European Archives of Psychiatry and Clinical Neuroscience (2019) 269:731–740
https://doi.org/10.1007/s00406-018-0878-2

ORIGINAL PAPER

Behavioral and electrophysiological responses to fairness norm violations in antisocial offenders

Sarah Verena Mayer[1,2] · Karsten Rauss[3] · Gilles Pourtois[4] · Aiste Jusyte[1,5] · Michael Schönenberg[1]

Received: 19 June 2017 / Accepted: 24 January 2018 / Published online: 3 February 2018

Abstract

Antisocial personality disorder is characterized by a stable, lifelong pattern of disregard for and violation of others' rights. Disruptions in the representation of fairness norms may represent a key mechanism in the development and maintenance of this disorder. Here, we investigated fairness norm considerations and reactions to their violations. To examine electrophysiological correlates, we assessed the medial frontal negativity (MFN), an event-related potential previously linked to violations of social expectancy and norms. Incarcerated antisocial violent offenders (AVOs, $n = 25$) and healthy controls (CTLs, $n = 24$) acted as proposers in the dictator game (DG) and ultimatum game (UG) and received fair vs. unfair UG offers from either another human (social context) or a computer (non-social context). Results showed that AVOs made lower offers in the DG but not the UG, indicating more rational and strategic behavior. Most importantly, when acting as recipients in the UG, acceptance rates were modulated by social context in CTLs, while AVOs generally accepted more offers. Correspondingly, ERP data indicated pronounced MFN amplitudes following human offers in CTLs, whereas MFN amplitudes in AVOs were generally reduced. The current data suggest intact fairness norm representations but altered reactions to their violation in antisocial personality disorder.

Keywords Medial frontal negativity · Ultimatum game · Dictator game · Social decision-making · Fairness norms · Antisocial personality disorder

Introduction

Persistent violations of social norms and rights of others are hallmarks of antisocial personality disorder (ASPD) and psychopathy, further characterized by deceitful behavior and a profound lack of remorse and empathy [1, 2]. Disruptions in the representation of fairness norms may represent an important mechanism that contributes to the emergence and maintenance of antisocial behavior. Despite a wealth of studies that delineated fairness behavior and its neural correlates in healthy populations, the link between antisocial behavior and representations of fairness norms is not well understood. Common approaches to experimentally investigate fairness norm considerations and reactions to their violations derive from economics: the dictator game (DG) and the ultimatum game (UG). In the DG [3], the proposer can divide a fixed amount of monetary units (MUs) between him/herself and an anonymous recipient who has no alternative but to accept the offer. Although rational choice would predict zero shares, a broad body of research on healthy proposers shows that the average offer is almost one-third of the

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00406-018-0878-2) contains supplementary material, which is available to authorized users.

✉ Michael Schönenberg
michael.schoenenberg@uni-tuebingen.de

1 Department of Clinical Psychology und Psychotherapy, University of Tübingen, Tübingen, Germany

2 Department of Psychiatry and Psychotherapy, Neurophysiology and Interventional Neuropsychiatry, University of Tübingen, Tübingen, Germany

3 Institute of Medical Psychology and Behavioral Neurobiology, University of Tübingen, Tübingen, Germany

4 Department of Experimental Clinical and Health Psychology, Ghent University, Ghent, Belgium

5 LEAD Graduate School and Research Network, University of Tübingen, Tübingen, Germany

total amount [4]. Research investigating the role of psychopathic traits in DG decision-making indicates an association between higher psychopathy scores and more self-centered behavior/ lower shares in community samples [5, 6] and psychopathic inmates [7]. Thus, this evidence suggests less altruistic behavior in individuals with psychopathic traits.

In the UG [8], the proposer is also given a fixed amount of MUs to share at any rate; however, the recipient can either accept or reject the offer. In case of an acceptance, the money is divided as proposed, in case of a rejection, both players get zero MUs. It is well documented that even moderately unfair offers below 30% are commonly rejected in one of two cases, whereas proposers usually offer a share of 40–50% [5, 8, 9]. These results suggest that other factors than profit-maximization, such as emotional reactions to fairness norm violations, affect decision-making thus overruling rational strategies even at the expense of personal gains. This is further supported by research showing that the variation of context, i.e., social and non-social, is a key factor that determines decision-making. Unfair offers made by another human player (social context) receive higher rejection rates than offers generated by a computer (non-social context) and are accompanied by an increased physiological arousal which presumably reflects negative emotional responses to unfair treatment by another human [10, 11]. Only few studies to date have employed the UG in populations with pronounced psychopathic traits, with the participants acting in the role of the recipient in most studies. The only study to additionally report UG data as proposers shows comparable offers between primary psychopaths and healthy controls [7]. Studies on UG gambling behavior in the role of the recipient yielded largely inconsistent findings, with one study indicating no group differences between psychopathic offenders and healthy controls [12] and another pointing to lower acceptance rates in primary psychopaths [7]. In contrast, a previous investigation in a community sample reported increased acceptance rates following unfair offers in individuals with high psychopathic traits [13], which indicates an insensitivity to unfairness in high psychopathy scorers, whereas another community sample study found no group differences [14]. The only prior study that manipulated the context (i.e., compared offers generated by a computer with offers made by another human) reported differences in the behavioral adjustment to social context variables between psychopathic offenders, non-psychopathic offenders, and healthy controls [12].

Evidence from neuroeconomics highlights the crucial involvement of the anterior cingulate cortex (ACC) in decision-making. This region subserves conflict monitoring and regulation which arises when multiple or concurrent responses compete with one another or the goal/ task at hand [15–17]. Increased ACC activity has been shown following intentional, but not unintentional unfair offers [11], which may mirror a conflict between rational, i.e., profit-maximizing and emotional motives, highlighting the role of negative emotions during decision-making. This medial frontal region is densely interconnected with deeper limbic structures [18] and has been repeatedly found to be structurally and functionally altered in psychopathy and ASPD [19–26].

The medial frontal negativity (MFN) is a negative-going event-related potential (ERP) that has its dipole source in the ACC and is thought to be reflective of ACC activity and the motivational significance of an event [27]. It is measured at fronto-central scalp sites as difference wave that results from activation following unfair minus fair offers and peaks between 200 and 400 ms after offer presentation. The MFN is linked to the subjective evaluation of negative outcomes [27–30] and is sensitive to violations of social expectancy and norms, as experienced when confronted with an unfair offer in the UG [31–33]. Boksem and De Cremer [34] were the first to investigate associations between self-reported moral standards and MFN activity in healthy individuals who received fair vs. unfair offers in classical UG scenarios. According to their results, MFN amplitudes were more pronounced following unfair offers as compared to fair offers, and this effect was most prominent among individuals with high normative standards, and thus reflective of fairness norm violations.

