Cerebral Cortex September 2008;18:1987–1990
doi:10.1093/cercor/bhm237
Advance Access publication December 24, 2007

FEATURE ARTICLE

Studying the Neurobiology of Social Interaction with Transcranial Direct Current Stimulation—The Example of Punishing Unfairness

Daria Knoch[1,2,3], Michael A. Nitsche[4], Urs Fischbacher[1], Christoph Eisenegger[1], Alvaro Pascual-Leone[5] and Ernst Fehr[1,3]

[1]Institute for Empirical Research in Economics, University of Zürich, 8006 Zürich, Switzerland, [2]Department of Neurology, University Hospital Zürich, 8092 Zürich, Switzerland, [3]Collegium Helveticum, 8092 Zürich, Switzerland, [4]Department Clinical Neurophysiology, Georg-August-University, 37099 Goettingen, Germany and [5]Berenson-Allen Center for Noninvasive Brain Stimulation, Department of Neurology, Beth Israel Deaconess Medical Center, Boston, MA 02215, USA

The first 2 authors equally contributed to this work.

Studying social behavior often requires the simultaneous interaction of many subjects. As yet, however, no painless, noninvasive brain stimulation tool existed that allowed the simultaneous affection of brain processes in many interacting subjects. Here we show that transcranial direct current stimulation (tDCS) can overcome these limits. We apply right prefrontal cathodal tDCS and show that subjects' propensity to punish unfair behavior is reduced significantly.

Keywords: brain stimulation technique, prefrontal cortex, social interaction, transcranial direct current stimulation

Introduction

Social neuroscience (Adolphs 2003) and neuroeconomics (Glimcher and Rustichini 2004) examine the neural mechanisms of complex social behaviors such as trusting other people (Delgado, Frank, et al. 2005; Delgado, Miller, et al. 2005; King-Casas et al. 2005; Kosfeld et al. 2005), contributing to public goods, participating in market exchanges (Knutson et al. 2007), or the altruistic punishment of defectors in social exchanges (de Quervain et al. 2004). Currently, most work in this area is based on neuroimaging tools that allow for the examination of the neural correlates of social behaviors. These tools, although indispensable, do not permit causal inferences about the effect of brain processes on human behavior. In contrast, brain stimulation techniques such as transcranial magnetic stimulation (TMS) interfere with the activity of defined areas in the human cortex noninvasively, thus enabling researchers to observe the behavioral impact of an increase or decrease in the cortical excitability of the stimulated brain region. However, both brain imaging tools and TMS are difficult to apply "simultaneously" to larger groups of, say, 5 or more interacting subjects. Although hyperscanning is possible in theory (King-Casas et al. 2005), scanning more than 3 or 4 people simultaneously is at present beyond the scope of most or all laboratories. Likewise, we are not aware of any study that applied TMS simultaneously to a group of people who interact with each other.

Many social interaction experiments, however, require the simultaneous interaction of a number of subjects. For example, it is often necessary for larger groups of 5, 10, or even 20 people to interact simultaneously with each other in public goods or market experiments. Likewise, it is important in experiments examining altruistic behaviors that subjects interact only once with many different partners. As the absence of interaction partners during the experiment may raise suspicion among the subjects and may, therefore, change their behaviors (Frohlich et al. 2001), the best credible implementation of social interactions is achieved through the simultaneous presence of all subjects during the experiment. For example, if a subject doubts that a real person defected in a social exchange but instead suspects that the experimenter merely fabricated the defection, the subject has little reason to spend resources on punishing a possibly nonexistent defector.

Currently, no brain stimulation method is available that enables the neuroscientific study of "simultaneous" social interactions in larger experimental groups. In the present study, we applied transcranial direct current stimulation (tDCS) to investigate whether this method is a feasible tool for examining how individual brain processes may affect the social decision making process. tDCS induces changes in cortical excitability by means of a weak electrical field applied transcranially, which de- or hyperpolarizes neuronal membranes on a subthreshold level. Anodal tDCS increases, whereas cathodal tDCS decreases, excitability (Nitsche and Paulus 2001). tDCS does thus not directly elicit action potentials by means of suprathreshold resting membrane potential change but renders neuronal populations more or less ready to fire in response to additional inputs. In other words, it changes the likelihood that an incoming action potential will result in postsynaptic firing (Bindman et al. 1964, Purpura and McMurtry 1965, Wagner et al. 2007). It was demonstrated that the neurophysiological and functional effects of tDCS are generally restricted to the area under the electrodes (Nitsche et al. 2003, 2007). In a visuomotor coordination experiment, tDCS of both polarities influenced performance, where excitability-enhancing tDCS improved performance in the learning phase, probably due to improving long-term potentiation-like plasticity, whereas excitability-diminishing cathodal tDCS improved performance in the overlearned state of the task by increasing the signal-to-noise ratio, as suggested by a control experiment (Antal et al. 2004a, b). Taken together, these studies suggest that the effects of tDCS are spatially and functionally specific, not only on the neurophysiological but also on the behavioral level. Recent research on prefrontal cognitive functions has shown that prefrontal tDCS can modulate working memory and probabilistic classification learning, depending on stimulation polarity (Kincses et al. 2004; Fregni et al. 2005). However, no study has yet demonstrated that tDCS can affect behavior in social interactions.

