Neural correlates of social decision-making in severely antisocial adolescents

Wouter van den Bos,[1,2,3] Pauline Vahl,[1,4] Berna Güroğlu,[1,3] Félice van Nunspeet,[1,3] Olivier Colins,[4] Monica Markus,[4] Serge A.R.B. Rombouts,[1,5,6] Nic van der Wee,[1,6] Robert Vermeiren,[1,4] and Eveline A. Crone[1,3,7]

[1]Leiden Institute for Brain and Cognition, Leiden, the Netherlands 2300 RC, [2]Center for Adaptive Rationality (ARC), Max-Planck-Institute for Human Development, Berlin, Germany 14195, [3]Leiden University, Department of Psychology, Leiden, the Netherlands, 2333 AK, [4]Department of Child and Adolescent Psychiatry, Leiden University Medical Centre, Leiden, the Netherlands 2300 RC, [5]Department of Radiology, Leiden University Medical Centre, Leiden, the Netherlands 2300 RC, [6]Department of Psychiatry, Leiden University Medical Centre, Leiden, the Netherlands 2300 RC, and [7]Department of Psychology, University of Amsterdam, Amsterdam, the Netherlands 1018 WS.

Neurobiological and behavioral findings suggest that the development of delinquent behavior is associated with atypical social-affective processing. However, to date, no study has examined neural processes associated with social interactions in severely antisocial adolescents. In this study we investigated the behavioral and neural processes underlying social interactions of juvenile delinquents and a matched control group. Participants played the mini-Ultimatum Game as a responder while in the MRI scanner. Participants rejected unfair offers significantly less when the other player had 'no alternative' compared with a 'fair' alternative, suggesting that they took the intentions of the other player into account. However, this effect was reduced in the juvenile delinquents. The neuroimaging results revealed that juvenile delinquents showed less activation in the temporal parietal junction (TPJ) and inferior frontal gyrus (IFG). However, the groups showed similar activation levels in the dorsal anterior cingulate cortex (dACC) and the right anterior insula (AI) when norms were violated. These results indicate that juvenile delinquents with severe antisocial behavior process norm violations adequately, but may have difficulties with attending spontaneously to relevant features of the social context during interactions.

INTRODUCTION

During adolescence there is a steep increase in antisocial behavior, some studies reporting a peak of 10-fold the base rate of delinquent behavior in children, followed by a decline that starts around emerging adulthood (Moffit, 1993; Moffit and Scott, 2009). Adolescent antisocial behavior generates large costs to society: juvenile delinquents not only inflict serious physical and psychological harm on others, but also greatly increase the risk for negative outcomes for themselves (Loeber et al., 2000, 2009). Furthermore, severely antisocial behavior in adolescence is associated with increases in risk for negative outcomes later in life, such as academic failure, peer rejection and career criminality (Patterson et al., 1989; Maughan and Rutter, 2001).

A number of studies have suggested that severely antisocial behavior is the result of atypical social information processing (Crick and Dodge, 1994; Happe and Frith, 1996). The social information processing (SIP) model suggests that atypical processing of social information, either during encoding or action selection, may lead to aberrant social behavior. One example is the tendency to attribute hostile intentions to ambiguous social cues, which may lead to inappropriately aggressive responses (Dodge and Frame, 1982; Graham et al., 1992). Another set of studies has reported a relation between severely antisocial behavior in adolescence and inferring the affective state of others (Sharp, 2008; Fairchild et al., 2009a; Jones et al., 2010; Schwenck et al., 2011; Sebastian et al., 2012, 2013). In addition, several brain imaging studies on clinical and non-clinical adolescent populations suggest that juvenile antisocial behavior may be related to reduced affective sensitivity or

problems with emotion regulation (Sterzer et al., 2005, 2007; Herpertz et al., 2008; Marsh and Blair, 2008; Passamonti et al., 2010).

