ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Schunk, Daniel; Fehr, Ernst; Knoch, Daria; Schneider, Frédéric; Hohmann, Martin

Conference Paper

Disrupting the prefrontal cortex diminishes the human ability to build a good reputation

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie - Session: Trust, Honesty, and Reputation: Neuroeconomic Evidence, No. C5-V1

Provided in Cooperation with:
Verein für Socialpolitik / German Economic Association

Suggested Citation: Schunk, Daniel; Fehr, Ernst; Knoch, Daria; Schneider, Frédéric; Hohmann, Martin (2010) : Disrupting the prefrontal cortex diminishes the human ability to build a good reputation, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie - Session: Trust, Honesty, and Reputation: Neuroeconomic Evidence, No. C5-V1, Verein für Socialpolitik, Frankfurt a. M.

This Version is available at:
https://hdl.handle.net/10419/37512

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Disrupting the prefrontal cortex diminishes the human ability to build a good reputation

Daria Knoch[a,1], Frédéric Schneider[b,1], Daniel Schunk[b,1], Martin Hohmann[a], Ernst Fehr[b,c]

[a] Social and Affective Neuroscience, Faculty of Psychology, University of Basel, Birmannsgasse 8, 4055 Basel, Switzerland

[b] Institute for Empirical Research in Economics, University of Zurich, Blümlisalpstrasse 10, 8006 Zurich, Switzerland

[c] Collegium Helveticum, Schmelzbergstrasse 25, 8092 Zurich, Switzerland

[1] These authors contributed equally to this work

Classification

SOCIAL SCIENCES: Economic Sciences

BIOLOGICAL SCIENCES: Neuroscience

Manuscript length: 17 text pages; 4 figure pages; 0 table pages

Correspondence and requests should be addressed to D.K. (e-mail: daria.knoch@unibas.ch) and E.F. (e-mail: efehr@iew.uzh.ch)

Reputation formation pervades human social life. In fact, many people go to great lengths to acquire a good reputation, although building a good reputation is costly in many cases. We know little, however, about the neural underpinnings of this important social mechanism. In the present study we show that disruption of the right, but not the left, lateral prefrontal cortex (PFC) with low-frequency repetitive transcranial magnetic stimulation (rTMS) diminishes subjects' ability to build a favorable reputation. This effect occurs even though subjects' ability to behave altruistically in the absence of reputation incentives remains intact, and even though they are still able to recognize both the fairness standards necessary for acquiring and the future benefits of a good reputation. Thus, subjects with a disrupted right lateral PFC no longer seem to be able to resist the temptation to defect, even though they know that this has detrimental effects on their future reputation. This suggests an important dissociation between the knowledge about one's own best interests and the ability to act accordingly in social contexts. These results link findings on the neural underpinnings of self-control and temptation with the study of human social behavior, and they may help explain why reputation formation remains less prominent in most other species with less developed prefrontal cortices.

Introduction. Humans are unique to the extent to which social norms regulate their lives and reputation formation is a powerful mechanism in generating norm compliance. "Reputation is what you are in the light; character is what you are in the dark", says a Chinese proverb. In other words, although much norm compliance is voluntary, there is ample evidence that people are more likely to comply with norms when they feel observed by others. In such a situation ("in the light"), individuals signal their quality as cooperators to future interaction

partners, thereby forming a good reputation.

Reputation formation is characterized by two features: First, the signals for building a good reputation (in human societies) are costly in many cases; otherwise they would be "cheap talk", and therefore of no informational value for the potential interaction partner. Second, this process of costly reputation formation is characterized by a trade-off between the current benefits of defection and one's future benefits of a good reputation.

Evidence for the crucial role of a good reputation in social decision-making comes from empirical studies showing that individuals increase their levels of cooperation and are more likely to comply with norms when they know that others observe their behavior, and that individuals cooperate with those whom they observe cooperating with others[1-16]. Hence, the concern for reputation profoundly affects our daily social interactions and motivates many important decisions in our lives.

Although reputation formation mechanisms are ubiquitous in social exchange, their neurobiological substrate remains largely unknown. Moreover, a universal question arises, one with relevance not only to cognitive neuroscience but also to fields of research in evolutionary biology, developmental psychology, and behavioral economics: Which skills are required in order to acquire a good reputation? Intuitively, we assume that there must be a self-control capacity because forming a reputation typically requires an individual to overcome the temptation to defect in order to gain future reputation benefits. From a neurobiological perspective, we thus assume the involvement of the PFC, as this region has been shown to be involved in self-control processes [17-19].

