Increasing honesty in humans with noninvasive brain stimulation

Michel André Maréchal[a,1,2], Alain Cohn[b,1,2], Giuseppe Ugazio[c], and Christian C. Ruff[a,2]

[a]Department of Economics, University of Zurich, 8006 Zurich, Switzerland; [b]Booth School of Business, University of Chicago, Chicago, IL 60637; and [c]Department of Psychology, Harvard University, Cambridge, MA 02138

Edited by Susan T. Fiske, Princeton University, Princeton, NJ, and approved March 13, 2017 (received for review September 6, 2016)

Honesty plays a key role in social and economic interactions and is crucial for societal functioning. However, breaches of honesty are pervasive and cause significant societal and economic problems that can affect entire nations. Despite its importance, remarkably little is known about the neurobiological mechanisms supporting honest behavior. We demonstrate that honesty can be increased in humans with transcranial direct current stimulation (tDCS) over the right dorsolateral prefrontal cortex. Participants (n = 145) completed a die-rolling task where they could misreport their outcomes to increase their earnings, thereby pitting honest behavior against personal financial gain. Cheating was substantial in a control condition but decreased dramatically when neural excitability was enhanced with tDCS. This increase in honesty could not be explained by changes in material self-interest or moral beliefs and was dissociated from participants' impulsivity, willingness to take risks, and mood. A follow-up experiment (n = 156) showed that tDCS only reduced cheating when dishonest behavior benefited the participants themselves rather than another person, suggesting that the stimulated neural process specifically resolves conflicts between honesty and material self-interest. Our results demonstrate that honesty can be strengthened by noninvasive interventions and concur with theories proposing that the human brain has evolved mechanisms dedicated to control complex social behaviors.

honesty | cheating | social decision making | brain stimulation | dorsolateral prefrontal cortex

Dishonest behavior is pervasive and carries important economic and societal consequences (1–6). For example, illegal tax evasion is thought to account for over 5% of the world's gross domestic product (7), and total bribes to public officials are estimated at over US $1 trillion annually (8). Furthermore, recent business scandals such as the Volkswagen emission fabrications and several interest rate manipulations in the financial industry have eroded trust in the integrity of the corporate world (5). Although there are formal laws and regulations to limit cheating, honest behavior is often difficult or impossible to monitor and enforce. Individuals therefore face an internal trade-off between honesty and personal material gain.

The conflict between honesty and material self-interest is a central feature of human social life, and honesty is exalted in virtually all world religions and moral value systems. Despite substantial interest in the origins and determinants of honest behavior in biology (9), behavioral sciences (2, 10), and economics (4, 11), little is known about the neural processes that enable humans to resolve conflicts between honesty and personal financial gain. Understanding the neural processes involved in these "costly" displays of honesty could offer important new perspectives on the evolutionary origins and development of honest behavior (9, 12) and may also aid in designing interventions for enhancing lie detection (13) and the treatment of pathological cheating (14).

The neural basis of honesty remains largely unexplored in humans because previous studies have almost exclusively relied on instructed-lying paradigms, which examine deception ability rather than dishonest behavior (15). Participants in these studies are explicitly instructed by an experimenter to make untruthful statements, and they do not benefit materially from lying. Thus, participants neither genuinely decide to be honest nor face a trade-off between honest behavior and material gain. Other studies have used signaling games to study the neural basis of deception (16, 17), but such tasks potentially confound honesty with strategic motives [e.g., if senders believe opponents will do the opposite of what they recommend, then a sender will actually "deceive" the opponent by telling the truth (18)]. Only one neuroimaging study has investigated cheating in a setting that involved a moral trade-off between honesty and financial gain (19). In that study, honest behavior correlated with brain activity in a network comprising areas of the right dorsolateral prefrontal cortex (rDLPFC). However, these correlational findings cannot determine whether heightened neural activity genuinely causes honest behavior or simply reflects a functionally irrelevant by-product of honest behavior.