We applied tDCS to the prefrontal cortices of a group of subjects acting in the role of a responder in an ultimatum bargaining game. A "proposer" in this game is paired with a "responder," and the former can propose how to split an available amount of money. The responder can then either accept or reject the offer. In case of a rejection, both players earn nothing; in case of acceptance, the amount of money is split as proposed. Strong evidence suggests that many people reject low offers in the ultimatum game (UG) because they view them as unfair (Camerer 2003). Thus, a responder who is tendered an unfair offer faces a tension between economic self-interest, which suggests accepting even a low offer, and fairness motives, which favor rejecting low offers. We deliberately chose the UG because we know from a recent study (Knoch et al. 2006) that disrupting the function of the right dorsolateral prefrontal cortex (DLPFC) by means of low-frequency rTMS increases the acceptance rate of unfair offers relative to a placebo stimulation, whereas rTMS of the left DLPFC does not affect behavior significantly. This gives us a clear hypothesis for the present study because both cathodal tDCS and low-frequency rTMS tend to reduce cortical excitability of the stimulated brain region. The question then is whether—within the constraints of painless, noninvasive, stimulation—tDCS is sufficiently powerful to override the strong fairness motives that drive rejections in the UG.

Materials and Methods

Participants (mean age 23, range 21–26 years) neither suffered from neurological nor psychiatric conditions, nor did they take chronic or acute medications. All were naive to tDCS, the UG, and the nature of the experiment; they were further unaware of the experimental variable tested. Participants gave informed written consent prior to entering the study; the local ethics committee approved this study. Six responders where stimulated simultaneously in 1 session.

We conducted the tDCS study with 128 subjects in the role of the proposer (no tDCS) and 64 subjects (all right-handed men) in the role of the responder who received either cathodal (i.e., excitability reducing) tDCS ($n = 30$) or placebo tDCS ($n = 34$) to the right DLPFC. During an experimental session (Fig. 1), 6 responders played 12 UGs each with 12 different, anonymous, proposers, that is, each responder "faced" any given proposer only once and was never informed of the bargaining partner's identity. We deliberately chose 1-shot interactions because no strategic spillovers across periods occur with this structure. This is particularly important if "true" preferences are to be elicited. We can therefore rule out the possibility that the observed rTMS effect is due to induced ignorance of possible reputational effects. The responders had to agree on the division of 20 Swiss Francs (CHF; CHF 1 ≈ € 0.65). The proposer could make 1 and only 1 proposal how to allocate the CHF 20 by making an offer of CHF 4, 6, 8, or 10 to the responder. If the responder accepted, each player received the amount the proposer suggested. If the responder rejected, neither player received any money. Subjects received instructions prior to tDCS stimulation that explained the rules of the game. Each subject was required to complete a series of test questions successfully after reading the instructions to verify comprehension. We also ensured that the responders faced the same distribution of offers across the 12 UGs regardless of the treatment condition. This design feature ensures that behavioral differences across treatments cannot be caused by different offer distributions. After the responders had played 12 UGs, they were shown a list of all possible offers and asked to report on a 7-point scale to what extent they perceived an offer as fair or unfair (1 = very unfair; 7 = very fair). These fairness assessments took place during tDCS stimulation. We implemented the random payment method in our experiment, that is, 6 of the 12 trials were randomly selected for payment. The experiment was conducted with the z-tree software (Fischbacher 2007).

Figure 1. Experimental setting: 12 proposers and 6 responders were in the same laboratory during an experimental session. Responders received tDCS (placebo or cathodal stimulation over the right DLPFC), whereas experimenters sat between each pair of responders to control the tDCS devices.