Here, we investigated for the first time fairness norm considerations and reactions to their violations, as well as associated neural correlates in ASPD. For this purpose, antisocial violent offenders (AVOs) and healthy controls (CTLs) acted as proposers both in the DG and the UG. Subsequently, we presented them with fair vs. unfair offers in a series of UG scenarios and measured offer-locked ERPs. Moreover, we manipulated the type of proposer to investigate the influence of social context on decision-making.

Based on the results of previous studies, we expected AVOs to behave more rationally and profit-oriented compared to CTLs. For the DG, this should be reflected in an overall lower offer rate in AVOs compared to CTLs. For the UG in the role of the proposer, however, we expected comparable offers from both AVOs and CTLs, reflecting intact fairness norm representations. In the role of the recipient (UG), we expected overall higher acceptance rates for AVOs compared to CTLs. Moreover, we expected AVOs to accept more offers irrespective of proposer type, whereas CTLs should show higher rejection rates for unfair offers by another human, as we expected them to be more sensitive to social context information.

Finally, we aimed to extend earlier behavioral findings on decision-making in antisocial populations by linking reactions to unfairness with amplitude variations of the MFN. Based on the reported association between fairness norm violations and more pronounced MFN amplitudes, we expected reduced MFN amplitudes in AVOs compared to

CTLs. Additionally, we expected social context to modulate the MFN, i.e., attenuated MFN amplitudes following unfair computer offers as compared to unfair human offers should be observable.

Methods and materials

Participants and measures

Twenty-six males incarcerated for violent offenses in a German correctional facility (Justizvollzugsanstalt Adelsheim) were recruited through advertisement on the black board within the facility (see Supplementary Table 1 for detailed information on convictions and prison sentences). Twenty-five healthy, age-matched male controls were recruited from a vocational school. Inclusion criteria were 18–25 years of age, no current psychiatric illness or criminal record (controls). Controls were tested in the laboratory of the department of psychology, AVOs were assessed in designated rooms of the correctional facility. Trained psychologists from our research group carried out all assessments; all participants gave written informed consent and received monetary compensation. The study protocol was approved by the local ethics committee and was conducted in accordance with the Declaration of Helsinki.

The ability for deductive reasoning and problem solving was measured via the *Wiener Matrizen Test 2* (WMT 2) [35] to control for IQ. Core psychopathic traits were assessed with the *Youth Psychopathic Traits Inventory* (YPI) [36, 37]. Aggressive behavior was measured via the *Buss-Perry Aggression Questionnaire* (BP) [38]. Categorical diagnosis of Axis-I psychopathology and ASPD were assessed by trained psychologists using the *Mini International Neuropsychiatric Interview 6.0.0* [MINI; 39, 40].

Procedure

Dictator game

After completing the IQ measure, questionnaires, and the psychiatric interview, participants were introduced to the DG. They were informed that they would play a game with individuals who had already undergone the experiment and played a classical DG with each player. Furthermore, they were informed that the MUs they kept would be converted to real money and added to their reimbursement. On each trial, a model face was presented for 500 ms and participants were asked to divide 10 MUs between themselves and the respective player (between 1 and 10 MUs), after which they received feedback information on the screen (e.g., “You get 5 points; the other player gets 5 points”). The participants underwent a total of 72 DG trials, with each model identity serving as the recipient once.

Ultimatum game

Following the DG, participants played one UG in the role of the proposer and were instructed to split 10 MUs between themselves and a hypothetical other player. This one-shot UG trial served to control for sufficient understanding of fairness norms. Subsequently, participants played 144 UG trials in the role of the recipient. They were instructed that some offers would be made by other players who previously participated in the experiment, while other offers would be generated randomly by a computer. Thus, participants received offers from other humans (face condition, 72 trials) and offers from a computer (computer condition, 72 trials). In each condition, half of the trials were fair (i.e., 36 × 5 MUs were presented) and the other half varied regarding the degree of unfairness (i.e., 9 × 4 MUs, 9 × 3 MUs, 9 × 2 MUs, and 9 × 1 MUs). Trial types (degree of fairness) and conditions (face vs. computer) were presented in randomized order. To familiarize participants with the task, each of them performed four practice trials (two computer and two cartoon-faced trials) prior to the main experiment. Participants were reminded that earned MUs would be added to their reimbursment at the end of the experiment. Each trial started with the presentation of a fixation cross in the middle of the computer screen for 500 ms (see Fig. 1) followed by a picture of a computer/ one of the 72 faces for 2000 ms. Next, the corresponding offer appeared for 3000 ms under the picture before the “accept" and "reject” options were presented until response. At the end of each trial, feedback regarding the division of the MUs was provided for 1500 ms followed by a 1000 ms inter-trial interval. The final screen contained information about participants’ overall winnings which they received on top of their reimbursement.

Stimuli and software

Pictures of 72 male models with neutral facial expressions were selected from the Radboud Faces Database [41] and the Karolinska Directed Emotional Faces [42, 43]. To adjust the pictures of the different databases, faces were matched for size (495 × 619 pixels) and luminance using Adobe Photoshop CS4® (Adobe Systems Inc., San Jose, USA) and presented in random order against a black background on a 15.4-inch WXGA wide TFT LCD laptop. For the presentation of visual stimuli, Presentation Software Version 16.4® (Neurobehavioral Systems, USA) was used. For the 72 computer trials in the UG (computer condition), a schematic picture of a computer was presented (1446 × 900 pixels). For the five practice trials, pictures of two comic characters were used.

