

Psychological Science

<http://pss.sagepub.com/>

A Neural Marker of Costly Punishment Behavior

Daria Knoch, Lorena R.R. Gianotti, Thomas Baumgartner and Ernst Fehr
Psychological Science 2010 21: 337 originally published online 1 February 2010
DOI: 10.1177/0956797609360750

The online version of this article can be found at:
<http://pss.sagepub.com/content/21/3/337>

Published by:



<http://www.sagepublications.com>

On behalf of:



[Association for Psychological Science](http://www.sagepublications.com)

Additional services and information for *Psychological Science* can be found at:

Email Alerts: <http://pss.sagepub.com/cgi/alerts>

Subscriptions: <http://pss.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

A Neural Marker of Costly Punishment Behavior

Daria Knoch¹, Lorena R.R. Gianotti^{1,2}, Thomas Baumgartner^{1,3}, and Ernst Fehr^{3,4}

¹Social and Affective Neuroscience, Department of Psychology, University of Basel; ²The KEY Institute for Brain-Mind Research, University Hospital of Psychiatry, University of Zurich; ³Institute for Empirical Research in Economics, University of Zurich; and ⁴Collegium Helveticum, Zurich, Switzerland

Psychological Science
 21(3) 337–342
 © The Author(s) 2010
 Reprints and permission:
sagepub.com/journalsPermissions.nav
 DOI: 10.1177/0956797609360750
<http://pss.sagepub.com>



Abstract

Human readiness to incur personal costs to punish norm violators is a key force in the maintenance of social norms. The willingness to punish is, however, characterized by vast individual heterogeneity that is poorly understood. In fact, this heterogeneity has so far defied explanations in terms of individual-level demographic or psychological variables. Here, we use resting electroencephalography, a stable measure of individual differences in cortical activity, to show that a highly specific neural marker—baseline cortical activity in the right prefrontal cortex—predicts individuals' punishment behavior. The analysis of task-independent individual variation in cortical baseline activity provides a new window into the neurobiology of decision making by bringing dispositional neural markers to the forefront of the analysis.

Keywords

costly punishment behavior, prefrontal cortex, resting EEG, individual differences, decision making, source localization

Received 1/22/09; Revision accepted 7/25/09

Civilized human life depends on the maintenance of elementary social norms, many of which are enforced by individuals' willingness to sanction violations even at personal cost. A large amount of cultural and individual variation characterizes costly punishment behavior, however (Camerer, 2003; Henrich et al., 2005; Herrmann, Thoni, & Gächter, 2008). Recent attempts to explain cultural variation in terms of key economic and group variables have been relatively successful (Henrich et al., 2005; Herrmann et al., 2008), but the sources of individual variation are still very poorly understood. Typically, individual variables such as gender, income, wealth, or education have low predictive power, and their influence varies strongly, depending on the details of the study (Camerer, 2003). Recent evidence has suggested a genetic component of up to 40% of individual variation in costly punishment behavior. In a study by Wallace, Cesarini, Lichtenstein, and Johannesson (2007), in which both monozygotic and dizygotic twins played the ultimatum game, the behavioral correlation between monozygotic twins was much larger than in dizygotic twins.

This study is consistent with the idea that dispositional variation might explain individual variation in costly punishment, but so far there is no evidence indicating that stable psychological personality traits or stable neural characteristics affect punishment behavior. Therefore, we examined whether

dispositional differences in neural baseline activity explained individual variation in punitive behavior. Answering this question required a measure that was both stable over time and allowed the examination of individuals' neurophysiological characteristics at rest. We used resting electroencephalography (EEG) to measure tonic cortical activity, which is stable over time and can therefore capture dispositional individual differences (e.g., Kondacs & Szabo, 1999; Näpflin, Wildi, & Sarnthein, 2007; Tomarken, Davidson, Wheeler, & Kinney, 1992).