Direct current was induced using 2 saline-soaked surface sponge electrodes (active right cathodal electrode area = 35 cm^2, reference electrode area = 100 cm^2) and delivered by a battery-driven, constant current stimulator. For technical details of the stimulator, contact P.S. Boggio (boggio@usp.br). To allow a functionally unipolar tDCS, we used a large reference, which has been demonstrated to be functionally inert without diminishing the efficacy of tDCS under the stimulation electrode (Nitsche et al. 2007). For stimulation over the right DLPFC, the cathode electrode was placed over F4 (electroencephalography 10/20 system) and the reference electrode over the left orbit. Participants received a constant current of 1.5 mA intensity with cathodal polarity over the stimulation electrode for active stimulation. tDCS started 4 min before the task began and was delivered during the whole course of the UG, which lasted less than 10 min. For sham stimulation, the electrodes were placed at the same positions as for active stimulation, but the stimulator was only turned on for 30 s; participants thus felt the initial itching sensation associated with tDCS but received no active current for the rest of the stimulation period. This method of sham stimulation has been shown to be reliable (Gandiga et al. 2006).

Results

As expected, the acceptance rates varied strongly across offers. Offers of 4 were accepted on average in 35.4% of the trials, whereas the acceptance rate for offers of 6 was 75.5%, offers of 8 were accepted in 96.8% and offers of 10 in 100% of the cases. If we examine the behavior of the 2 treatment groups separately (Fig. 2*a*), however, we observe treatment differences in the acceptance rate of unfair offers. During placebo tDCS, the acceptance rate for the most unfair offer was 25.4%, considerably less than the acceptance rate of 46.6% during active cathodal tDCS of the right DLPFC. The differences across the 2 groups are significant for the most unfair offer (Mann–Whitney *U* test, $Z = -2.244$, $P = 0.025$). The same results hold true if we conduct a repeated-measures analysis of variance of treatment (right DLPFC, placebo) × offer (4, 6, 8, 10). We find a main effect of treatment ($F = 4.17$, $P = 0.046$) and no significant interaction between treatment × offer.

Can changes in subjects' fairness judgments explain the observed effects, or does tDCS prevent the behavioral implementation of these judgments? We elicited subjects' fairness judgments with regard to different offers on a 7-point scale after the behavioral experiment and found that subjects in the 2 treatment groups showed no differences with respect to fairness judgments for any of the offers (Fig. 2*b*). Thus, cathodal tDCS of right DLPFC induces subjects to behave more in line with their economic self-interest by increasing the acceptance rate of unfair offers, although it does not affect fairness judgments.

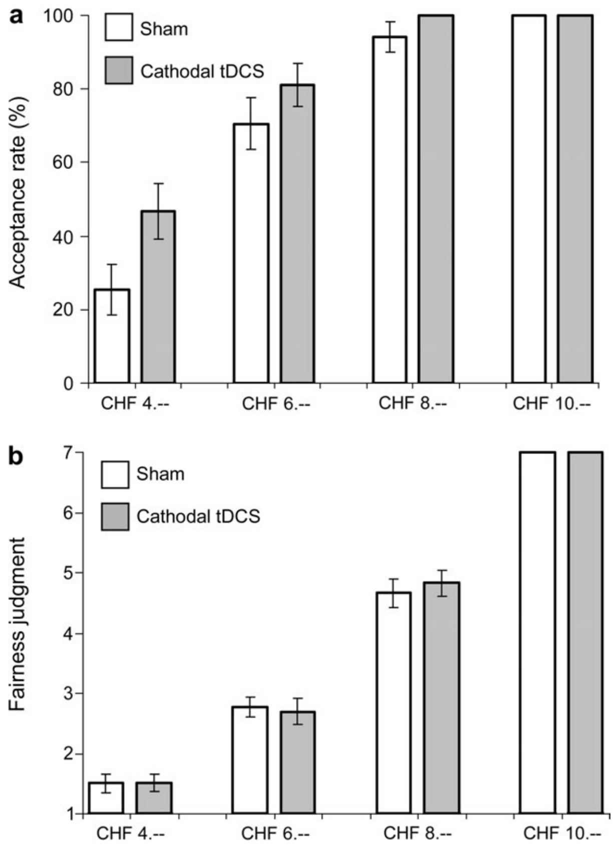

Figure 2. Responders' behavioral responses to all offers and fairness judgments. (*a*) Acceptance rates (means ± standard error of the mean) across the 2 treatment groups. Subjects whose function of the right DLPFC is disrupted by cathodal tDCS exhibit a much higher acceptance rate than those who received placebo tDCS. (*b*) Perceived unfairness across treatments (1 = very unfair; 7 = very fair). There were no group differences in fairness judgments.