Here, we present causal evidence for a neural mechanism that regulates honest behavior by applying transcranial direct current stimulation (tDCS) in 145 subjects confronted with a real trade-off between honesty and personal material gain. tDCS is a noninvasive method to modulate neural excitability in healthy humans by applying weak electric currents to the scalp (20) (a detailed description of the experimental procedures is available in *SI Appendix*). To enhance neural excitability exogenously, we applied anodal tDCS (n = 49) over the rDLPFC region

Significance

Honesty affects almost every aspect of social and economic life. We conducted experiments in which participants could earn considerable amounts of money by cheating on a die-rolling task. Cheating was substantial but decreased by more than one-half during transcranial direct current stimulation over the right dorsolateral prefrontal cortex. This stimulation-induced increase in honesty was functionally specific: It did not affect other types of behavioral control related to self-interest, risk-taking, and impulsivity. Moreover, cheating was only reduced when it benefited the participants themselves rather than another person. Thus, the human brain implements specialized processes that enable us to remain honest when faced with opportunities to cheat for personal material gain. Importantly, these processes can be strengthened by external interventions.

Author contributions: M.A.M., A.C., and C.C.R. designed research; G.U. performed research; M.A.M., A.C., and C.C.R. analyzed data; and M.A.M., A.C., G.U., and C.C.R. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]M.A.M. and A.C. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: christian.ruff@econ.uzh.ch, alain.cohn@chicagobooth.edu, or michel.marechal@econ.uzh.ch.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1614912114/-/DCSupplemental.

previously found to be active when subjects decided to behave honestly (figure S1*B* of ref. 19) (*SI Appendix*, Fig. S1). By applying excitatory tDCS over this specific region, we aimed to strengthen behaviorally relevant neural activity and thereby enhance honest behavior. We also measured behavior in two additional groups where we applied tDCS either to decrease neural excitability [cathodal (n = 49)] or to leave it unchanged [sham (n = 47)]. Random assignment to conditions generated three groups that were matched in socioeconomic status, cognitive ability, and personality. Moreover, the three stimulation conditions were conducted in a double-blind manner, were perceived similarly by the participants, and did not differ reliably in terms of participants' mood, alertness, and calmness. Thus, any effects of the three tDCS interventions on honest behavior cannot be explained in terms of preexisting group differences or changes in beliefs and emotions (*SI Appendix*).

During stimulation, we measured cheating using an incentivized and unobtrusive die-rolling task (10, 11) that has been shown to predict rule-violating behavior reliably in real-world settings (21, 22). The die-rolling task was embedded in a battery of control tasks (*SI Appendix*) that served two purposes: First, they allowed us to measure and control for other aspects of choice behavior that may be affected by the stimulation (23–27). Second, they helped to disguise the purpose of the experiment by reducing the salience of the die-rolling task. In the die-rolling task, subjects were instructed to report the outcomes of 10 die rolls using a computer interface. Each roll could result with 50% probability in either a gain of 9 Swiss francs or no change in payoff. Before each roll, a computer screen indicated which outcomes would yield the monetary payoff. Given that the participants could earn up to 90 Swiss francs (about US $90 at the time of testing) in this task, they faced a substantial material incentive to overreport the number of successful die rolls. To ensure fully private decisions, participants completed the task anonymously outside of the experimenter's and other participants' view. Although this prevented us from identifying individual acts of cheating with certainty, we can determine the degree of cheating associated with each tDCS intervention by comparing the mean percentage of reported successful die rolls against the 50% benchmark implied by fully honest reporting.

Cheating was substantial in the neurally ineffective sham condition (Fig. 1*A*). Compared with the honesty benchmark of 50%, participants reported 68% successful die rolls on average (95% confidence interval: 63–74%). This result implies that cheating occurred in 37% of all responses, assuming that participants never misreported to their disadvantage (*SI Appendix*). Additional simulations and Bayesian analyses demonstrate that given our sample size and observations, the probability that the subjects reported completely honestly is virtually zero (*SI Appendix*, Fig. S3). Fig. 1*B* shows the binomial distribution of successful die rolls expected if everyone behaved honestly and the empirically observed distribution for sham tDCS. A total of 8.5% of the subjects reported successful outcomes for all 10 die rolls (thereby maximizing their earnings), which is significantly higher than the 0.1% expected under the binomial distribution ($P < 0.001$, binomial test). Subjects who claimed nine, eight, and seven successful die rolls were also significantly overrepresented ($P < 0.001$, $P = 0.001$, and $P = 0.002$; binomial tests), suggesting that many of them cheated on some but not all possible occasions. Such incomplete cheating is commonly observed in similar paradigms (10, 11).