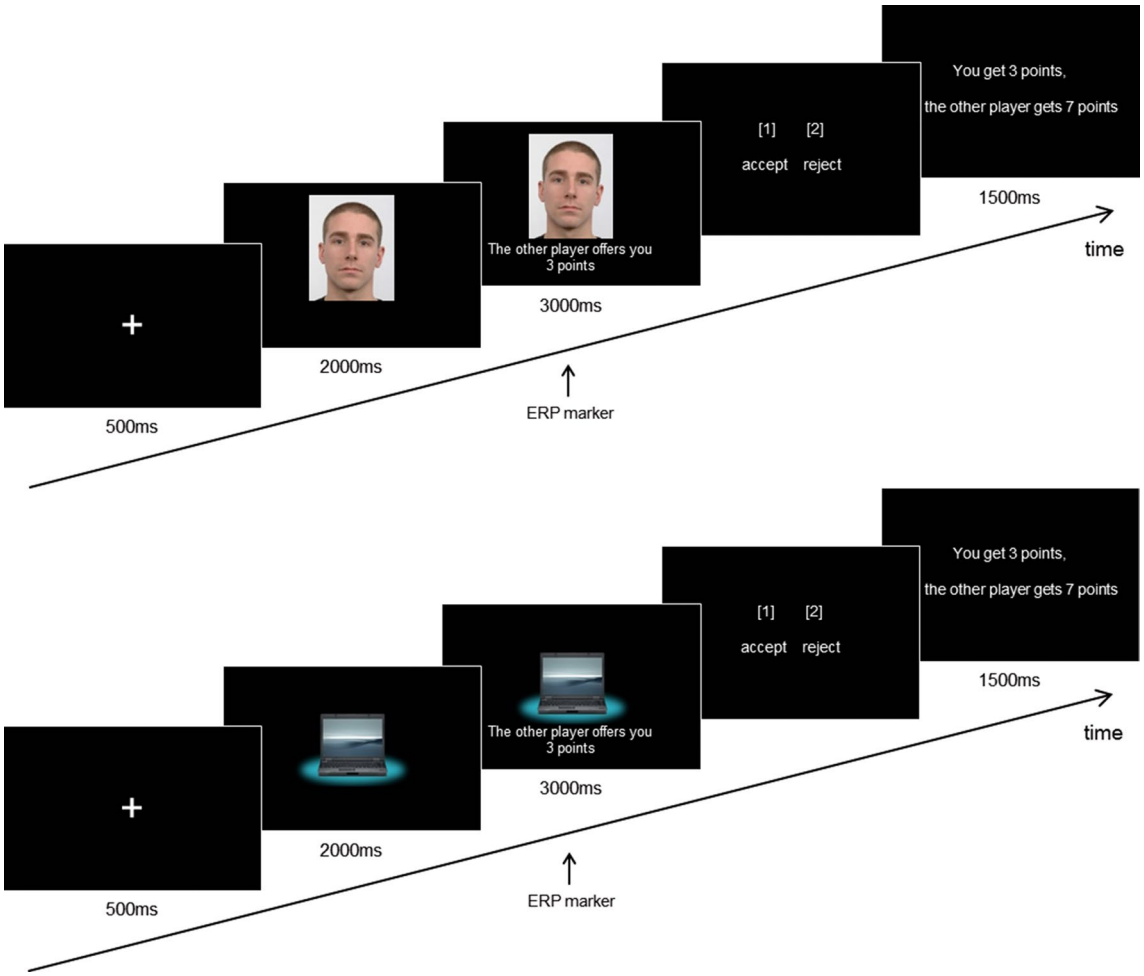

Fig. 1 Exemplary single-shot ultimatum game (UG) trial in which participants act as recipients and can either accept or reject the (unfair) offer coming from another human (upper timeline) or computer (lower timeline). ERPs were locked to offer presentation

Apparatus and electrophysiological recordings

EEG data was recorded from 64 active electrodes mounted in an elastic cap (Brain Products, Munich, Germany) according to the international 10–20 system. One of the 64 electrodes was relocated below the right eye to track vertical eye-movements and blinks. Sampling rate was set to 500 Hz, data was referenced to FCz and re-referenced offline to the average of all electrodes. Impedance levels were kept below 5 kΩ. EEG signals were both low-pass filtered with a 40 Hz cut-off and high-pass filtered with 0.01 Hz cut-off (roll-offs of 12 and 24 dB/oct, respectively). Eye movement artifacts were corrected using semiautomatic Independent Component Analysis [ICA; 44]. Data were then semi-automatically inspected and artefacts rejected based on the following criteria: maximum allowed voltage steps of 50 µV/ms, maximal allowed absolute difference of 200 µV in 200 ms, maximal/minimal allowed amplitude of ± 200 µV, and lowest variability of activity of 0.5 µV in 100 ms. Epochs of 1000 ms (including a 200-ms pre-stimulus baseline) were extracted and baseline-corrected (from − 200 to 0 ms) before we created averaged waveforms for each condition. All analyses were carried out using Brain Vision Analyzer 2.0.4 Software. Statistical analyses were conducted using PASW Statistics 21 software for Windows (SPSS Inc., Chicago, IL, USA).

Results

Participant characteristics

All AVOs fulfilled diagnostic criteria for ASPD and scored significantly higher on psychopathy and aggression scales than CTLs (Table 1). Two participants (one AVO and one CTL) were excluded because their acceptance rates in the fair condition of a standard, neutral UG scenario averaged three standard deviations below the mean of the respective

Table 1 Demographic and diagnostic description

	AVOs ($n=25$)	CTLs ($n=24$)	Statistics	
Age	19.68 (1.07)	19.58 (1.50)	$t_{47}=-0.26$; $p=0.796$	$\eta^2=0.00$
WMT 2 sum score	6.64 (2.53)	8.75 (3.18)	$t_{47}=2.58$; $p=0.013$	$\eta^2=0.12$
YPI				
Grandiose-Manipulative	42.36 (10.03)	39.38 (10.39)	$t_{47}=-1.02$; $p=0.311$	$\eta^2=0.02$
Callous-Unemotional	36.84 (7.96)	31.62 (5.69)	$t_{47}=-2.63$; $p=0.012$	$\eta^2=0.13$
Impulsive-Irresponsible	41.88 (8.38)	34.25 (6.80)	$t_{47}=-3.49$; $p=0.001$	$\eta^2=0.21$
Total sum score	121.08 (21.86)	105.25 (18.45)	$t_{47}=-2.73$; $p=0.009$	$\eta^2=0.14$
BPAQ				
Physical Aggression	32.88 (4.76)	21.04 (6.77)	$t_{47}=-7.11$; $p=0.000$	$\eta^2=0.52$
Verbal Aggression	18.24 (3.80)	15.38 (2.99)	$t_{47}=-2.93$; $p=0.005$	$\eta^2=0.15$
Anger	17.76 (4.67)	13.13 (4.42)	$t_{47}=-3.57$; $p=0.001$	$\eta^2=0.21$
Hostility	25.68 (5.54)	21.21 (5.52)	$t_{47}=-2.83$; $p=0.007$	$\eta^2=0.15$
Total sum score	94.56 (12.85)	70.75 (15.46)	$t_{47}=-5.87$; $p=0.000$	$\eta^2=0.42$

Illustrated are mean values and standard deviations in parenthesis

WMT 2 Wiener Matrizen Test 2, *YPI* Youth Psychopathic Traits Inventory, *BPAQ* Buss–Perry Aggression Questionnaire, *AVOs* antisocial violent offenders, *CTLs* healthy controls

group thereby indicating poor task understanding. EEG data of five AVO and six CTLs had to be excluded due to excessive artefacts or technical difficulties. The final sample thus consisted of 25 AVOs and 24 CTLs for the behavioral data, and EEG recordings of 20 AVOs and 18 CTLs.