The ultimatum game nicely illustrates that individuals indeed use costly punishment to enforce social norms. In this bargaining game, a proposer is given a sum of money and makes a proposal of how to split the money between himself or herself and a responder. The responder can either accept the offer, implying the money is divided accordingly, or reject it, and then both players receive nothing. When confronted with a low offer, the responder faces a trade-off between accepting the offered money or rejecting the unfair offer and thus enforcing

Corresponding Authors:

Daria Knoch or Lorena Gianotti, Social and Affective Neuroscience, Department of Psychology, University of Basel, Birmannsgasse 8, CH-4055 Basel, Switzerland

E-mail: daria.knoch@unibas.ch or lorena.gianotti@unibas.ch

a fairness norm. Neuroimaging studies of the ultimatum game have found the lateral prefrontal cortex (PFC) to be involved in regulating reactions to unfair offers (Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003; Tabibnia, Satpute, & Lieberman, 2008). However, these studies have focused on brain activity during the decision-making process rather than examining individuals' pretask neurophysiological characteristics. With resting EEG, we were able to measure cortical activity before task performance.

Recent evidence from brain stimulation studies (Knoch et al., 2007; Knoch, Pascual-Leone, Meyer, Treyer, & Fehr, 2006; van 't Wout, Kahn, Sanfey, & Aleman, 2005) demonstrates that disruption of right lateral PFC function can modulate costly punishment behavior. Subjects chose significantly less costly punishment after disruption of the right lateral PFC compared with disruption of the left lateral PFC or with placebo stimulation (Knoch et al., 2006). Based on this evidence, we hypothesized that differences in the neural functioning of the PFC would explain individual differences in willingness to punish: The higher an individual's baseline cortical activity level in the right PFC, the more the individual would punish unfair behavior.

To investigate whether individual differences in tonic cortical activity level in the PFC predict individuals' propensity to punish in the ultimatum game, we measured brain activity using resting EEG in 20 subjects before they engaged in several one-shot ultimatum games in the role of the responder.

Materials and Method

Subjects

Twenty self-declared right-handed female students (mean age = 26.3 years, $SD = 6.2$) recruited at the University of Zurich participated in the study. All subjects were right-handed according to the Chapmans' Handedness Inventory (scores ranged between 13 and 17; Chapman & Chapman, 1987) and had no history of neurological or psychiatric disorder or alcohol or drug abuse. The study was approved by the local ethics committee. Subjects were remunerated with 30 Swiss francs (CHF 30; CHF 1 = \$1 U.S.) for participating, in addition to the money earned in the ultimatum game.

Procedure

Upon arriving at the laboratory, subjects signed an informed consent form for participation. Subjects were seated in a sound- and electrically shielded EEG recording chamber with dim illumination and intercom connection to the experimenter. They were instructed that EEG recording was to be done during resting with open or closed eyes. The protocol consisted of 20 s eyes open followed by 40 s eyes closed, repeated four times. Only data from the 160-s eyes-closed condition were further analyzed. Thirty minutes after the recording of the resting EEG, all subjects received written instructions for the ultimatum game.

Ultimatum game

Subjects in the role of the responder played 12 ultimatum games each, with 12 different anonymous proposers. We deliberately chose one-shot interactions because no strategic spillovers across trials occur with this structure. This is particularly important if true preferences are to be elicited. The responders had to agree on the division of CHF 20. The proposer made one proposal of how to allocate the CHF 20 by making an offer of CHF 4, 6, 8, or 10 to the responder. CHF 10 is obviously the fairest offer, because it splits the stake size equally, whereas CHF 4 is the most unfair offer. If the responder accepted, each player received the amount the proposer suggested. If the responder rejected, neither player received any money. The offers that were presented to the responders had in fact been made by the proposers in behavioral pilot experiments. After the pilot experiments, we asked these proposers whether we were allowed to use their offers again in subsequent experiments. These proposers also received the money that resulted from the subsequent use of their offers in the EEG experiment. During the EEG experiment, the responders were located in a room at the University Hospital of Psychiatry, University of Zurich, and they were led to believe that the proposers were making their decisions in the experimental laboratory of the Institute for Empirical Research (IER) in Economics, University of Zurich. To lend credibility to this matching process, we used subjects who had already participated in other experiments at the IER. We implemented the random payment method in our experiment; 6 of the 12 trials were randomly selected for payment.