To test whether enhanced neural excitability promotes honest behavior, we compared the distribution of die rolls in anodal and sham tDCS. Anodal stimulation over the rDLPFC substantially reduced the average percentage of successful die rolls to 58% ($P = 0.005$, rank-sum test; Fig. 1*A*). [Note that the effect of anodal tDCS on honest behavior is also statistically significant if we adjust the P value for multiple hypothesis testing across the control tasks using a conservative Bonferroni correction ($P^* = 0.049$).] This result corresponds to an implied cheating rate of 15%, a figure that is nearly 60% lower than that observed in the sham condition. In *SI Appendix*, we report complementary simulations and Bayesian analyses, showing that the observed differences between the die roll successes for anodal and sham tDCS are very unlikely to have been generated by chance. For example, our simulations show that only five of 10,000 experiments with a similarly sized sample of completely honest subjects would show a similar or larger tDCS effect than the effect we observed.

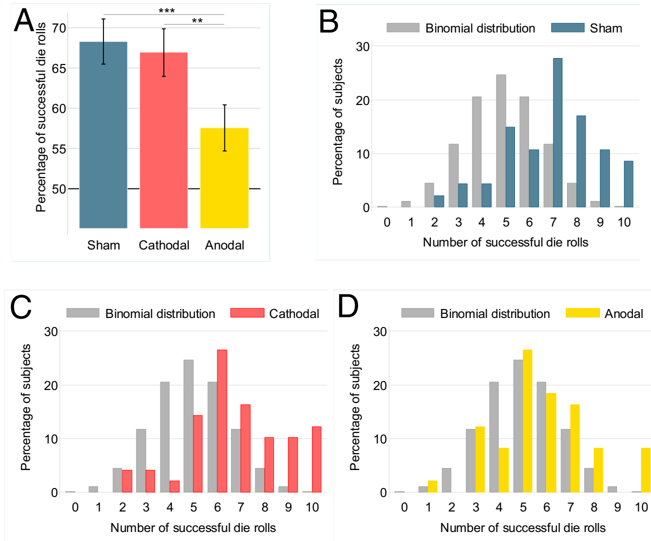

Fig. 1. Effects of tDCS on reporting in the die-rolling task. (*A*) Error bars indicate ±1 SEM. Asterisks indicate significance levels: **$P < 0.05$, ***$P < 0.01$. The self-reported percentage of successful die rolls is significantly reduced for anodal (*D*) compared with sham (*B*) and cathodal (*C*) tDCS ($P = 0.005$ and $P = 0.018$, rank-sum tests; n = 96 and n = 98). The empirical distribution for the cathodal and sham groups is skewed toward higher numbers of successful die rolls compared with the binomial (honest) distribution. The distribution for the anodal group more closely resembles the binomial distribution.

In the analysis of the different reported outcomes, we no longer found significant overreporting of nine, eight, and seven successful die rolls in the anodal condition ($P = 1.000$, $P = 0.168$, and $P = 0.369$; binomial tests; Fig. 1*D*). However, the fraction of subjects who reported the maximum outcome of 10 successful rolls remained essentially unchanged at 8.2%. This finding suggests that anodal tDCS predominantly reduced cheating in participants who actually pondered the trade-off between honesty and financial gain, but not in participants who were committed to maximizing their payoff.

Although anodal tDCS decreased cheating, we did not find opposite behavioral effects of cathodal tDCS (Fig. 1 *A* and *C*). Participants in the cathodal condition reported 67% successful die rolls (95% confidence interval: 61–73%), which was not significantly different from the success rate reported in the sham stimulation ($P = 0.635$, rank-sum test) but was significantly higher than the rate reported in anodal tDCS ($P = 0.018$, rank-sum test). There are two plausible explanations for why cathodal tDCS did not increase cheating. First, several studies suggest that cathodal tDCS induces less stable cognitive behavioral effects than anodal tDCS (28). Second, the high cheating rate in the sham condition [which is similar, for example, to the cheating rate in a sample of prisoners (21)] entails that there was little room for tDCS to increase incomplete cheating further. Thus, cathodal stimulation may not induce transitions from incomplete