Discussion

We find that tDCS is a powerful tool for examining how individual brain processes may affect the outcomes of social interactions. For reasons of clear predictions regarding the expected effect, we chose the UG and applied tDCS to only 1 interaction partner (the responder). The increased acceptance rate during active cathodal tDCS of the right DLPFC could be interpreted as a reduced ability to resist the economic temptation to accept unfair offers. Our findings are also interesting in the light of evidence, suggesting that patients with right prefrontal lesions are characterized by the inability to behave in normatively appropriate ways despite the fact that they possess the social knowledge that is necessary for normative behavior (Damasio 1995). Note that if we suggest that the right DLPFC is involved in overriding self-interest motives, we do not necessarily imply that this brain region directly suppresses other brain areas that represent self-interest. Instead, we believe that the right DLPFC is involved in top-down control (or executive control), the overall effect of which is a reduction in the weight of self-interested impulses on an individual's action. Thus, rather than directly suppressing neural activities that represent self-interested impulses, the DLPFC may be part of a network that modulates the relative impact of fairness motives and self-interest goals on decision making, and the final outcome of this modulation may therefore be a weakening of the impact of self-interest motives on decision making. Another possible interpretation is that prefrontal cathodal tDCS disrupted and/or disturbed the negative emotional reaction to unfair offers. The right DLPFC seems to be involved in regulating emotional responses in general. For example, Ochsner et al. (2004) have shown that both the downregulation and upregulation of negative emotions activate the DLPFC/anterior cingulated cortex network. Disrupting the right DLPFC function may have prevented the prefrontal cortex from being able to react to emotional feeling states. Indeed, some studies favor an effect of prefrontal tDCS on affectivity (Lippold and Redfearn 1964). However, the tDCS paradigm used in these studies differs relevantly from that applied here. In those studies, the frontopolar, and not the dorsolateral prefrontal cortex, was stimulated, and the reference electrode was positioned at the knee. This might have caused a relevant brainstem stimulation, which might explain the results, as the authors suggest. In a recently conducted study, which applied the same tDCS protocol as used in the present experiment, tDCS did not influence mood in healthy subjects (unpublished results of our group).

Regardless of which interpretation ultimately prevails, the primary aim of this study was to show the great potential of tDCS—a technique that seems to be gaining in popularity (Hallett 2007)—for brain stimulation studies investigating social interaction.

One important advantage of tDCS over rTMS becomes apparent in the context of studying the impact of tDCS on fairness judgments. Prefrontal rTMS is associated with potential side effects (Robertson et al. 2003; Abler et al. 2005), including discomfort and irritation, which means that a behavioral effect could be due to these side effects rather than the reduction (or increase) in neuronal excitability. A potential solution to this problem is offline rTMS, where the experimental task is performed "after" stimulating the brain. In particular, low-frequency rTMS for the duration of several minutes leads to a suppression of activity in the stimulated brain region that outlasts the duration of the rTMS train for several minutes (Robertson et al. 2003). The duration and the strength of this aftereffect are subject to some uncertainty, however. Thus, it is not entirely clear to what extent subjects' fairness judgments were still made under sufficiently reduced neuronal excitability under offline rTMS, which was concluded before the start of the UG (Knoch et al. 2006). In contrast, active tDCS is painless and virtually unnoticeable; behavioral tasks can therefore be performed "during" tDCS. This has the great advantage that the duration and the strength of the aftereffect are of no consequence. Thus, we can be absolutely sure that subjects' fairness judgments took place under the effect of cathodal, excitability-reducing tDCS. The relatively large tDCS electrodes used in this experiment, however, complicate speculation on which specific brain areas are influenced. In particular, a restriction of the stimulation efficacy to a specific area is improbable. However, the main aim of this study was to test the principal efficacy of tDCS in influencing task performance. Because the spatial resolution of rTMS and functional magnetic resonance imaging is also restricted, we decided to use these large electrodes to minimize the probability of missing the relevant area.

To summarize, tDCS has the distinct advantage that it can be centrally administered simultaneously to many interacting subjects, it is noninvasive and painless (Gandiga et al. 2006), it provides a reliable sham condition (Gandiga et al. 2006), and—as we have shown here—it can nevertheless change important social behaviors. In addition, it is inexpensive and easy to apply. These properties make tDCS a powerful tool for studying the neuronal mechanisms of social interaction, enabling researchers to establish exciting links between neuronal events in individual brains, individual behaviors, and the outcomes of complex social interactions.

Notes

This paper is part of the research priority program at the University of Zürich on the "Foundations of Human Social Behavior—Altruism versus Egoism." *Conflict of Interest*: None declared.