Fairness judgments

At the end of the ultimatum game, responders were shown a list of all possible offers and were 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*).

EEG recording and processing

Fifty-eight electrodes were placed following the 10-10 montage covering the entire scalp, as recommended by the International Federation of Clinical Neurophysiology Standards for Digital Recording of Clinical EEG (Nuwer et al., 1998). The electrode at Cz (vertex of the head) was used as recording reference. Horizontal and vertical eye movements were recorded with electrodes at the left and right outer canthi and left infra-orbital. Impedances were kept below 10 k Ω . The signals were amplified (bandpass 0.5–125 Hz) and digitized (256 samples/s) using a 64-channel EEG/event-related potential system.

After artifact detection, all available artifact-free 2-s EEG epochs (on average, 70.0 ± 8.3 epochs available per subjects) were recomputed against the average reference and subjected to conventional spectral analyses (see Methodological Details in the Supplemental Material available on-line). Power spectra

were integrated for the following seven independent frequency bands (Kubicki, Herrmann, Fichte, & Freund, 1979): delta (1.5–6 Hz), theta (6.5–8 Hz), alpha1 (8.5–10 Hz), alpha2 (10.5–12 Hz), beta1 (12.5–18 Hz), beta2 (18.5–21 Hz), and beta3 (21.5–30 Hz).

In the next step, standardized low-resolution brain electromagnetic tomography (sLORETA; Pascual-Marqui, 2002) was used to estimate the intracerebral electrical sources that generated the scalp-recorded activity in each of the seven frequency bands (see Methodological Details in the Supplemental Material available on-line).

Statistical analyses

The main goal of this study was to assess links between cortical activity and costly punishment behavior. Accordingly, a correlation approach was used for the whole-brain analyses. The localization of global, widespread correlations (cluster analysis) between cortical activity and costly punishment behavior was assessed by applying the exceedance proportion test (Friston et al., 1990), for each frequency band and each offer separately. The nonparametric randomization approach (Nichols & Holmes, 2002) was used for estimating empirical probability distributions and the corresponding corrected (for multiple comparisons) critical probability thresholds.

To formally test whether findings were specific to one hemisphere, laterality analyses were conducted. Homologous contralateral region was identified (by reversing the x -coordinates), and the averaged current density was calculated across all voxels within the identified cluster of interest. Pearson correlation was then computed between the contralateral region and costly punishment behavior. Finally, the two correlations, one for each hemisphere, were compared using Meng's test (Meng, Rosenthal, & Rubin, 1992) for comparing dependent correlation coefficients.

Results

As expected, the acceptance rates varied strongly across offers. Offers of CHF 4 were accepted on average in 48.3% ($\pm 41.1\%$) of the cases, whereas the acceptance rate for offers of CHF 6 was 86.7% ($\pm 27.4\%$), and offers of CHF 8 and 10 were accepted in 100% of the cases. Using sLORETA as a source-localization technique to estimate intracerebral sources underlying scalp-recorded resting EEG, we found that in the alpha1 frequency band (8.5–10 Hz), 41 voxels showed positive significant correlations between current density and acceptance rate for the most unfair offer. As shown in Figure 1, these voxels all fell into one cluster in the right lateral PFC, including Brodmann's areas 44, 45, 46, and 47. The significant positive correlation between current density within the cluster (i.e., averaged current density across voxels within the cluster) and acceptance rate of the most unfair offer, $r(18) = .71$, $p < .05$, accounted for 50% of the variance in the costly punishment behavior (see Fig. 2). Meng's test for dependent correlations confirmed the laterality effect: The correlation between acceptance rate of unfair offers and current density in right lateral PFC was significantly stronger than the correlation between acceptance rate of unfair offers and current density in the homologous left lateral PFC ($Z = 2.24$, $p < .03$).

Moreover, our findings were highly specific to the right lateral PFC because no significant correlations were found in any other EEG frequency bands, and in no other brain region was resting alpha activity correlated with the acceptance rate of unfair offers. We found no significant correlations between the acceptance rates of the other three offers (CHF 6, 8, and 10) and cortical baseline activity in any of the seven frequency bands.