Both groups were comparable with respect to age, but differed in terms of intelligence levels. However, WMT 2 sum scores were not related to offer rates or acceptance rates in the DG or UG (see Supplementary Table 2).

Behavioral data

For the dictator game, we compared offers between groups via *t* tests for independent samples. On average, AVOs ($M=2.36$; SD = 1.32) offered significantly less MUs in the DG as compared to CTLs ($M=3.63$; SD = 1.22; $t_{47}=3.49$, $p=0.001$; Fig. 2). While 48.83% of AVOs decided on giving the minimum of 1 MU, only 18.23% of CTLs gave 1 MU. For the ultimatum game in the role of the proposer, we compared UG offers between groups via *t*-tests and found that AVOs and CTLs allocated a comparable amount of MUs ($M=4.48$; SD = 1.73 vs. $M=5.05$; SD = 1.20; $t_{44}=1.26$, $p=0.213$; Fig. 2).

The ultimatum game trials in the role of the recipient were analyzed using a 2 (within-subjects factor proposer: face/computer) × 2 (within-factor fairness: fair/unfair) × 2 (between-subjects factor group: AVOs/CTLs) general linear model (GLM) for repeated measures, with acceptance rates as dependent variable. Analysis revealed a significant main effect for proposer ($F_{1,47}=5.91$; $p=0.019$), with higher acceptance rates for computer relative to face offers, a significant main effect for fairness ($F_{1,47}=160.93$; $p<0.001$), with higher acceptance rates

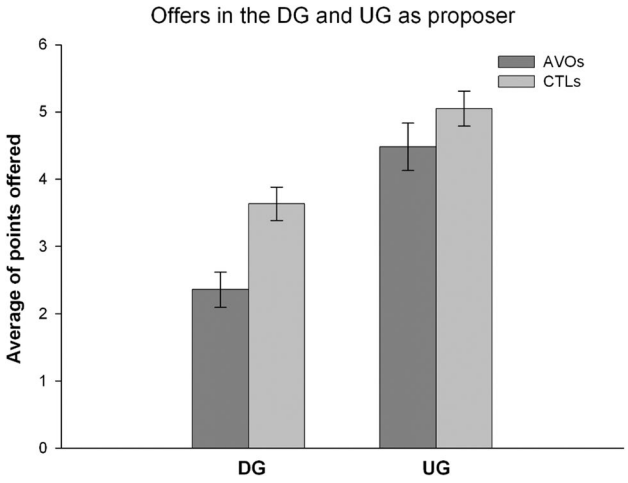

Fig. 2 Offers in the dictator game (DG) and ultimatum game (UG) as proposer for both groups (vertical bars represent standard errors of mean). AVOs gave significantly less than CTLs in the DG ($p=0.001$), but made comparable offers in the UG ($p=0.213$). *AVOs* antisocial violent offenders, *CTLs* healthy controls

for fair as compared to unfair offers, and a significant proposer x group interaction ($F_{1,47}=4.96$; $p=0.031$; Fig. 3), with lower acceptance rates in CTLs in the face condition. Post-hoc paired *t* tests revealed significant differences between the overall acceptance rates between the face and the computer condition for CTLs ($t_{23}=2.60$, $p=0.016$), but not for AVOs ($t_{24}=0.22$, $p=0.828$). We did not find a main effect for group ($F_{1,47}=1.40$; $p=0.243$), a fairness x group interaction ($F_{1,47}=0.64$; $p=0.428$), a proposer × fairness interaction ($F_{1,47}=0.30$; $p=0.585$), or a proposer × fairness × group interaction ($F_{1,47}=0.27$; $p=0.609$).

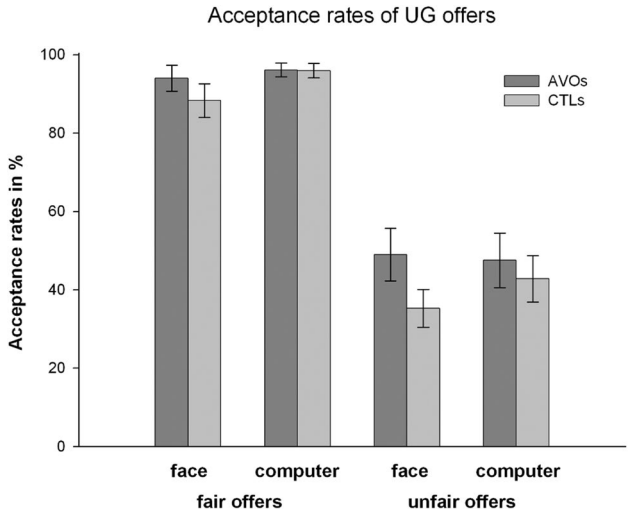

Fig. 3 Acceptance rates for fair and unfair offers (5 points vs. 4, 3, 2, and 1 points) in the face and computer condition in the ultimatum game (UG) as recipient (vertical bars represent standard errors of mean). We found significant main effects for proposer ($p = 0.019$) and fairness level ($p < 0.001$), and a proposer × group interaction ($p = 0.031$) in the absence of any other significant effects. *AVOs* antisocial violent offenders, *CTLs* healthy controls

Correlation analyses between overall acceptance rates and psychopathy or aggression scores also did not yield any significant results (see Supplementary Table 3).

Electrophysiological data

To identify the MFN, we subtracted grand-average data obtained in CTLs in the fair-face condition from those of the unfair-face condition. The resulting difference waves revealed more negative voltages in the unfair-face condition at fronto-central electrodes between 240 and 280 ms. Based on previous findings [28–30, 32], we interpreted this as the MFN in our dataset. We thus analyzed MFN amplitudes by extracting individual mean activity between 240 and 280 ms, averaged over electrodes FC1, FCz, and FC2 [28].

