, Redfearn JW. 1964. The action of brief polarizing currents on the cerebral cortex of the rat (1) during current flow and (2) in the production of long-lasting after effects. J Physiol. 172:369–382.

Birbaumer N, Veit R, Lotze M, Erb M, Hermann C, Grodd W, Flor H. 2005. Deficient fear conditioning in psychopathy: a functional magnetic resonance imaging study. Arch Gen Psychiatry. 62:799–805.

Blair RJR. 2007. The amygdala and ventromedial prefrontal cortex in morality and psychopathy. Trends Cogn Sci. 11:387–392.

Blair RJR, Jones L, Clark F, Smith M. 1997. The psychopathic individual: a lack of responsiveness to distress cues? Psychophysiology. 34:192–198.

Creutzfeldt OD, Fromm GH, Kapp H. 1962. Influence of transcortical d-c currents on cortical neuronal activity. Exp Neurol. 5:436–452.

Davatzikos C, Ruparel K, Fan Y, Shen DG, Acharyya M, Loughead JW, Gur RC, Langleben DD. 2005. Classifying spatial patterns of brain activity with machine learning methods: application to lie detection. NeuroImage. 28:663–668.

Decety J, Sommerville JA. 2003. Shared representations between self and other: a social cognitive neuroscience view. Trends Cogn Sci. 7:527–533.

Dieckhöfer A, Waberski TD, Nitsche M, Paulus W, Buchner H, Gobbele R. 2006. Transcranial direct current stimulation applied over the somatosensory cortex—differential effect on low and high frequency SEPs. Clin Neurophysiol. 117:2221–2227.

Farah MJ. 2002. Emerging ethical issues in neuroscience. Nat Neurosci. 5:1123–1129.

Fecteau S, Knoch D, Fregni F, Sultani N, Boggio P, Pascual-Leone A. 2007. Diminishing risk-taking behavior by modulating activity in the prefrontal cortex: a direct current stimulation study. J Neurosci. 27:6212–6218.

Gandiga PC, Hummel FC, Cohen LG. 2006. Transcranial DC stimulation (tDCS): a tool for double-blind sham-controlled clinical studies in brain stimulation. Clin Neurophysiol. 117:845–850.

Ganis G, Kosslyn SM, Stose S, Thompson WL, Yurgelun-Todd DA. 2003. Neural correlates of different types of deception: an fMRI investigation. Cereb Cortex. 13:830–836.

Gartside IB. 1968. Mechanisms of sustained increases of firing rate of neurones in the rat cerebral cortex after polarization: role of protein synthesis. Nature. 220:383–384.

Greene JD, Sommerville RB, Nystrom LE, Darley JM, Cohen JD. 2001. An fMRI investigation of emotional engagement in moral judgment. Science. 293:2105–2108.

Hare RD, Frazelle J, Cos DN. 1978. Psychopathy and physiological responses to threat of an aversive stimulus. Psychophysiology. 15:165–172.

Heatherton T, Wyland CL, Macrae CN, Demos KE, Denny BT, Kelley WM. 2006. Medial prefrontal activity differentiates self from close others. SCAN. 1:18–25.

Herwig U, Satrapi P, Schoenfeldt-Lecuona C. 2003. Using the 10-20 EEG system for positioning of transcranial magnetic stimulation. Brain Topogr. 16:95–99.

Hilgetag CC, Théoret H, Pascual-Leone A. 2001. Enhanced visual spatial attention ipsilateral to rTMS-induced 'virtual lesions' of human parietal cortex. Nat Neurosci. 4:953–957.

Hummel F, Celnik P, Giraux P, Floel A, Wu WH, Gerloff C, Cohen LG. 2005. Effects of non-invasive cortical stimulation on skilled motor function in chronic stroke. Brain. 128:490–499.

Karim AA, Kammer T, Lotze M, Hinterberger T, Godde B, Cohen L, Birbaumer N. 2003. Effects of repetitive transcranial magnetic stimulation (rTMS) on slow cortical potentials (SCP). Suppl Clin Neurophysiol. 56:331–337.

Karim AA, Lotze M, Schneider M, Weber C, Braun C, Birbaumer N. 2006. Inhibition of the anterior prefrontal cortex improves deceptive behavior. Psychophysiology. 43:S50.

Karim AA, Schneider M, Krippl M, Birbaumer N. 2009. Neurobiology of deception. In: Müller JL, editor. Neurobiology of forensic disorders. Stuttgart: Hogrefe. Forthcoming.

Karim AA, Schuler A, Hegner YL, Friedel E, Godde B. 2006. Facilitating effect of 15-Hz repetitive transcranial magnetic stimulation on tactile perceptual learning. J Cogn Neurosci. 18:1577–1585.

Knoch D, Nitsche MA, Fischbacher U, Eisenegger C, Pascual-Leone A, Fehr E. 2008. Studying the neurobiology of social interaction with transcranial direct current stimulation—the example of punishing unfairness. Cereb Cortex. 18:1987–1990.

Knoch D, Pascual-Leone A, Meyer K, Treyer V, Fehr E. 2006. Diminishing reciprocal fairness by disrupting the right prefrontal cortex. Science. 314:829–832.

Kobayashi M, Hutchinson S, Théoret H, Schlaug G, Pascual-Leone A. 2004. Repetitive TMS of the motor cortex improves ipsilateral sequential simple finger movements. Neurology. 62:91–98.

Koenigs M, Young L, Adolphs R, Tranel D, Cushman F, Hauser M, Damasio A. 2007. Damage to the prefrontal cortex increases utilitarian moral judgements. Nature. 446:908–911.