The observed effect also cannot be attributed to individual differences with regard to fairness judgments. After the ultimatum game experiment, we elicited subjects' fairness

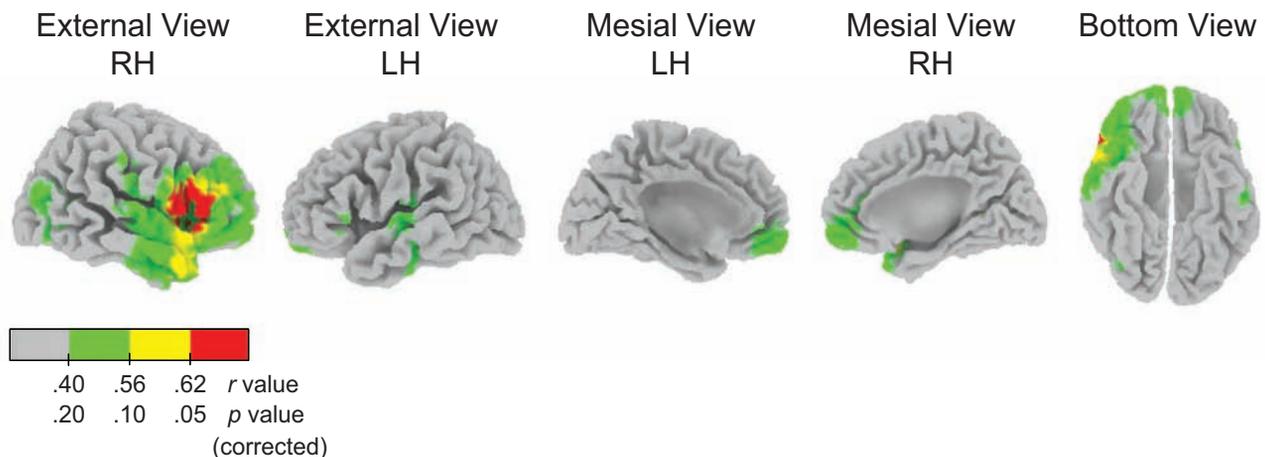


Fig. 1. Regions of the cortex showing source localization of costly punishment behavior. Images of the correlations between the acceptance rates of the most unfair offer (4 Swiss francs) and baseline alpha1 activity (A/m^2) are shown. Results are displayed on the fiducial cortical surface (boundary midway through cortical thickness; Dickson, Drury, & Van Essen, 2001). The cortex is shown in gray scale. RH = right hemisphere; LH = left hemisphere.