Four previous neuroimaging studies have examined reputation[20-23]. Two of these studies did not address the neural underpinnings of the process of individual reputation formation,

i.e. they did not focus on the individual who forms a reputation. Instead, they examined individuals who made decisions based on reputation information about another individual[20,21]. In one of the studies, for example, subjects played iterative trust games with three partners whose (fictional) profiles make them seem morally good, bad, or neutral[21]. The study shows that information about the interaction partner's moral reputation affects the investors' reward prediction error signal in the caudate nucleus during reciprocal exchange. Another study showed that reward-related brain areas were activated when a subject learned that others perceive his or her reputation as good[22]. Finally, one study used hyperscanning fMRI while two interacting partners played an iterated trust game and showed that the peak activation of the caudate nucleus underwent a temporal shift as the reputation of the interaction partner developed[23].

However, none of these studies provides causal evidence about the brain processes involved in costly reputation formation. Functional imaging methods, although indispensable, do not permit causal inferences about the effect of brain processes on human behavior because the observed neural activations could be spuriously correlated with task performance and need not necessarily play a causal role in task execution[24,25]. In contrast, brain stimulation techniques such as transcranial magnetic stimulation interfere non-invasively with the activity of defined areas in the human cortex, thus enabling researchers to observe the behavioral impact of an increase or decrease in the cortical excitability of the stimulated brain region. 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 about half the duration of the stimulation[26,27]. Here, we investigated the effect of disrupting the PFC by means of rTMS on subjects' reputation formation.

We chose a version of the trust game (Fig. 1A) as a vehicle for investigating the effects of rTMS on costly reputation formation. Subjects played 15 periods of this trust game with randomly re-matched partners each period (see Supplementary Information).

We implemented two treatment conditions, an “anonymous condition” and a “reputation condition”. In the anonymous condition, the trustee's previous decisions are unknown to the current investor, while the investor has information about the trustee in the reputation condition (Fig. 1B). He can observe the trustee’s decisions in the previous three periods (i.e., how many times the trustee chose to back-transfer “nothing”, “a quarter” or “equalize payoff”). Hence, a trustee is likely to acquire a bad reputation by choosing to back-transfer “nothing”, whereas a trustee improves his reputation by choosing “equalize payoff”. As a trustee who transfers nothing is unlikely to receive high transfers from the investors in future periods, the trustees have an incentive to make relatively high back-transfers in the reputation condition. This reputation incentive therefore generates a motivational conflict for the trustees. A trustee could maximize his short-run self-interest by choosing to transfer nothing back to the investor in the current period, but this action is likely to have detrimental effects for his reputation and decreases future investors’ willingness to transfer money to him. Therefore, in order to reap the benefits from a good reputation in future periods, a trustee must constrain his immediate self-interest and forgo the current option of transferring back nothing.

In contrast, the strategic incentive for behaving in a cooperative manner is fully absent in the anonymous condition because the investors have no information about the trustees’ past behavior. In terms of the Chinese proverb cited at the beginning the trustees in the anonymous condition act in perfect darkness and only their “character” plays a role. Thus,

the anonymous condition measures how much the trustee is willing to return voluntarily to the investor (which may be viewed as a form of altruistic behavior). This amount reflects the trustee's preference for back-transferring if there are no strategic reputation incentives. Consequently, if the trustee returns amount X to the investor in the anonymous condition, then the trustee is apparently not willing to return more than X. However, the trustee very well might return more than X in the reputation condition, as strategic incentives for reputation formation are then present, i. e. the trustee must override his immediate self-interest in this case to build a good reputation.

How will the disruption of the PFC with low-frequency rTMS affect the trustees' behavior? As the lateral PFC has been shown to be reliably involved in overriding prepotent responses and self-control processes[17-19], and as costly reputation formation requires overriding immediate benefits, disrupting this brain region should functionally weaken the self-control capacity and should thus lead to a lower back-transfer in the reputation condition compared to the other stimulation groups. In contrast, little or no self-control effort is involved in the anonymous condition because the trustee has no reputational incentive to back-transfer more than his immediate preference dictates. Therefore, we would expect little difference between the stimulation groups for the anonymous condition. Moreover, as the right lateral PFC in particular has been shown to be involved in control capacities[17,19], we hypothesized that disruption of the right, but not the left, lateral PFC will lead to difficulties in resisting the temptation to go for the immediate benefit and thus reduce the ability to form a good reputation.