A GLM with proposer (face/computer) and fairness (fair/unfair) as within-subjects factors, and group (CTLs/AVOs) as between-subjects factor showed a significant fairness x group interaction ($F_{1, 36} = 5.73$, $p = 0.022$), in the absence of any other main ($F_{1, 36} = 1.79$, $p = 0.190$ for proposer; $F_{1, 36} = 0.05$, $p = 0.818$ for fairness; and $F_{1, 36} = 0.13$, $p = 0.722$ for group) or interaction effects ($F_{1, 36} = 0.11$, $p = 0.748$ for proposer × group; $F_{1, 36} = 1.76$, $p = 0.193$ for proposer × fairness; and $F_{1, 36} = 0.39$, $p = 0.534$ for proposer × fairness × group; see Fig. 4). Post-hoc tests indicated that the fairness × group interaction was driven by opposite effects of fairness in the two groups: in CTLs, more negative voltages were observed for unfair offers (-0.34 ± 0.22 µV), whereas more positive voltages were seen in AVOs ($+0.28 \pm 0.14$ µV). Taken by themselves, these effects were only marginally significant (CTLs, $t_{17} = -1.53$, $p = 0.073$ one-tailed; AVOs, $t_{19} = 1.96$, $p = 0.065$), and there were no group differences when fair ($t_{36} = 0.23$, $p = 0.822$) and unfair offers ($t_{36} = -0.883$, $p = 0.383$) were considered separately.

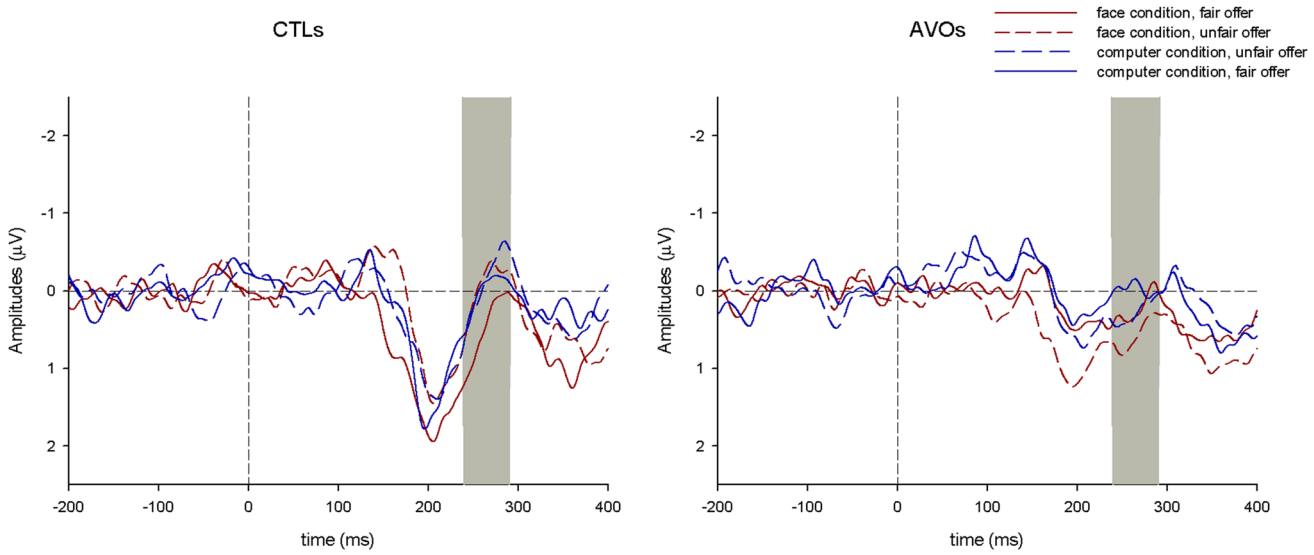

Fig. 4 Medial frontal negativity amplitudes (240–280 ms, grey) from pooled electrodes FC1, FCz, and FC2 elicited for fair and unfair offers in the face and computer condition in the ultimatum game (UG) as recipient separately for CTLs (left column) and AVOs (right column). We found a significant fairness x group interaction ($p = 0.022$), all other effects were non- significant. *AVOs* antisocial violent offenders, *CTLs* healthy controls

In contrast to what we observed in the behavioral data, the interaction between proposer and group was non-significant. Due to this difference between the two sources of data, we nevertheless conducted exploratory analyses within conditions. This was done to clarify whether this result indicates a dissociation between fairness processing at the level of the MFN and its behavioral expression in terms of acceptance rates. Results indicated a significant fairness x group interaction only for human proposers ($F_{1, 36} = 4.92$, $p = 0.033$), but not in the PC condition ($F_{1, 36} = 0.95$, $p = 0.336$). Thus, while the electrophysiological data do not support the notion of group differences in proposer-specific fairness processing, some information on both the source and valence of a particular offer seems to be retained at the level of the MFN.

Additionally, we correlated MFN amplitudes and gambling behavior in the UG as recipient, however, we did not find any correlations either for AVOs ($p = 0.342$ for the face condition, $p = 0.709$ for the computer condition) or for CTLs ($p = 0.872$ for the face condition, $p = 0.955$ for the computer condition).

While our main focus was on the MFN, Fig. 4 suggests earlier differences during the P200 interval [45, 46]. To address this early difference, mean amplitude values were extracted between 180 and 220 ms and averaged across the set of electrodes used for the analysis of the MFN (i.e., FC1, FCz, and FC2). Using the same GLM approach as for the MFN data, we observed a marginally significant group difference ($F_{1, 36} = 3.24$, $p = 0.080$, with higher P200 amplitudes observed in CTLs), and a significant fairness x group interaction ($F_{1, 36} = 7.64$, $p = 0.009$). No other main or interaction effects were significant (proposer, $F_{1, 36} = 0.98$, $p = 0.33$; fairness, $F_{1, 36} < 0.01$, $p = 0.954$; proposer × group, $F_{1, 36} = 0.37$, $p = 0.55$; proposer × fairness, $F_{1, 36} = 0.02$, $p = 0.880$; proposer × fairness x group, $F_{1, 36} = 1.07$, $p = 0.308$). Post-hoc analysis indicated that the fairness x group interaction was again due to opposite effects of fairness in the two groups, with a significant difference for CTLs ($+0.45 \pm 0.20$ µV, $t_{17} = -2.21$, $p = 0.041$), and a marginal effect for AVOs (-0.47 ± 0.26 µV, $t_{19} = -1.83$, $p = 0.083$). In addition, when considering fair and unfair offers separately, we observed a significant group difference only for the fair condition (CTLs, 1.55 ± 1.23 µV; AVOs, 0.37 ± 1.33 µV; $t_{36} = 2.84$, $p = 0.007$; unfair: $p > 0.57$). This suggests that earlier group differences in neural responses to fair offers may set the stage for differential processing of unfair offers at the level of the MFN.