Address correspondence to email: dknoch@iew.uzh.ch.

References

Abler B, Walter H, Wunderlich A, Grothe J, Schonfeldt-Lecuona C, Spitzer M, Herwig U. 2005. Side effects of transcranial magnetic stimulation—Biased task performance in a cognitive neuroscience study. Brain Topogr. 17:193–196.

Adolphs R. 2003. Cognitive neuroscience of human social behaviour. Nat Rev Neurosci. 4:65–178.

Antal A, Nitsche MA, Kincses TZ, Kruse W, Hoffmann KP, Paulus W. 2004a. Facilitation of visuo-motor learning by transcranial direct current stimulation of the motor and extrastriate visual areas in humans. Eur J Neurosci. 19:2888–2892.

Antal A, Nitsche MA, Kruse W, Kincses TZ, Hoffmann KP, Paulus W. 2004b. Direct current stimulation over V5 enhances visuomotor coordination by improving motion perception in humans. J Cogn Neurosci. 16:521–527.

Bindman LJ, Lippold OCJ, Redfearn JWT. 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.

Camerer CF. 2003. Behavioral game theory—Experiments in strategic interaction. Princeton (NJ): Princeton University Press.

Damasio AR. 1995. Descartes' error: Emotion, reason and the human brain. New York: Harper Collins.

de Quervain DJF, Fischbacher U, Treyer V, Schelthammer M, Schnyder U, Buck A, Fehr E. 2004. The neural basis of altruistic punishment. Science. 305:1254–1258.

Delgado MR, Frank RH, Phelps EA. 2005. Perceptions of moral character modulate the neural systems of reward during the trust game. Nat Neurosci. 8:1611–1618.

Delgado MR, Miller MM, Inati S, Phelps EA. 2005. An fMRI study of reward-related probability learning. Neuroimage. 24:862–873.

Fischbacher U. 2007. z-Tree: Zürich toolbox for ready-made economic experiments. Exp Econ. 10:171–178.

Fregni F, Boggio PS, Nitsche M, Bermpohl F, Antal A, Feredoes E, Marcolin MA, Rigonatti SP, Silva MT, Paulus W, et al. 2005. Anodal transcranial direct current stimulation of prefrontal cortex enhances working memory. Exp Brain Res. 166:23–30.

Frohlich N, Oppenheimer J, Moore JB. 2001. Some doubts about measuring self-interest using dictator experiments: The costs of anonymity. J Econ Behav Organ. 46:271–290.

Gandiga P, Hummel F, Cohen L. 2006. Transcranial DC stimulation (tDCS): A tool for double-blind sham-controlled clinical studies in brain stimulation. Clin Neurophysiol. 117:845–850.

Glimcher PW, Rustichini A. 2004. Neuroeconomics: The consilience of brain and decision. Science. 306:447–452.

Hallett M. 2007. Transcranial magnetic stimulation: A primer. Neuron. 55:187–199.

Kincses TZ, Antal A, Nitsche MA, Bartfai O, Paulus W. 2004. Facilitation of probabilistic classification learning by transcranial direct current stimulation of the prefrontal cortex in the human. Neuropsychologia. 42:113–117.

King-Casas B, Tomlin D, Anen C, Camerer CF, Quartz SR, Montague PR. 2005. Getting to know you: Reputation and trust in a two-person economic exchange. Science. 308:78–83.

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.

Knutson B, Rick S, Wimmer GE, Prelec D, Loewenstein G. 2007. Neural predictors of purchases. Neuron. 53:147–156.

Kosfeld M, Heinrichs M, Zak P, Fehr E. 2005. Oxytocin increases trust in humans. Nature. 435:673–676.

Lippold OC, Redfearn JW. 1964. Mental changes resulting from the passage of small direct currents through the human brain. Br J Psychiatry. 110:768–772.

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

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. Facilitation of implicit motor learning by weak transcranial direct current stimulation of the primary motor cortex in the human. J Cogn Neurosci. 15:619–626.

Ochsner KN, Ray RD, Cooper JC, Robertson ER, Chopra S, Gabrieli JDE, Gross JJ. 2004. For better or for worse: Neural systems supporting the cognitive down- and up-regulation of negative emotion. Neuroimage. 23:483–499.

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

Robertson EM, Theoret H, Pascual-Leone A. 2003. Studies in cognition: The problems solved and created by transcranial magnetic stimulation. J Cogn Neurosci. 15:948–960.

Wagner T, Valero-Cabre A, Pascual-Leone A. 2007. Noninvasive brain stimulation. Annu Rev Biomed Eng. 9:527–565.