It is important to note that the trustee in the reputation condition knows that the investor only has information about his three previous choices (i.e., "nothing", "a quarter" or

"equalize payoff"), but not about how high the corresponding previous investors' investments were in the last three periods. For example, if a trustee receives an investment of 1 point and chooses "equalize payoff" in order to form a good reputation, he actually back-transfers 2.5 points. This is because he in fact received 4 points (1 point, quadrupled by the experimenter) and back-transferring 2.5 of those 4 points together with the initial endowment of 10 points leaves both the investor and the trustee with equal amounts totaling 11.5 points. If a trustee receives an investment of 10 points and chooses "equalize payoff", his back-transfer is 25 points, and both players end up with a total of 25 points. Because future investors will only observe the choice "equalize payoff", and not the amount actually transferred by the investor, the trustee's reputational benefit is the same in both cases. The immediate costs, however, are different: 2.5 points in the first case, compared to 25 points in the second case. Therefore, the costs of reputation formation (i.e., the number of points the trustee has to forego in order to form a good reputation) vary with the size of the investment while the effect of a particular choice on a trustee's reputation is always the same, regardless of the received investment. In other words, while the future reputational value of a trustee's choice is independent of the investor's investments, the immediate cost of reputation formation, and thus the temptation to maximize one's short-run self-interest, varies with the size of the investment. Therefore, the self-control effort necessary to constrain short-run self-interest is likely to be much higher in case of a large investment compared to a small investment where reputation formation is almost costless. This variation in the temptation to maximize one's short-run self-interest by paying back nothing enables us to investigate whether the effect of disrupting the lateral PFC depends on the degree of self-control required for reputation formation.

This feature of our design puts important constraints on the interpretation of a possible effect of right lateral PFC disruption. If, for example, rTMS of right PFC primarily reduces the trustees' back-transfers in the reputation condition at low investment levels, an interpretation of this effect in terms of reduced self-control abilities is less convincing because little temptation to defect exists at low investment levels. However, if rTMS primarily affects their back-transfers at high investment levels where the temptation to defect is high, an interpretation in terms of reduced self-control makes a lot of sense. In order to examine whether the lateral PFC activity is crucial in the trustees' reputation formation, we applied low-frequency rTMS over the dorsolateral prefrontal cortex (DLPFC) for 15 minutes to healthy subjects in the role of the responder (see Methods).

Results. Our results show that reputation formation pays off for the trustees in the long run, since investors give more points to trustees who cooperated in the past than to defectors. Trustees have a 71% probability of receiving a 10 points investment if they had always equalized payoffs, but this drops to a probability of less than 6% if they always chose to back-transfer "nothing". Consequently, a strategy of cooperating in the first 14 periods and defecting in the last one (rational cooperation) is on average 43% more profitable (371 points) than always defecting (260 points). Hence, the trustees have an incentive to constrain their short-run self-interest and to back-transfer a high amount in the reputation condition because the investors condition their investments on the trustee's past actions. Accordingly, our results show that trustees cared a lot about their reputation when reputation formation was possible. Subjects sent back on average 24.9% of the transferred amount in the anonymous condition, while in the reputation condition they transferred back 43.8% on average (Fig. 2).

Of primary interest are back-transfers with regard to the investors' highest investment because the temptation to follow short-run self-interest and, therefore, the requirement for self-control effort is greatest in this case. Focusing on the reputation condition (Fig. 3A), we see that the back-transfer for the highest investment was 41.2% following sham rTMS and 48.0% after real rTMS of the left DLPFC. These results contrast sharply with the back-transfer of 29.7% after rTMS of the right DLPFC. The differences in back-transfers across the stimulation groups are significant in the reputation condition (GLS regression, $P<0.001$ for the difference between right DLPFC and left DLPFC, and $P = 0.015$ for the difference between right DLPFC and sham condition, for details see Supplementary Information). In contrast, we found no significant differences in back-transfers between the three stimulation groups in the anonymous condition (Fig. 3B; GLS regression, $P = 0.816$ for the difference between right DLPFC and left DLPFC, and $P = 0.232$ for the difference between right DLPFC and sham).

In other words, while disruption of the right DLPFC significantly reduces back-transfers in the reputation condition in cases of highest investment, it does not do so in the anonymous condition. Thus, there is a significant differential effect of rTMS across stimulations (right DLPFC, left DLPFC, sham) in the reputation condition, but not in the anonymous condition (see Supplementary Information for additional statistical analyses).