Discussion

The aim of the present study was to explore fairness norm considerations, reactions to fairness norm violations, and underlying electrophysiological correlates in individuals who persistently violate social norms relative to healthy controls. The main findings can be summarized as follows: First, AVOs shared less MUs in the DG as compared to CTLs; however, when acting as proposer in the UG, shares between groups were comparable and reached almost 50%. These findings indicate intact fairness norm representations in AVOs as they were able to adjust their behavior to follow a given norm if it corresponded to their personal interests. Second, when acting as recipients in the UG, we found gambling behavior to be modulated by social context in CTLs, but not in AVOs, who generally accepted more offers irrespective of the proposer type. This provides further evidence for a more rational and profit-maximizing behavior in AVOs compared to CTLs. Third, ERP findings indicated attenuated MFN amplitudes in AVOs irrespective of social context, while in CTLs, MFN amplitudes were more pronounced, i.e., more negative deflections for unfair compared to fair offers. Exploratory analyses indicated that this effect was modulated by proposer type, i.e., unfair offers made by human proposers resulted in the largest MFN deflections in healthy controls. Taken together, the data provide first evidence for impairments in the processing of social norm violations on both behavioral and neural levels in ASPD.

Behavioral results

When acting as proposers in the DG and UG, AVOs exhibited more strategic, and thus profit-maximizing behaviors compared to CTLs. This indicates that AVOs know about fairness norms and their importance to others, and that they are capable of complying to a given norm, at least when it serves their purposes. These findings are consistent with previous research suggesting an association between certain psychopathic traits and more strategic behavior in gambling paradigms [5, 47]. Importantly, our results correspond with a study on primary psychopaths by Koenigs et al. [7], who also reported approximately fair splits when participants acted as proposers in the UG but lower offers in the DG.

The current findings also show that in the role of recipients in the UG, acceptance rates were modulated by social context in CTLs, which is in line with previous findings in healthy individuals, suggesting higher acceptance rates for computer offers relative to human offers [10, 11, 48]. However, this effect was not evident in AVOs. While AVOs generally accepted more offers irrespective of the social context, CLTs rejected human offers more often than computer offers. Based on previous findings [10, 11], this may be indicative of emotional and punishing behavior. Our results may thus indicate that AVOs' decisions were rather rational, as they continued to pursue their strategy of profit maximization irrespective of whether the offer was made by a social or non-social proposer. Only one previous study directly compared differences between social vs. non-social

proposers on gambling behavior in psychopathic offenders and reported differences in the behavioral adjustment to social context variables between psychopathic offenders, non-psychopathic offenders, and healthy controls [12]. Our findings of strategic behavior during different game settings in antisocial individuals are in line with results in healthy individuals indicating that psychopathic traits, such as Machiavellianism [6, 47] and coldheartedness [5], are predictive of less emotional and more rational decision-making behavior. Moreover, psychopathic traits have been associated with self-centered behavior, as individuals high in psychopathic traits pursue strategies that focus on their own short-term winnings rather than on the establishment of functional long-term interactions [13, 24, 47].

Medial frontal negativity

With respect to the MFN, we were able to replicate previous results of more negative deflections following unfair as compared to fair offers in the UG [34, 49, 50] and extended existing findings by manipulating the degree of social context: In CTLs, exploratory analyses indicated that more pronounced MFN amplitudes following unfair as compared to fair offers were only evident following the human, but not computer trials. These results contradict the idea that the MFN is merely reflective of the final outcome evaluation on a good-bad dimension [51, 52]. If this were true, the MFN should also be present following computer offers, since outcomes were numerically identical irrespective of proposer type. Our results are better explained in terms of reinforcement learning based on expectancy violation which predicts increased ACC activity when outcomes are worse than expected [53]. Boksem and De Cremer [34] showed more pronounced MFN amplitudes following unfair as compared to fair UG offers and further extended this theory to the social nature of expectancy violations. The association between moral standards and MFN amplitudes indicate that the violation of subjective fairness norms leads to a negative emotional response, i.e., "social pain", which in turn is reflected in the MFN as consequence of increased ACC activity following unfair treatment by another human player [11, 14]. Accordingly, in our study, it is likely that the presentation of a face generated the expectation of a human player who follows social norms, and thus behaves fairly (which corresponds to a fair offer in the UG, see Chang and Sanfey [54]). Corresponding to this model and in line with our results, MFN amplitudes increased following unfair offers, i.e., when expectations were not met and fairness norms were violated. On the contrary, when participants assumed that an offer was randomly generated by a computer, no adherence to fairness norms was expected and MFN amplitudes were attenuated. Since MFN amplitudes are thought to be the reflective of ACC activity [27, 55] and the ACC is involved in the processing of the affective aversive components of both physical and social pain [15, 17, 56, 57], we believe this to be the most parsimonious explanation of the present ERP results. Accordingly, Sanfey et al. [11] found increased ACC activations following unfair as compared to fair offers, and this effect was more prominent for proposals made by humans as compared to computer.

Most importantly, AVOs did not show increased MFN amplitudes following unfair offers following either human or computer offers. One possible explanation is that they perceived the scenarios simply as business transactions, thereby expecting low offers and lacking the experience of social pain following unfair offers. Our behavioral findings, namely low offer rates in AVOs as proposers in the DG and higher acceptance rates as recipients regardless of the proposer type in the UG, support this assumption. In addition, previous evidence from an community sample indicates that high psychopathy scorers perceived unfair offers as less unfair although rejection rates were comparable [14]. Alternatively, the absence of electrophysiological correlates in AVOs could reflect a diminished emotional reactivity to unfairness. Osumi and Ohira [13] found higher acceptance rates and attenuated skin conductance responses following unfair offers in students high in psychopathic traits indicative of diminished emotional reactions. This insensitivity to unfairness resulted in more rational and strategic behaviors in individuals with high psychopathic traits. This neuronal pattern is thus consistent with our behavioral results and the callous-unemotional nature of ASPD and psychopathy and is further corroborated by evidence on dysfunctional ACC activity in these individuals [see 25, 26].