Lang N, Siebner HR, Ward NS, Lee L, Nitsche MA, Paulus W, Rothwell JC, Lemon RN, Frackowiak RS. 2005. How does transcranial DC stimulation of the primary motor cortex alter regional neuronal activity in the human brain? Eur J Neurosci. 22:495–504.

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

Liebetanz D, Nitsche MA, Tergau F, Paulus W. 2002. Pharmacological approach to the mechanisms of transcranial DC-stimulation-induced after-effects of human motor cortex excitability. Brain. 125:2238–2247.

Lotze M, Veit R, Anders S, Birbaumer N. 2007. Evidence for a different role of the ventral and dorsal medial prefrontal cortex for social reactive aggression: an interactive fMRI study. Neuroimage. 34:470–478.

Lykken DT. 1959. The GSR in the detection of guilt. J Appl Psychol. 43:385–388.

Lykken DT. 1960. The validity of the guilty knowledge technique: the effects of faking. J Appl Psychol. 44:258–262.

MacLeod CM. 1991. Half a century of research on the Stroop effect: an integrative review. Psychol Bull. 109:163–203.

Miranda PD, Lomarev M, Hallett M. 2006. Modeling the current distribution during transcranial direct current stimulation. Clin Neurophysiol. 117:1623–1629.

Moll J, de Oliveira-Souza R, Eslinger PJ, Bramati IE, Mourao-Miranda J, Andreiuolo PA, Pessoa L. 2002. The neural correlates of moral sensitivity: a functional magnetic resonance imaging investigation of basic and moral emotions. J Neurosci. 22:2730–2736.

Moll J, Zahn R, de Oliveira-Souza R, Krueger F, Grafman J. 2005. The neural basis of human moral cognition. Nat Neurosci. 6:799–809.

Nitsche MA, Doemkes T, Karaköse A, Antal A, Liebetanz D, Lang N, Tergau F, Paulus W. 2007. Shaping the effects of transcranial direct current stimulation of the human cortex. J Neurophysiol. 97:3109–3117.

Nitsche MA, Nitsche MS, Klein CC, Tergau F, Rothwell JC, Paulus W. 2003. Level of action of cathodal DC polarisation induced inhibition of the human motor cortex. Clin Neurophysiol. 114:600–604.

Nitsche MA, Paulus W. 2000. Excitability changes induced in the human motor cortex by weak transcranial direct current stimulation. J Physiol. 527:633–639.

Nitsche MA, Paulus W. 2001. Sustained excitability elevations induced by transcranial DC motor cortex stimulation in humans. Neurology. 57:1899–1901.

Nitsche MA, Schauenburg A, Lang N, Liebetanz D, Exner C, Paulus W, Tergau F. 2003. Fascilitation of implicit motor learning by weak transcranial direct current stimulation of the primary motor cortex in the human. J Cogn Neurosci. 15:619–626.

Okamoto M, Dan H, Sakamoto K, Kazuhiro T, Shimizu K, Kohono S, Oda I, Isobe S, Suzuki T, Kohayama K, et al. 2004. Three-dimensional probabilistic anatomical carnio-cerebral correlation via the international 10-20 system oriented for transcranial functional brain mapping. Neuroimage. 21:99–111.

Oldfield RC. 1971. The assessment and analysis of handedness: the Edinburgh Inventory. Neuropsychologia. 9:97–113.

Phan KL, Magalhaes A, Ziemlewicz TJ, Fitzgerald DA, Green C, Smith W. 2005. Neural correlates of telling lies: a functional magnetic resonance imaging study at 4 Tesla. Acad Radiol. 12:164–172.

Priori A, Mameli 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.

Purpura DP, McMurtry JG. 1965. Intracellular activities and evoked potential changes during polarization of motor cortex. J Neurophysiol. 28:166–185.

Raine A, Yang Y. 2006. Neural foundations to moral reasoning and antisocial behavior. SCAN. 1:203–213.

Rockstroh B, Elbert T, Canavan A, Lutzenberger W, Birbaumer N. 1989. Slow cortical potentials and behaviour. Baltimore (MD): Urban & Schwarzenberg.

Rush S, Driscoll DA. 1968. Current distribution in the brain from surface electrodes. Anesth Analg. 47:717–723.

Siebner HR, Lang N, Rizzo V, Nitsche MA, Paulus W, Lemon RN, Rothwell JC. 2004. Preconditioning of low-frequency repetitive transcranial magnetic stimulation with transcranial direct current stimulation: evidence for homeostatic plasticity in the human motor cortex. J Neurosci. 24:3379-3385.

Stroop JR. 1935. Studies of interference in serial verbal reactions. J Exp Psychol. 18:643-662.

Stuss DT, Gallup GG, Jr., Alexander MP. 2001. The frontal lobes are necessary for 'theory of mind'. Brain. 124:279-286.

Swick D, Jovanovic J. 2002. Anterior cingulate cortex and the Stroop task: neuropsychological evidence for topographic specificity. Neuropsychologia. 40:1240-1253.

Terzuolo CA, Bullock TH. 1956. Measurement of imposed voltage gradient adequate to modulate neuronal firing. Proc Natl Acad Sci USA. 42:687-694.

Veit R, Flor H, Erb M, Hermann C, Lotze M, Grodd W, Birbaumer N. 2002. Brain circuits involved in emotional learning in antisocial behavior and social phobia in humans. Neurosci Lett. 328: 233-236.

Verschuere B, Crombez G, De Clercq A, Koster EH. 2005. Psychopathic traits and autonomic responding to concealed information in a prison sample. Psychophysiology. 42:239-245.

Yang Y, Raine A, Lencz T, Bihrle S, Lacasse L, Colletti P. 2005. Prefrontal white matter in pathological liars. Br J Psychiatry. 187:320-325.