Interestingly, those subjects in the reputation condition who received rTMS to the right DLPFC transferred similar amounts back to the investor as those in the anonymous condition (compare Fig. 3A and 3B; $P = 0.667$, *t*-test). Hence, disrupting the right DLPFC completely removed the behavioral impact of the reputation condition while it had no effect on behavior when reputation formation was not possible. Moreover, there are no significant differences

across stimulation groups for lower investments, where the temptation to yield to short-run gains and thus the recruitment of self-control effort is lower (all $P > 0.193$).

rTMS of the right DLPFC limited subjects' ability to override immediate short-run benefits. However, rTMS neither changed subjects' perception of the prevailing fairness norm, nor their ability to assess the consequences of past and current trustee behaviors on future investments, which we elicited immediately after the experiment (see Methods). First, subjects in all three stimulation groups judged the scenario of transferring back nothing in response to an investment of 7 as rather unfair, and there are no differences in fairness judgments across groups ($P = 0.376$, Kruskal-Wallis test). Second, rTMS of the right DLPFC did not change subjects' ability to assess the consequences of past and current trustee behaviors, since subjects in the different stimulation groups predict the same investments by future investors in response to a given profile of past back-transfers ($P = 0.950$, Kruskal-Wallis test). Moreover, if rTMS of the right DLPFC had impaired subjects' general ability to perform complex calculations, then we would have observed differences across stimulation groups also for the lower investments, but our results show a behavioral effect only for highest investments. Thus, the disruption of the right DLPFC has an effect on the behavioral ability to form a good reputation, even though it affected neither subjects' ability to perform complex cognitive operations nor their recognition of the prevailing fairness norm nor their ability to assess the future consequences of back-transfer behaviors.

We also checked whether individual differences in impulsivity and the propensity to reciprocate kind or hostile acts can explain our results. We find that neither dispositional differences in subjects' reciprocity norm nor individual differences in impulsivity across treatment groups can explain the behavioral differences across conditions: there was no

difference across treatments for impulsivity (BIS scale: $P = 0.827$, BAS scale: $P = 0.967$, Kruskal-Wallis tests) and no difference across treatments for the reciprocity norm (positive reciprocity scale: $P = 0.741$, negative reciprocity scale: $P = 0.971$, Kruskal-Wallis tests).

Discussion. The results reported above indicate a highly specific, lateralized effect of a disrupted function of the lateral PFC on the ability to form a reputation for being trustworthy. First, we do not find any differences between the stimulation groups in the anonymous condition where the incentives for reputation formation are absent. In this condition, only subjects' preferences for altruistic behaviors can induce them to repay trust, implying that an interference with the function of the right lateral PFC leaves subjects' altruistic propensities to behave in a trustworthy manner unchanged. This contrasts, second, with an rTMS effect in those circumstances in which costly reputation formation requires a particularly strong recruitment of self-control effort, i.e. when the investors make a high investment. In this situation, the incentive to yield to the short-run costs for building a reputation are largest, suggesting an interpretation of the rTMS effect in terms of the reduced ability to recruit the required self-control resources. Third, the absence of any rTMS effect on subjects' ability to recognize the prevailing fairness norm supports this. Thus, despite the fact that subjects are well aware of the existing fairness norm, and although they have pecuniary incentives to obey this norm in the reputation condition, they nevertheless do not follow it, suggesting that rTMS causes a specific inability to constrain short-run temptations rather than a cognitive inability to perceive the normative demands involved in the situation. Fourth, the fact that rTMS has no effect on subjects' ability to assess the future consequences of past back-transfers further supports our interpretation. Subjects across all three stimulation conditions have the same knowledge about the future benefits of high current back-transfers, but only

subjects with transiently disrupted right DLPFC function are less able to constrain their short-run self-interest and thus exploit this knowledge.

Taken together, these results support the hypothesis that right, but not left, lateral PFC activity is crucial in the ability to forego immediate benefits in order to form a good reputation. Subjects whose right lateral PFC is disrupted behave as if they were not concerned about their reputation when reputation formation required forgoing a large current benefit, suggesting that they are less able to pay an immediate cost for future social reputation benefits although their ability to assess these benefits cognitively remains intact. These findings suggest an important dissociation at the neurobiological level between the knowledge about what is in one's own best interest in social interaction situations and the ability to act accordingly. Moreover, by providing causal evidence on the role of the prefrontal cortex for costly reputation formation, our findings may also help explain why reputation mechanisms are rare in other species, whose prefrontal regions are less developed. However, brain areas do not act in isolation of each other in highly complex processes such as reputation formation, but rather must work together as a network. Future studies could combine low frequency rTMS and fMRI to understand how different brain regions interact on the functional-anatomical level in reputation formation.

Materials and Methods.

Transcranial magnetic stimulation. We applied low-frequency rTMS for 15 minutes to 87 healthy subjects in the role of the trustee (see Supplemental Information for further details). In order to investigate a possible hemispheric laterality in the role of lateral PFC on trustees' decisions, we applied rTMS to the right DLPFC or to the left DLPFC. The existence of a

stimulation group receiving rTMS to the right DLPFC and a control group that receives rTMS to the left DLPFC is also important because this controls for the potential side effects of rTMS[28], including discomfort, irritation, and mood changes. In addition, we had a further control condition where we applied sham stimulation for 15 minutes to the right or left DLPFC. As mentioned above, we implemented an anonymous condition and a reputation condition. Hence, the experiment has a 2x3 design with the factors "condition" (anonymous, reputation) and "stimulation" (left rTMS, right rTMS, sham) leading to six experimental groups (see Table S1). We randomly assigned subjects to one of the six groups.