Limitations

Although our study was carefully planned and executed, several limitations have to be mentioned. For one, we only investigated male violent offenders and future studies should also include female samples to examine whether the effects reported here are gender specific. Further, in the current study, we did not employ a direct assessment of emotional reactions to unfairness. Thus, interpretations of the present findings in terms of emotionality remain speculative. Future studies should include more explicit measures of emotional responses to unfairness. In addition, due to practical constraints, we did not acquire high-resolution eye-movement data. Differential patterns of eye-movements in response to the offer text could thus have affected our EEG data. However, we note that in order to explain our MFN results, such differences would have to be specific for a particular group x fairness condition. Nevertheless, future studies should employ state-of-the-art eye-tracking methods to address this potential confound. Finally, the behavioral data

in the present study indicate differences between groups with regard to the role of social context, while we did not find an interaction with social context for the electrophysiological data. Although exploratory analyses of the MFN amplitudes indicated a similar pattern, these findings must be interpreted with caution, and the conclusions regarding the interplay between social context and fairness at the level of the MFN remain speculative. Additional studies in larger samples will be required to unravel the neural underpinnings of how proposer and fairness information influences behavioral choices in healthy subjects and antisocial offenders.

Conclusion

In sum, the present results suggest that antisocial behavior is associated with more rational and self-centered gambling behavior which is underpinned by electrophysiological data. However, additional studies are needed to clarify the role of emotions in decision-making in ASPD which were not assessed in the present study. Further studies should replicate and further corroborate these findings and investigate whether disruptions in neural systems are causally related to the profound lack of empathic responding and persistent hurtful, deceitful, and antisocial behavior in psychopathy and ASPD.

Acknowledgement The authors would like to thank Elisabeth Künzel and Angelika Bertsche for their support in data collection. Moreover, we would like to thank the staff of the JVA Adelsheim, especially Dr. Wolfgang Stelly, for their support in implementing the study.

Funding S. V. M. was supported by the Postgraduate Research Grants Program of Baden-Württemberg and subsequently by the FAZIT foundation. A. J. was supported by the Promotion of Junior Researchers Program at the University of Tübingen and the LEAD Graduate School [GSC1028], a project of the Excellence Initiative of the German federal and state governments. Aside from financial support, no further contributions were made by the funders.

Compliance with ethical standards

Conflict of interest None of the authors has any conflict of interest.

References

1. American Psychiatric Association (2000). Diagnostic and statistical manual of mental disorders, textrevision (DSM-IV-TR). American Psychiatric Association, Washington, DC
2. Cleckley HM (1941) The mask of sanity: an attempt to clarify some issues about the so called psychopathic personality. Aware Journalism
3. Kahneman D, Knetsch JL, Thaler RH (1986) Fairness and the assumptions of economics. J Bus 59:285–300
4. Engel C (2011) Dictator games: a meta study. Exp Econ 14:583–610
5. Berg JM, Lilienfeld SO, Waldman ID (2013) Bargaining with the devil: Using economic decision-making tasks to examine the heterogeneity of psychopathic traits. J Res Pers 47:472–482
6. Spitzer M, Fischbacher U, Herrnberger B, Grön G, Fehr E (2007) The neural signature of social norm compliance. Neuron 56:185–196
7. Koenigs M, Kruepke M, Newman J (2010) Economic decision-making in psychopathy: a comparison with ventromedial prefrontal lesion patients. Neuropsychologia 48:2198–2204
8. Güth W, Schmittberger R, Schwarze B (1982) An experimental analysis of ultimatum bargaining. J Econ Behav Organ 3:367–388
9. Bolton GE, Zwick R (1995) Anonymity versus punishment in ultimatum bargaining. Games Econ Behav 10:95–121
10. Van't Wout M, Kahn RS, Sanfey AG, Aleman A (2006) Affective state and decision- making in the ultimatum game. Exp Brain Res 169:564–568
11. Sanfey AG, Rilling JK, Aronson JA, Nystrom LE, Cohen JD (2003) The neural basis of economic decision-making in the ultimatum game. Science 300:1755–1758
12. Radke S, Brazil IA, Scheper I, Bulten BH, De Bruijn ER (2013) Unfair offers, unfair offenders? Fairness considerations in incarcerated individuals with and without psychopathy. Front Hum Neurosci 7:406
13. Osumi T, Ohira H (2010) The positive side of psychopathy: Emotional detachment in psychopathy and rational decision-making in the ultimatum game. Pers Individ Differ 49:451–456
14. Vieira JB, Almeida PR, Ferreira-Santos F, Barbosa F, Marques-Teixeira J, Marsh AA (2013) Distinct neural activation patterns underlie economic decisions in high and low psychopathy scorers. Soc Cogn Affect Neurosci 9:1099–1107
15. Eisenberger NI, Lieberman MD, Williams KD (2003) Does rejection hurt? An fMRI study of social exclusion. Science 302:290–292
16. Bush G, Luu P, Posner MI (2000) Cognitive and emotional influences in anterior cingulate cortex. Trends Cogn Sci 4:215–222
17. Shackman AJ, Salomons TV, Slagter HA, Fox AS, Winter JJ, Davidson RJ (2011) The integration of negative affect, pain and cognitive control in the cingulate cortex. Nat Rev Neurosci 12:154–167
18. Ghashghaei H, Hilgetag C, Barbas H (2007) Sequence of information processing for emotions based on the anatomic dialogue between prefrontal cortex and amygdala. Neuroimage 34:905–923
19. Glenn AL, Yang Y, Raine A, Colletti P (2010) No volumetric differences in the anterior cingulate of psychopathic individuals. Psychiatry Res Neuroimaging 183:140–143
20. Kiehl KA (2006) A cognitive neuroscience perspective on psychopathy: evidence for paralimbic system dysfunction. Psychiatry Res 142:107–128
21. Kiehl KA, Smith AM, Hare RD, Mendrek A, Forster BB, Brink J, Liddle PF (2001) Limbic abnormalities in affective processing by criminal psychopaths as revealed by functional magnetic resonance imaging. Biol Psychiatry 50:677–684
22. Glenn AL, Raine A, Schug RA (2009) The neural correlates of moral decision-making in psychopathy. Mol Psychiatry 14:5–6
23. Ly M, Motzkin JC, Philippi CL, Kirk GR, Newman JP, Kiehl KA, Koenigs M (2012) Cortical thinning in psychopathy. Am J Psychiatry 169:743–749
24. Rilling JK, Glenn AL, Jairam MR, Pagnoni G, Goldsmith DR, Elfenbein HA, Lilienfeld SO (2007) Neural correlates of social cooperation and non-cooperation as a function of psychopathy. Biol Psychiatry 61:1260–1271
25. Koenigs M (2012) The role of prefrontal cortex in psychopathy. Rev Neurosci 23:253–262
26. Yang Y, Raine A (2009) Prefrontal structural and functional brain imaging findings in antisocial, violent, and psychopathic individuals: a meta-analysis. Psychiatry Res Neuroimaging 174:81–88