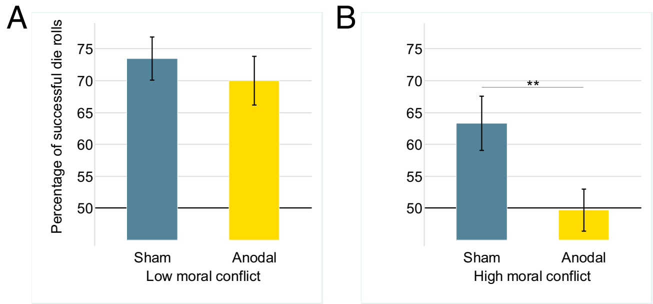

Fig. 2. Effect of anodal tDCS for subjects who experience low and high moral conflict. (*A*) For subjects who assign a low moral value to honesty (below the median of the group, which corresponds to the point of indifference on the Likert scale), the self-reported percentage of successful die rolls does not significantly differ between anodal tDCS and sham groups ($P = 0.327$, rank-sum test; $n = 42$). (*B*) For subjects who assign a high moral value to honesty (above or equal to the median), the difference in successful die rolls between sham and anodal tDCS groups is statistically significant ($P = 0.014$, rank-sum test; $n = 54$). Error bars indicate ±1 SEM. Asterisks indicate significance levels: $^{**}P < 0.05$.

cheating to the more extreme form of cheating in every possible instance. In any case, the results of the cathodal condition clearly show that any general side effects of electrical stimulation, such as possible discomfort or distraction, cannot account for the substantial reduction in cheating observed when stimulation polarity was reversed (20).

We explored possible mechanisms for why anodal tDCS increased honesty. Our task was designed so that participants had to trade off personal financial gain against the value they assigned to being honest. The stimulated neural process could therefore have strengthened honesty by (*i*) decreasing material self-interest (i.e., the subjective value of money), (*ii*) enhancing the value placed on honesty, or (*iii*) perturbing the choice process that trades off these two conflicting motives. We tested these mechanisms using control tasks that were administered to our participants while they were under the influence of the stimulation (*SI Appendix*).

To assess whether anodal tDCS reduced cheating by weakening material self-interest, we used a dictator game that required participants to split money between themselves and well-known charities. Several studies have documented that participants who behave selfishly in dictator games cheat more in other tasks (12). This finding is also evident in our data: Selfish behavior in the dictator game (i.e., the amount of money kept) was positively correlated with subjects' earnings from the die-rolling task (Spearman's rho = 0.266, $P = 0.001$). However, tDCS did not affect the amount of money kept in the dictator game ($P = 0.989$, Kruskal–Wallis test; Table 1), and controlling for dictator game behavior in a regression analysis did not change the effect of anodal tDCS on honest reporting (*SI Appendix*). This result suggests that the increase in honest behavior caused by anodal tDCS is not due to decreased material self-interest.

To test whether anodal tDCS inhibited cheating by increasing the value placed on honesty, we analyzed participants' moral beliefs under the influence of tDCS. Participants indicated the extent to which they considered misreporting in the die-rolling task to be morally inappropriate. This measure was negatively correlated with report rates in the die-rolling task (Spearman's rho = −0.448, $P < 0.001$), confirming that participants who highly valued honesty cheated less. However, tDCS did not affect this measure of moral values ($P = 0.507$, Kruskal–Wallis test). We also did not find that tDCS influenced participants' beliefs about the appropriateness of various forms of dishonest behavior in everyday-life situations ($P = 0.948$, Kruskal–Wallis test). Moreover, controlling for participants' ratings on these measures in a regression analysis did not alter the effect of anodal tDCS on cheating (*SI Appendix*). Thus, the reduction in cheating caused by anodal tDCS does not appear to reflect increased moral valuations of honesty.

We next examined whether tDCS stimulation is involved in resolving the trade-off between honesty and material self-gain. If tDCS stimulation were involved, then anodal tDCS should have primarily influenced individuals who were genuinely conflicted between honesty and material gain. As reported earlier, anodal tDCS indeed reduced incomplete cheating but did not alter the rate of complete cheating (Fig. 1 *B* and *D*). The latter is presumably associated with low conflict due to the complete dominance of financial over moral concerns. We corroborated this result by further examining the magnitude of the tDCS effect in participants who reported low or high moral conflict associated with cheating (Fig. 2). To this end, we used the median rating (corresponding to the point of indifference on the Likert scale) of how "morally inappropriate" participants considered cheating to be in the die-rolling task, to divide subjects into a low-conflict group and a high-conflict group. For low-conflict participants ($n = 42$), cheating was not affected by anodal tDCS ($P = 0.327$, rank-sum test). In contrast, high-conflict subjects ($n = 54$) cheated significantly less in the anodal tDCS group than in the sham group ($P = 0.014$, rank-sum test). Remarkably, responses for high-conflict subjects who received anodal tDCS were not statistically different from the 50% honesty benchmark ($P = 0.920$, *t* test; $n = 30$). These findings substantiate that tDCS only affected the trade-off between honesty and material self-interest for participants who were, in fact, conflicted between these two motives.