Measurement of fairness norms. Because the disruption of the PFC might also affect subjects' perception of what constitutes the social norm in a certain situation, we further elicited individuals' perception of fairness norms immediately after the trust game by confronting them with a hypothetical scenario. We asked participants to judge the fairness of a hypothetical trustee's behavior on a seven point scale from "very unfair" to "very fair". The scenario described an investor who invests 7 points while the trustee returns nothing.

Measurement of the ability to assess the consequences of past and current trustee behaviors. The disruption of the PFC might also affect subjects' ability to assess the consequences of a particular reputation, i.e. to assess the impact of one's actions on future social interaction, an abstract and cognitively demanding task. In order to rule out this explanation, we gave subjects another scenario to measure an individual's assessment of the potential consequences of a certain reputation. We asked subjects how many points (1, 4, 7 or 10 points) they would expect an investor to transfer to a trustee who had chosen to "equalize payoff" twice and to back-transfer "nothing" once.

Measurement of dispositional differences in impulsivity and reciprocity. Subjects completed personality questionnaires that assessed their impulsivity[29] – using the Behavioral Inhibition

System (BIS) and Behavioral Approach System (BAS) scales – and their personal norm of reciprocity.[30] These questionnaires were completed roughly 10 days after the experiment.

Further details about our experimental protocol, participant instructions and further analyses are contained in the Supplementary Material.

Acknowledgements

This paper is part of the research priority program at the University of Zurich on the "Foundations of Human Social Behavior – Altruism versus Egoism" and the research program of the Collegium Helveticum on the emotional foundations of moral behavior, which is supported by the Cogito Foundation. DK thanks the Swiss National Science Foundation (PP00P1-123381). The authors thank Antonio Rangel and Michel Marechal for helpful comments.

References

1. Nowak MA, Sigmund K (2005) Evolution of indirect reciprocity. *Nature* 437:1291–1298.
2. Milinski M, Semmann D, Krambeck, HJ (2002) Reputation helps solve the 'tragedy of the commons'. *Nature* 415:424–426.
3. Wedekind C, Milinski M (2000) Cooperation through image scoring in humans. *Science* 288:850–852.
4. Nowak MA, Sigmund K (1998) Evolution of indirect reciprocity by image scoring. *Nature* 393:573–577
5. Panchanathan K, Boyd R (2004) Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature* 432:499–502.
6. Brown M, Falk A, Fehr E (2004) Relational contracts and the nature of market interactions. *Econometrica* 72:747–780.
7. Camerer C, Weigelt K (1988) Experimental tests of a sequential equilibrium reputation model. *Econometrica* 56:1–36.
8. Fehr E, Brown M, Zehnder C (2009) On reputation: A microfoundation of contract enforcement and price rigidity. *Econ J* 119:333–353
9. Basua S, Dickhaut J, Hecht G, Towry K, Waymire G (2009) Recordkeeping alters economic history by promoting reciprocity. Proc Natl Acad Sci USA 106:1009–1014.
10. Houser D, Wooders J (2006) Reputation in auctions: Theory, and evidence from eBay. *J Econ Manag Strategy* 15:353–370.
11. Keser C, van Winden F (2000) Conditional cooperation and voluntary contributions to public goods. *Scand J Econ* 102:23–39
12. Fehr E, Gächter S (2000) Cooperation and punishment in public goods experiments. *Am Econ Rev* 90:980–994