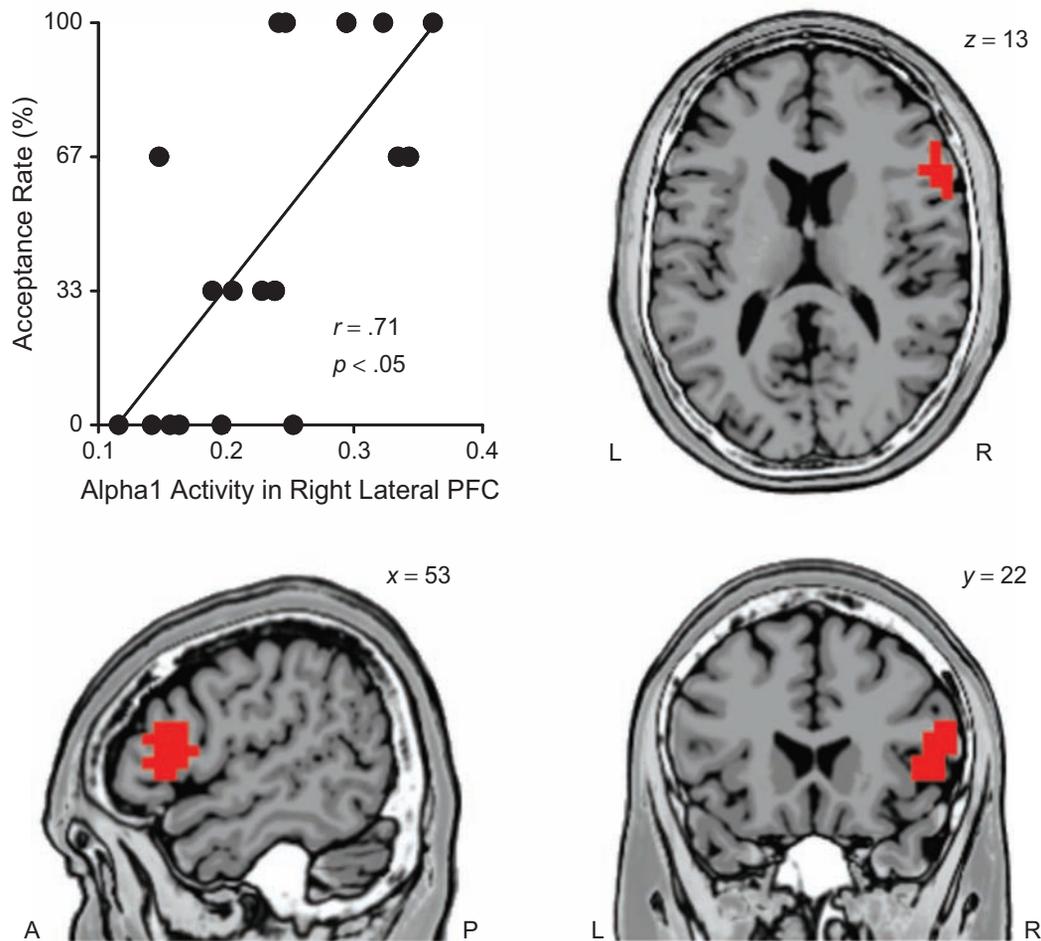


Fig. 2. Acceptance rate of the most unfair offer as a function of baseline alpha1 activity level (A/m^2) in the right lateral prefrontal cortex (PFC) and localizations of the 41 voxels that showed significant correlations between acceptance rate of the most unfair offer and baseline alpha1 activity level. The graph (upper left) shows the significant correlation between percentage of acceptance rates for 4 Swiss francs and alpha1 activity level. Note that resting alpha activity is an inverse indicator of cortical activation. The three images (clockwise from the top) are axial, coronal, and sagittal views of the brain. Locations of the 41 voxels that showed positive significant correlations are indicated in red. Coordinates are in millimeters, transformed into coordinates of the Montreal Neurological Institute atlas, and the origin is the anterior commissure. L = left; R = right; A = anterior; P = posterior.

judgments with regard to different offers. The relation between the cortical activity in the right PFC and the costly punishment behavior remained significant after controlling for fairness judgments, $r(18) = .72$, $p < .05$.

Discussion

We found a positive significant correlation between resting alpha1 current density in the right lateral PFC and acceptance rate for the most unfair offer in the ultimatum game. Because resting alpha activity is an inverse indicator of cortical activation (Shagass, 1972), our results indicate that the right prefrontal baseline hypoactivation predicts heightened acceptance rates of unfair offers. This finding supports the results of brain stimulation studies that applied a disruptive protocol and found an increased acceptance rate of unfair offers compared with sham stimulation or left PFC disruption (Knoch et al,

2006; van 't Wout et al., 2005). In those studies, however, it remained unclear whether the reduction (i.e., the change) of right lateral PFC activation during the task or the lower level of brain activation in this area after transcranial magnetic stimulation was responsible for the behavioral effect. The present findings suggest that the lower level of activation in right lateral PFC is per se responsible for a lower willingness to punish. Moreover, the fact that the relation between the cortical activity in the right PFC and the costly punishment behavior remains significant after controlling for fairness judgments indicates that the level of activation in this area predicts the punitive behavior independently of how fair or unfair the responder perceives a low offer to be.