27. Gehring WJ, Willoughby AR (2002) The medial frontal cortex and the rapid processing of monetary gains and losses. Science 295:2279–2282
28. Fukushima H, Hiraki K (2006) Perceiving an opponent's loss: Gender-related differences in the medial-frontal negativity. Soc Cogn Affect Neurosci 1:149–157
29. Alexopoulos J, Pfabigan DM, Göschl F, Bauer H, Fischmeister FPS (2013) Agency matters! Social preferences in the three-person ultimatum game. Front Hum Neurosci 7:312
30. Alexopoulos J, Pfabigan DM, Lamm C, Bauer H, Fischmeister FPS (2012) Do we care about the powerless third? An ERP study of the three-person ultimatum game. Front Hum Neurosci 6:59
31. Wu Y, Leliveld MC, Zhou X (2011) Social distance modulates recipient's fairness consideration in the dictator game: An ERP study. Biol Psychol 88:253–262
32. Wu Y, Zhou Y, van Dijk E, Leliveld MC, Zhou X (2011) Social comparison affects brain responses to fairness in asset division: an ERP study with the ultimatum game. Front Hum Neurosci 5:131
33. Van der Veen FM, Sahibdin PP (2011) Dissociation between medial frontal negativity and cardiac responses in the ultimatum game: effects of offer size and fairness. Cogn Affect Behav Neurosci 11:516–525
34. Boksem MA, De Cremer D (2010) Fairness concerns predict medial frontal negativity amplitude in ultimatum bargaining. Soc Neurosci 5:118–128
35. Formann AK, Waldherr K, Piswanger K (2011) Wiener Matrizen-Test 2. Manual. Beltz Test GmbH, Göttingen
36. Andershed H, Gustafson SB, Kerr M, Stattin H (2002) The usefulness of self-reported psychopathy-like traits in the study of antisocial behaviour among non-referred adolescents. Eur J Pers 16:383–402
37. Köhler DK, Kuska SK, Schmeck K, Hinrichs G, Fegert J (2010) Deutsche Version des Youth-Psychopathic-Traits-Inventory (YPI). In: Barkmann C, Schulte-Markwort M, Brähler E (eds) Klinisch-psychiatrische Ratingskalen für das Kindes- und Jugendalter. Hogrefe, Göttingen, pp 478–482
38. Buss AH, Perry M (1992) The aggression questionnaire. J Pers Soc Psychol 63:452–459
39. Lecrubier Y, Sheehan D, Weiller E, Amorim P, Bonora I, Harnett Sheehan K, Janavs J, Dunbar G (1997) The mini international neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. Eur Psychiatry 12:224–231
40. Ackenheil M, Stotz G, Dietz-Bauer R, Vossen A (1999) Deutsche Fassung des Mini- international neuropsychiatric interview. Psychiatrische Universitätsklinik München, München
41. Langner O, Dotsch R, Bijlstra G, Wigboldus DH, Hawk ST, van Knippenberg A (2010) Presentation and validation of the radboud faces database. Cogn Emot 24:1377–1388
42. Lundqvist, D., Flykt, A., Öhman, A. (1998). The Karolinska Directed EmotionalFaces - KDEF, CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet
43. Goeleven E, De Raedt R, Leyman L, Verschuere B (2008) The Karolinska directed emotional faces: a validation study. Cogn Emo 22:1094–1118
44. Jutten C, Herault J (1991) Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture. Sign Process 24:1–10
45. Dunn BR, Dunn DA, Languis M, Andrews D (1998) The relation of ERP components to complex memory processing. Brain Cogn 36:355–376
46. Noldy NE, Stelmack RM, Campbell KB (1990) Event-related potentials and recognition memory for pictures and words: The effects of intentional and incidental learning. Psychophysiology 27:417–428
47. Curry O, Chesters MJ, Viding E (2011) The psychopath's dilemma: The effects of psychopathic personality traits in one-shot games. Pers Individ Differ 50(6):804–809
48. Radke S, Güroğlu B, De Bruijn ER (2012) There's something about a fair split: intentionality moderates context-based fairness considerations in social decision-making. PloS One 7:e31491
49. Hewig J, Kretschmer N, Trippe RH, Hecht H, Coles MG, Holroyd CB, Miltner WH (2011) Why humans deviate from rational choice. Psychophysiology 48:507–514
50. Moser A, Gaertig C, Ruz M (2014) Social information and personal interests modulate neural activity during economic decision-making. Front Hum Neurosci 8:31
51. Hajcak G, Moser JS, Holroyd CB, Simons RF (2006) The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. Biol Psychol 71:148–154
52. Yeung N, Sanfey AG (2004) Independent coding of reward magnitude and valence in the human brain. J Neurosci 24:6258–6264
53. Holroyd CB, Coles MG (2002) The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. Psychol Rev 109:679–709
54. Chang LJ, Sanfey AG (2009) Unforgettable ultimatums? Expectation violations promote enhanced social memory following economic bargaining. Front Behav Neurosci 3:277–284
55. Miltner WH, Braun CH, Coles MG (1997) Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a "generic" neural system for error detection. J Cogn Neurosci 9:788–798
56. Rainville P, Duncan GH, Price DD, Carrier B, Bushnell MC (1997) Pain affect encoded in human anterior cingulate but not somatosensory cortex. Science 277:968–971
57. Sawamoto N, Honda M, Okada T, Hanakawa T, Kanda M, Fukuyama H, Konishi J, Shibasaki H (2000) Expectation of pain enhances responses to nonpainful somatosensory stimulation in the anterior cingulate cortex and parietal operculum/posterior insula: an event-related functional magnetic resonance imaging study. J Neurosci 20:7438–7445