13. Falk A, Gächter S, Kovacs J (1999) Intrinsic motivation and extrinsic incentives in a repeated game with incomplete contracts. *J Econ Psychol* 20:251–284
14. Cochard F, Nguyen Van P, Willinger M (2004) Trusting behavior in a repeated investment game. *J Econ Behav Organ* 55:31–44
15. Engelmann D, Fischbacher U (2009) Indirect reciprocity and strategic reputation building in an experimental helping game. *Games Econ Behav* (forthcoming)
16. Seinen I, Schram A (2006) Social status and group norms: Indirect reciprocity in a repeated helping experiment. *Eur Econ Rev* 50:581–602
17. Aron AR, Robbins TW, Poldrack RA (2004) Inhibition and the right inferior frontal cortex. *Trends Cogn Sci* 8:170–177.
18. Miller EK, Cohen JD (2001) An integrative theory of prefrontal function. *Annu Rev Neurosci* 24:167–202.
19. 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.
20. Takahashi H *et al.* (2008) Neural correlates of human virtue judgment. *Cereb Cortex* 18:1886–1891.
21. 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.
22. Izuma K, Saito DN, Sadato N (2008) Processing of social and monetary rewards in the human striatum. *Neuron* 58:284–294.
23. King-Casas B *et al.* (2005) Getting to know you: Reputation and trust in a two-person economic exchange. *Science* 308:784–83.
24. Walsh V, Cowey A (2000) Transcranial magnetic stimulation and cognitive neuroscience. *Nat Rev Neurosci* 1:73–79.
25. Sack AT, Linden DE (2003) Combining transcranial magnetic stimulation and functional imaging in cognitive brain research: Possibilities and limitations. *Brain Res Brain Rev* 43:41–56.
26. Robertson EM, Théoret H, Pascual-Leone A (2003) Studies in cognition: The problems solved and created by transcranial magnetic stimulation. *J Cogn Neurosci* 15:948–960.
27. Eisenegger C, Treyer V, Fehr E, Knoch D (2008) Time-course of "off-line" prefrontal rTMS effects – A PET study. *Neuroimage* 42:379–384.
28. Abler B *et al.* (2005) Side effects of transcranial magnetic stimulation biased task. Performance in a cognitive neuroscience study. *Brain Topogr* 17:193–196.
29. Carver CS, White TL (1994) Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *J Pers Soc Psychol* 67:319–333
30. Perugini M, Callucci M, Presaghi F, Ercolani AP (2003) The personal norm of reciprocity. *Eur J Personality* 17:251–283.

Supplementary Information

Tables S1, S2, S3, S4, S5

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Figure Legends

Figure 1A

Schematic representation of the task. Design of one period of the trust game. In each period, two anonymous individuals, a first mover (investor) and a second mover (trustee), receive an endowment of 10 points each. The investor must decide how many points he wants to transfer to the trustee. The experimenter quadruples the invested points and transfers them to the trustee, who then decides how many points he would like to back-transfer to the investor. In order to reduce the cognitive complexity of the game, the strategy space was limited for both the investor and the trustee. The investor could transfer 1, 4, 7, or 10 points (1 point equals 0.20 CHF which was about 0.18 US-$ at the time the experiment was conducted) and the trustee had three options: he could back-transfer nothing, a quarter of the received amount, or he could back-transfer an amount that equalized the period payoff between the investor and the trustee. The latter restriction also has the added advantage that the reputational implications of different trustee behaviors are transparent. For example, paying back nothing is unambiguously bad for the formation of a good reputation, while equalizing payoffs is unambiguously good.

Figure 1B

Trustee's decision screen in the reputation condition. On the top section of the screen, the trustee can see his decisions in the three previous periods. The information about past decisions is not in chronological order. The trustee also knows that in this condition the investor is informed about the trustee's previous three decisions before he makes a transfer decision. Thus, the trustee knows that the investor can condition her transfer on the trustee's previous three back-transfer decisions. This also means that the trustee's back-transfer in the current period affects the information the future investors receive about him, i.e. it affects his reputation in future periods. The trustee's decision screen also contains information about the current investor's transfer and the resulting points at the trustee's disposal in the mid section of the screen. The bottom section features three clickable buttons for the trustee's decision.

Figure 2

The trustees' behavioral responses. Comparison of mean of back-transfer in the reputation vs. the anonymous condition; data pooled over all stimulation groups.

Figure 3

The trustee's behavioral responses across stimulation conditions. **(A)** Mean back-transfer associated with the investor's highest investment in the reputation condition. Subjects whose right DLPFC is disrupted back-transfer significantly less points than those in the other two stimulation groups ($P < 0.02$). **(B)** Mean back-transfer associated with the investor's highest investment in the anonymous condition.