One possible interpretation is that costly punishment requires self-regulation, as sanctioning behavior conflicts with one's own economic advantage. There is abundant evidence of involvement of the right PFC in self-control from go/no-go

paradigms (Aron, Fletcher, Bullmore, Sahakian, & Robbins, 2003; Garavan, Ross, & Stein, 1999). Furthermore, clinical studies on drug abuse or nonsubstance addictions, in which impairments of decision making seem to reflect a breakdown of self-control processes (Starkstein & Robinson, 1997), support this view. In addition, we were recently able to show that hypoactivity in the right PFC predicts greater individual risk-taking behavior, which is an indicator of lower regulatory abilities (Gianotti et al., 2009). And a recent on-line EEG study revealed less neuronal activation in the right PFC during fast, risky driving in a driving simulator (Jäncke, Brunner, & Esslen, 2008).

Our findings are also congruent with the observation of self-regulation deficits in patients with frontotemporal dementia (FTD), showing that the relative involvement of the right versus the left hemisphere may influence symptoms, with left-sided FTD manifesting itself in language changes, and right-sided FTD being associated with disinhibited social conduct and other socially undesirable behaviors (Mychack, Kramer, Boone, & Miller, 2001). The fact that the variability among individuals with regard to punitive behavior corresponds to different levels of activity in the right PFC might also be interpreted in the sense that hypoactivity in this area may reflect an approach-related behavior (Davidson, 2004). While we suggest that the baseline cortical activity in the right PFC can be a predictor of individuals' punishment behavior, we do not want to imply that this brain area exclusively determines this behavior. Instead, the right PFC might well have very different roles in different functions. A caveat of our study is that participants were all female. A future study on a larger population should provide the opportunity to test for gender differences.

There is increasing evidence that individual differences in EEG power bands are strongly influenced by genetic factors (Zietsch et al., 2007). Alpha band power is highly heritable in the lateral PFC (85%–87%; Anokhin, Heath, & Myers, 2006) and could reflect neural properties that influence decision making in everyday social encounters. Consequently, our findings suggest a plausible mechanism that could explain the relatively high heritability of costly punishment behavior (Wallace et al., 2007). Future studies could investigate whether alpha activity in the lateral PFC might serve as a useful endophenotype for decision-making behavior that requires self-regulation effort.

Acknowledgments

The first two authors contributed equally to this work.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interests with respect to their authorship and/or the publication of this article.

Funding

This research was supported by a grant from the Swiss National Science Foundation (PP00P1-123381) to Daria Knoch.

Supplemental Material

Additional supporting information may be found at <http://pss.sagepub.com/content/by/supplemental-data>

References

- Anokhin, A.P., Heath, A.C., & Myers, E. (2006). Genetic and environmental influences on frontal EEG asymmetry: A twin study. *Biological Psychology, 71*, 289–295.
- Aron, A.R., Fletcher, P.C., Bullmore, E.T., Sahakian, B.J., & Robbins, T.W. (2003). Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. *Nature Neuroscience, 6*, 115–116.
- Camerer, C.F. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton, NJ: Princeton University Press.
- Chapman, L.J., & Chapman, J.P. (1987). The measurement of handedness. *Brain and Cognition, 6*, 175–183.
- Davidson, R.J. (2004). What does the prefrontal cortex “do” in affect: Perspectives on frontal EEG asymmetry research. *Biological Psychology, 67*, 219–233.
- Dickson, J., Drury, H., & Van Essen, D.C. (2001). “The surface management system” (SuMS) database: A surface-based database to aid cortical surface reconstruction, visualization and analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences, 356*, 1277–1292.
- Friston, K.J., Frith, C.D., Liddle, P.F., Dolan, R.J., Lammertsma, A.A., & Frackowiak, K.S. (1990). The relationship between global and local changes in PET scans. *Journal of Cerebral Blood Flow & Metabolism, 10*, 458–466.
- Garavan, H., Ross, T.J., & Stein, E.A. (1999). Right hemispheric dominance of inhibitory control: An event-related functional MRI study. *Proceedings of the National Academy of Sciences, USA, 96*, 8301–8306.
- Gianotti, L.R.R., Knoch, D., Faber, P.L., Lehmann, D., Pascual-Marqui, R.D., Diezi, C., et al. (2009). Tonic activity level in the right prefrontal cortex predicts individuals' risk taking. *Psychological Science, 20*, 33–38.
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., et al. (2005). “Economic man” in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences, 28*, 795–815.
- Herrmann, B., Thoni, C., & Gächter, S. (2008). Antisocial punishment across societies. *Science, 319*, 1362–1367.
- Jäncke, L., Brunner, B., & Esslen, M. (2008). Brain activation during fast driving in a driving simulator: The role of the lateral prefrontal cortex. *NeuroReport, 19*, 1127–1130.
- Knoch, D., Nitsche, M.A., Fischbacher, U., Eisenegger, C., Pascual-Leone, A., & Fehr, E. (2007). Studying the neurobiology of social interaction with transcranial direct current stimulation: The example of punishing unfairness. *Cerebral Cortex, 18*, 1987–1990.
- Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., & Fehr, E. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science, 314*, 829–832.
- Kondacs, A., & Szabo, M. (1999). Long-term intra-individual variability of the background EEG in normals. *Clinical Neurophysiology, 110*, 1708–1716.