In light of these findings, the question emerges of whether the stimulated neural process is specialized for resolving conflicts between material self-interest and honesty, or whether it reflects a general-purpose mechanism involved in any choice between conflicting response options (23). To answer this question, we examined how tDCS affected behavior in three control tasks that required choices between monetary payoffs associated with different levels of risk, ambiguity, and temporal delay, respectively. The stimulation did not affect choices in any of these tasks (columns 4–6 in Table 1 and *SI Appendix*). Moreover, controlling for participants' behavior in these tasks in a regression analysis did not alter the effect of anodal tDCS on cheating (*SI Appendix*). Thus, the neural mechanism affected by tDCS does not

Table 1. Effect of tDCS on other types of behavioral conflicts

Stimulation group	Statistics	Self-interest	Risk	Ambiguity	Impulsivity
Sham ($n = 47$)	Mean	77.433	40.099	29.220	4.716
	± SEM	3.800	3.438	3.060	0.507
Cathodal ($n = 49$)	Mean	79.207	44.639	39.864	4.571
	± SEM	3.295	4.235	3.960	0.553
Anodal ($n = 49$)	Mean	81.027	38.422	36.435	4.857
	± SEM	2.626	3.275	3.462	0.509
Kruskal–Wallis	*P* value	0.989	0.626	0.139	0.826

Self-interest measures the average percentage of the endowment subjects kept for themselves in the three dictator games. Risk (Ambiguity) is the percentage of the endowment invested into a lottery with known (unknown) outcome probabilities. Impulsivity is the average number of impatient choices (0–20) made in three delay discounting tasks. Details are provided in *Materials and Methods*. The Kruskal–Wallis test results in the last row show that tDCS did not have any significant influence on any of these measures of conflict resolution, demonstrating that the stimulated neural process specifically resolves conflicts between honesty and material self-interest.

generally appear to affect choices involving financial trade-offs but, rather, specifically resolves conflicts between material self-interest and the motivation to behave honestly.

A final question we address is whether anodal tDCS over the rDLPFC also reduces cheating when the beneficiary is another person rather than oneself. Testing for such a tDCS effect on prosocial cheating is crucial, because it establishes whether the affected neural mechanism is specific to the conflict between honesty and material self-interest rather than controlling cheating in general (regardless of whether the outcomes benefit oneself or others). This test also addresses potential concerns that anodal tDCS may reduce cheating by biasing participants to opt for a response strategy that is less effortful and complex, because reporting the true or default outcome may be easier than generating false responses to earn money (15). In our design, self-interested cheating and prosocial cheating are matched for cognitive complexity because both require participants to generate fake responses. To test these accounts, we conducted an additional tDCS experiment with 156 participants [anodal (n = 78) and sham (n = 78)] for which we modified the die-rolling task so that subjects could not earn any money for themselves; instead, their earnings were credited to another anonymous participant. All other aspects of the experimental design and procedure were identical to the previous experiment (*SI Appendix*).

In line with previous findings (12, 29), participants undergoing sham tDCS cheated even when the associated gains were assigned to an anonymous recipient (Fig. 3 *A* and *B*). On average, they reported 61% successful outcomes (95% confidence interval: 56–66%), which corresponds to a cheating rate of 22%. A substantial fraction of subjects therefore cheated for purely prosocial reasons, even though the level of cheating was somewhat less pronounced (P = 0.044, rank-sum test) than in the main experiment where cheating served participants' own interest. However, as illustrated in Fig. 3 *A* and *C*, anodal tDCS did not reduce prosocial cheating: 62% of the die rolls were reported as successful, which does not significantly differ from the responses under sham tDCS (P = 0.805, rank-sum test). Moreover, the negative effect of anodal tDCS on dishonest reporting was significantly stronger in the main experiment than in the prosocial die-rolling task (P = 0.017, Wald test). The robustness of this difference between the experiments was again clearly confirmed with simulations and Bayesian analyses, which show that the difference in tDCS effects between the two experiments is highly unlikely to result from random fluctuations in the die roll outcomes (*SI Appendix*). These results show that the tDCS-induced enhancement of honesty in the main experiment cannot be explained by differences in cognitive effort associated with cheating, and they indicate that rDPLFC activity is specifically involved in the resolution of conflicts between honesty and self-interest, rather than affecting all forms of dishonest behavior.