Figures

Figure 1A

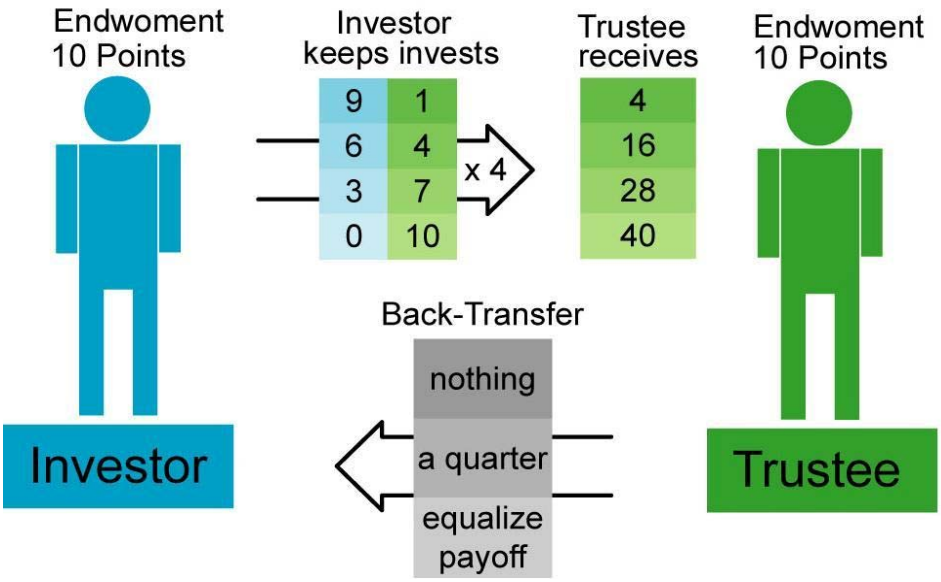


Figure 1B

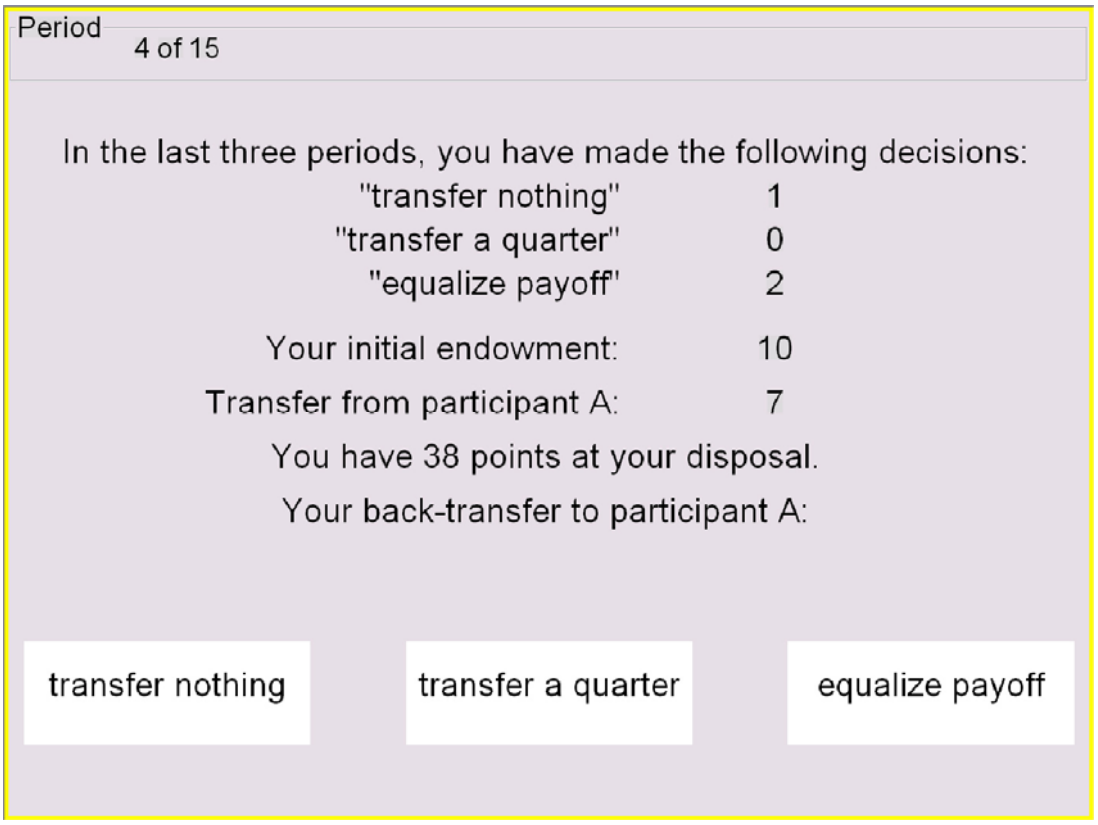

Period
4 of 15
In the last three periods, you have made the following decisions:
"transfer nothing" 1
"transfer a quarter" 0
"equalize payoff" 2
Your initial endowment: 10
Transfer from participant A: 7
You have 38 points at your disposal.
Your back-transfer to participant A:
transfer nothing
transfer a quarter
equalize payoff