- Kubicki, S., Herrmann, W.M., Fichte, K., & Freund, G. (1979). Reflections on the topics: EEG frequency bands and regulation of vigilance. *Pharmacopsychiatry*, *12*, 237–245.
- Meng, X., Rosenthal, R., & Rubin, D.B. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin*, *111*, 172–175.
- Mychack, P., Kramer, J.H., Boone, K.B., & Miller, B.L. (2001). The influence of right frontotemporal dysfunction on social behavior in frontotemporal dementia. *Neurology*, *56*, S11–S15.
- Näpflin, M., Wildi, M., & Sarnthein, J. (2007). Test–retest reliability of resting EEG spectra validates a statistical signature of persons. *Clinical Neurophysiology*, *118*, 2519–2524.
- Nichols, T.E., & Holmes, A.P. (2002). Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping*, *15*, 1–25.
- Nuwer, M.R., Comi, G., Emerson, R., Fuglsang-Frederiksen, A., Guerit, J.M., Hinrichs, H., et al. (1998). IFCN standards for digital recording of clinical EEG. *Electroencephalography and Clinical Neurophysiology*, *106*, 259–261.
- Pascual-Marqui, R.D. (2002). Standardized low-resolution brain electromagnetic tomography (sLORETA): Technical details. *Methods and Findings in Experimental Clinical Pharmacology*, *24*(Suppl. D), 5–12.
- Sanfey, A.G., Rilling, J.K., Aronson, J.A., Nystrom, L.E., & Cohen, J.D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science*, *300*, 1755–1758.
- Shagass, C. (1972). Electrical activity of the brain. In N.S. Greenfield & R.A. Sternbach (Eds.), *Handbook of psychophysiology* (pp. 263–328). New York: Holt, Rinehart and Winston.
- Starkstein, S.E., & Robinson, R.G. (1997). Mechanism of disinhibition after brain lesions. *Journal of Nervous and Mental Disease*, *185*, 108–114.
- Tabibnia, G., Satpute, A.B., & Lieberman, M.D. (2008). The sunny side of fairness: Preference for fairness activates reward circuitry (and disregarding unfairness activates self-control circuitry). *Psychological Science*, *19*, 339–347.
- Tomarken, A.J., Davidson, R.J., Wheeler, R.E., & Kinney, L. (1992). Psychometric properties of resting anterior EEG asymmetry: Temporal stability and internal consistency. *Psychophysiology*, *29*, 576–592.
- van 't Wout, M., Kahn, R.S., Sanfey, A.G., & Aleman, A. (2005). Repetitive transcranial magnetic stimulation over the right dorso-lateral prefrontal cortex affects strategic decision-making. *NeuroReport*, *16*, 1849–1852.
- Wallace, B., Cesarini, D., Lichtenstein, P., & Johannesson, M. (2007). Heritability of ultimatum game responder behavior. *Proceedings of the National Academy of Sciences, USA*, *104*, 15631–15634.
- Zietsch, B.P., Hansen, J.L., Hansell, N.K., Geffen, G.M., Martin, N.G., & Wright, M.J. (2007). Common and specific genetic influences on EEG power bands delta, theta, alpha, and beta. *Biological Psychology*, *75*, 154–164.