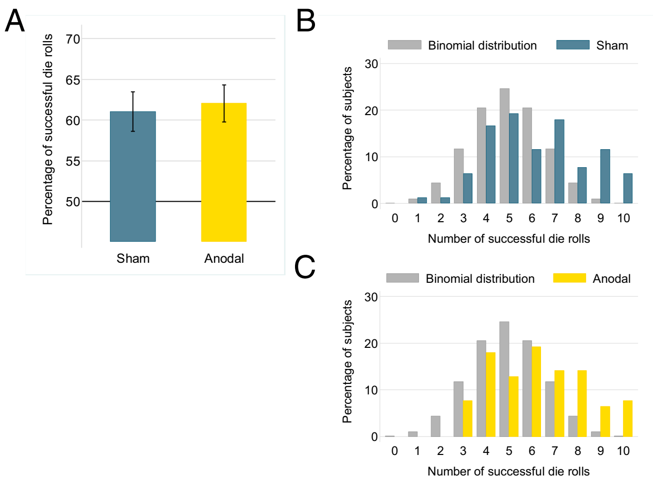

Fig. 3. Effect of anodal tDCS on reporting in the prosocial die-rolling task. (*A*) Error bars indicate ±1 SEM. The self-reported percentage of successful die rolls does not significantly differ between the anodal tDCS and sham groups (P = 0.805, rank-sum test; n = 156). The distributions for the anodal tDCS (*C*) and sham (*B*) groups are similar, and both are skewed toward higher numbers of successful die rolls compared with the binomial (honest) distribution.

Our results demonstrate that neural mechanisms involving the rDLPFC play a causal role in modulating honesty when individuals stand to gain from dishonest behavior. These neural processes are functionally independent of other forms of behavioral trade-offs, such as behavioral trade-offs related to risk (24), ambiguity (25), or delayed rewards (26, 27). Such specialization suggests a dedicated neurobiological process that enables humans to resist the self-interested temptation to cheat, consistent with proposals that complex social structures in primate groups have led to the evolution of neural processes dedicated to the control of social behavior (30). This finding also concurs with recent evidence from twin studies suggesting that moral beliefs about dishonesty are partially inherited (31).

Although the neural process enhanced by tDCS was clearly functionally specific, it seems unlikely that it operates independently. Previous studies suggest that the targeted DLPFC area may interact with other brain areas related to emotions, such as the amygdala (32), and brain regions devoted to behavioral control, such as the anterior cingulate and dorsomedial prefrontal cortex (19). The idea that honest behavior may be controlled by an interconnected network of neural processes could be tested in the future by combining neurostimulation with functional neuroimaging measures (33). Irrespective of these considerations, the current demonstration of a dedicated neurobiological basis for honesty may have important implications for jurisdiction and legal systems, in which the biological limits on the responsibility for legal transgressions are intensely debated (34). Moreover, our findings of a malleable neural process that influences honesty may be important for the development of measures to promote honesty (13). However, our finding that tDCS only enhanced honesty in individuals who experienced a conflict when cheating may prevent establishing such measures for the treatment of pathologies coined by an absence of such conflicts (14).

Materials and Methods

All participants were neurologically and psychiatrically healthy, as ascertained by standardized questionnaires, and did not take any medication at the time of testing. They all gave informed consent, and the procedures received ethical approval from the Ethics Committee of the Canton of Zurich (KEK 2010-0326/3). A total of 145 university students participated in the main study. They were invited to group sessions with 10–12 participants each. In each session, participants were randomly assigned to one of three stimulation groups [anodal (n = 49), cathodal (n = 49), and sham (n = 47) tDCS]. Neither the participants nor the experimenters knew who would receive active or sham stimulation.

The tDCS was applied by means of a multichannel stimulator (NeuroConn) and pairs of standard sponge electrodes soaked in saline solution. One of these electrodes (5 cm × 7 cm) was placed based on frameless stereotaxy over the rDLPFC region of interest found active during honest reporting in a previous fMRI study (19). The other electrode (10 cm × 10 cm) was placed over the vertex. tDCS was applied at an intensity of 1.5 mA for 30 min (in the anodal and cathodal groups) or was switched off after 60 s (in the sham group) to mimic the tingling sensations at the start of the stimulation (20). To minimize the sensations at stimulation onset, the current was linearly ramped up (at the start) and down (at the end) over periods of 20 s.