Figure 2

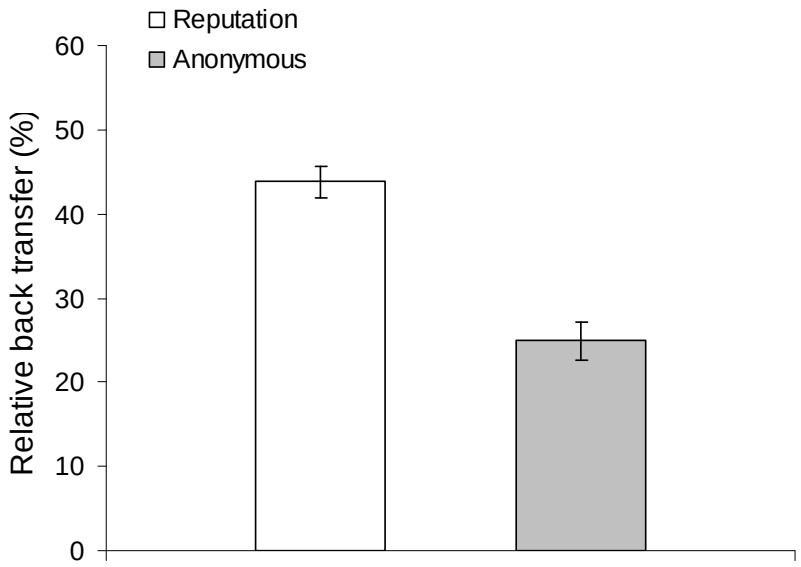
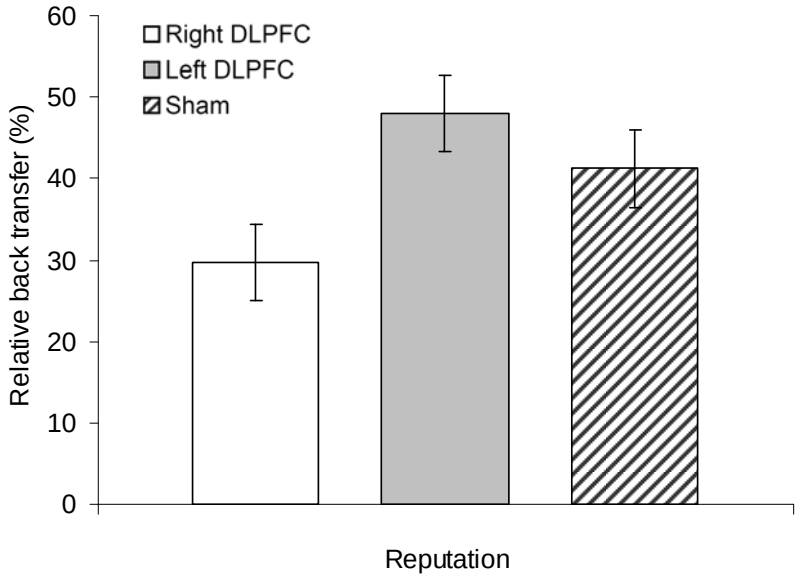
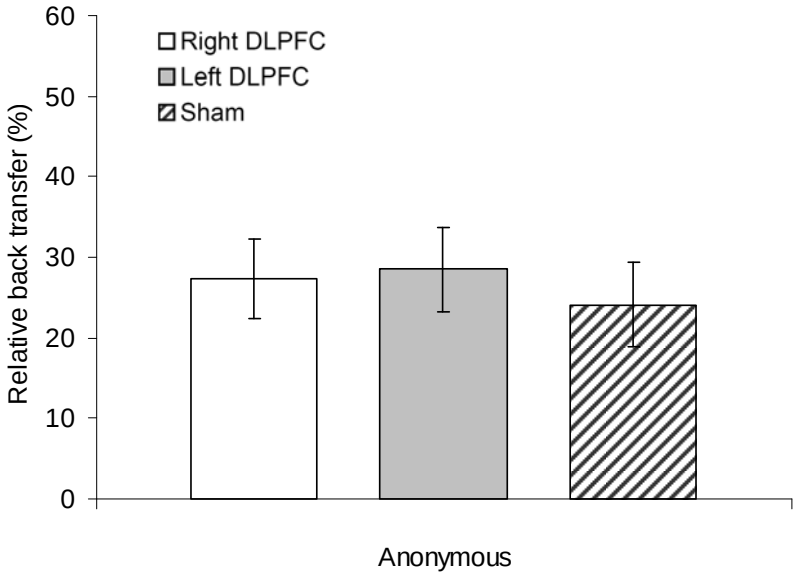

Reputation
Anonymous
Relative back transfer (%)
60
50
40
30
20
10
0

Figure 3

A


60
50
40
30
20
10
0
Relative back transfer (%)
Right DLPFC
Left DLPFC
Sham
Reputation

B


60
50
40
30
20
10
0
Relative back transfer (%)
Right DLPFC
Left DLPFC
Sham
Anonymous