During tDCS, participants completed a series of experimental tasks, the order of which was counterbalanced across sessions. Honest behavior was assessed with a game of chance requiring subjects to roll a six-sided die

10 times and report the outcomes of the die rolls. In each round, half of the rolled numbers resulted in a payoff of 9 Swiss francs, whereas the remaining numbers yielded no payoff. The payoff scheme for each round was displayed on a computer screen that participants also used to enter the outcome of their roll and the associated payoff. Because participants were fully shielded from view, no one (including the experimenters) could detect whether individual subjects misreported the outcomes of their die rolls. However, it is possible to detect cheating at the group level by comparing the mean percentage of successful die rolls reported by the subjects with the 50% benchmark if everyone reported honestly (35).

The other experimental measures acquired in each session included a dictator game to measure selfish behavior, an investment task to measure preferences for risky and ambiguous outcomes, and a delay discounting task to measure impulsivity. These additional tasks helped to disguise the purpose of the experiment by reducing the prominence of the die-rolling task, and they allowed us to control for other aspects of choice behavior that may be affected by the stimulation (24, 26, 27, 36). After stimulation was switched off [but while participants were still under the lasting influence of physiological tDCS aftereffects (20, 37)], we collected various questionnaire measures of our participants' beliefs about the tasks and their everyday behavior. At the end of the experiment, the computer randomly selected one of the four experimental tasks and the resulting individual choice-dependent payoffs were paid to the participants.

We also conducted a prosocial cheating experiment with an additional 156 participants [anodal ($n = 78$) and sham ($n = 78$)] for which we modified the die-rolling task so that subjects could not earn any money for themselves; instead, their earnings were credited to another anonymous participant. All other aspects of the experimental design and procedure were identical to the main experiment.

The honesty data were analyzed by means of comprehensive regression analyses that controlled for potential group differences in various participant-related measures and potential tDCS effects on the control tasks. These analyses were complemented by direct group comparisons focusing on specific aspects of the honesty data. Moreover, numerical simulations and Bayesian analyses were used to ascertain that our findings are highly unlikely to reflect purely random variations in the outcomes of our specific sample. A detailed description of the sample, experimental procedures, and statistical analyses is given in *SI Appendix* (21).

ACKNOWLEDGMENTS. We thank Charles Efferson, Thomas Epper, and David Tannenbaum for helpful comments and the staff of the Laboratory for Social and Neural Systems Research for practical support. Financial support from the Swiss National Science Foundation (CRSII3_141965) and Gottlieb Duttweiler Institute is gratefully acknowledged.

1. Wang JT-y, Spezio M, Camerer CF (2010) Pinocchio's pupil: Using eyetracking and pupil dilation to understand truth telling and deception in sender-receiver games. *Am Econ Rev* 100:984–1007.
2. Irlenbusch B, Villeval MC (2015) Behavioral ethics: How psychology influenced economics and how economics might inform psychology. *Curr Opin Psychol* 6:87–92.
3. Weisel O, Shalvi S (2015) The collaborative roots of corruption. *Proc Natl Acad Sci USA* 112:10651–10656.
4. Gächter S, Schulz JF (2016) Intrinsic honesty and the prevalence of rule violations across societies. *Nature* 531:496–499.
5. Cohn A, Fehr E, Maréchal MA (2014) Business culture and dishonesty in the banking industry. *Nature* 516:86–89.
6. Artavanis N, Morse A, Tsoutsoura M (2016) Measuring income tax evasion using bank credit: Evidence from Greece. *Q J Econ* 1131:739–798.
7. The Tax Justice Network (2011) The Cost of Tax Abuse: A briefing paper on the cost of tax evasion worldwide. Available at www.taxjustice.net/wp-content/uploads/2014/04/Cost-of-Tax-Abuse-TJN-2011.pdf. Accessed September 6, 2016.
8. World Bank Institute (2004) The Costs of Corruption. Available at web.worldbank.org/WBSITE/EXTERNAL/NEWS/0,,contentMDK:20190187~menuPK:34457~pagePK:34370~piPK:34424~theSitePK:4607,00.html. Accessed September 6, 2016.
9. Trivers R (2011) *Deceit and Self-Deception: Fooling Yourself the Better to Fool Others* (Allen Lane, London).
10. Shalvi S, Dana J, Handgraaf MJJ, De Dreu CKW (2011) Justified ethicality: Observing desired counterfactuals modifies ethical perceptions and behavior. *Organ Behav Hum Decis Process* 115:181–190.
11. Fischbacher U, Föllmi-Heusi F (2013) Lies in disguise. An experimental study on cheating. *J Eur Econ Assoc* 11:525–547.
12. Maggian V, Villeval MC (2016) Social preferences and lying aversion in children. *Exp Econ* 19:663–665.
13. Wolpe PR, Foster KR, Langleben DD (2010) Emerging neurotechnologies for lie-detection: Promises and perils. *Am J Bioeth* 10:40–48.
14. Snyder S (1986) Pseudologia fantastica in the borderline patient. *Am J Psychiatry* 143:1287–1289.
15. Sip KE, Roepstorff A, McGregor W, Frith CD (2008) Detecting deception: The scope and limits. *Trends Cogn Sci* 12:48–53.
16. Zhu L, et al. (2014) Damage to dorsolateral prefrontal cortex affects tradeoffs between honesty and self-interest. *Nat Neurosci* 17:1319–1321.
17. Volz KG, Vogeley K, Tittgemeyer M, von Cramon DY, Sutter M (2015) The neural basis of deception in strategic interactions. *Front Behav Neurosci* 9:27.
18. Sutter M (2009) Deception through telling the truth. Experimental evidence from individuals and teams. *Economic J* 119:47–60.
19. Greene JD, Paxton JM (2009) Patterns of neural activity associated with honest and dishonest moral decisions. *Proc Natl Acad Sci USA* 106:12506–12511.
20. Woods AJ, et al. (2016) A technical guide to tDCS, and related non-invasive brain stimulation tools. *Clin Neurophysiol* 127:1031–1048.
21. Cohn A, Maréchal MA, Noll T (2015) Bad boys: How criminal identity salience affects rule violation. *Rev Econ Stud* 82:1289–1308.
22. Cohn A, Maréchal MA (2016) Laboratory measure of cheating predicts school misconduct. Working paper no. 205 (University of Zurich, Department of Economics, Zurich). Available at www.econ.uzh.ch/static/workingpapers.php?id=884.
23. Fedorenko E, Duncan J, Kanwisher N (2013) Broad domain generality in focal regions of frontal and parietal cortex. *Proc Natl Acad Sci USA* 110:16616–16621.
24. Mohr PNC, Biele G, Heekeren HR (2010) Neural processing of risk. *J Neurosci* 30:6613–6619.
25. Huettel SA, Stowe CJ, Gordon EM, Warner BT, Platt ML (2006) Neural signatures of economic preferences for risk and ambiguity. *Neuron* 49:765–775.
26. Hare TA, Camerer CF, Rangel A (2009) Self-control in decision-making involves modulation of the vmPFC valuation system. *Science* 324:646–648.
27. Figner B, et al. (2010) Lateral prefrontal cortex and self-control in intertemporal choice. *Nat Neurosci* 13:538–539.
28. Jacobson L, Koslowsky M, Lavidor M (2012) tDCS polarity effects in motor and cognitive domains: A meta-analytical review. *Exp Brain Res* 216:1–10.
29. Erat S, Gneezy U (2012) White lies. *Manage Sci* 58:723–733.
30. Dunbar RIM (2009) The social brain hypothesis and its implications for social evolution. *Ann Hum Biol* 36:562–572.
31. Loewen PJ, et al. (2013) The heritability of moral standards for everyday dishonesty. *J Econ Behav Organ* 93:363–366.
32. Garrett N, Lazzaro SC, Ariely D, Sharot T (2016) The brain adapts to dishonesty. *Nat Neurosci* 19:1727–1732.
33. Driver J, Blankenburg F, Bestmann S, Vanduffel W, Ruff CC (2009) Concurrent brain-stimulation and neuroimaging for studies of cognition. *Trends Cogn Sci* 13:319–327.
34. Aspinwall LG, Brown TR, Tabery J (2012) The double-edged sword: Does biomechanism increase or decrease judges' sentencing of psychopaths? *Science* 337:846–849.
35. Houser D, Vetter S, Winter J (2012) Fairness and cheating. *Eur Econ Rev* 56:1645–1655.
36. Ruff CC, Ugazio G, Fehr E (2013) Changing social norm compliance with noninvasive brain stimulation. *Science* 342:482–484.
37. Nitsche MA, Paulus W (2001) Sustained excitability elevations induced by transcranial DC motor cortex stimulation in humans. *Neurology* 57:1899–1901.