A major limitation of most current studies on antisocial behavior is that they are mainly based on passive (viewing) tasks, and explicit self-report of cognitive processes (Dodge, 2010; but see White et al., 2013). It is possible that different cognitive processes are involved in real social interactions. For instance, it is well known that adolescents do not differ from adults in explicit risk perception, but take more risks in real life situations (Steinberg, 2010). In addition, the relation between atypical social-affective processes and both antisocial behavior and callous and unemotional (CU) traits has largely been investigated in community samples. This relationship may be different for participants displaying antisocial behavior at the extremely severe end of the spectrum, such as juvenile delinquents. Thus, the aim of the current study is to further elucidate the neural processes underlying social interactions of severely antisocial adolescents using an ecologically valid social interaction paradigm.

In the past decade the use of economic games in combination with neuroimaging emerged as fruitful method for investigating the neural correlates of social cognitive processes underlying real social interactions in both normative (Graham et al., 1992) and clinical populations (Kishida et al., 2010). The advantage of these games is that their structural simplicity yields precise characterizations of complex social behavior, which enables researchers to delineate specific steps in social information processing. Previous fMRI studies with healthy adults suggest that, in social interactions, the affective response related to the detection of violations of social norms is associated with activation in the anterior insula (AI) and the dorsal anterior cingulate cortex (dACC) (Sanfey et al., 2003; Montague and Lohrenz, 2007; Sanfey, 2007). Furthermore, the understanding of intentions and thinking about others has been associated with temporo-parietal junction (TPJ) and the medial prefrontal cortex (mPFC) (Decety et al., 2009; Güroğlu et al., 2010). Finally, the regulation of behavior and selection
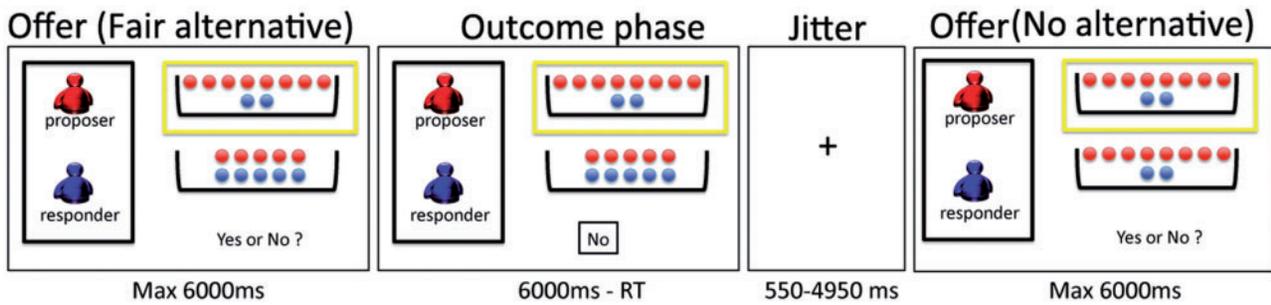
Fig. 1 Visual display of events presented in the mini Ultimatum Game. Trials started with a jittered fixation screen lasting 550–4950 ms. The left panel in the decision screen displayed the name of the proposer in red (here 'proposer') and the name of the responder (here 'responder'). Two offers each containing red and blue coins indicate the share for the proposer (red coins) and the responder (blue coins), the offer made by the proposer was encircled in yellow (here 8/2). The responder was given a maximum response time of 5000 ms to select Yes or No to accept or reject the offer. If they failed to respond within 5000 ms, a screen displaying 'Too late!' was presented for 1000 ms. Upon response, the feedback screen displayed the given response (here 'No') until 6000 ms after the start of the trial. Trials were randomized and presented with a jittered interstimulus interval (mean = 1530 s, min = 550 ms, max = 4950 ms; optimized with OptSeq2, (surfer.nmr.mgh.harvard.edu/optseq/). Both the 'fair alternative' and 'no alternative' condition are displayed.

of appropriate responses has been associated with the lateral prefrontal cortex (LPFC; Sanfey et al., 2003; Knoch et al., 2006).

For the current study, we recruited male adolescents showing severely antisocial behavior from forensic institutions, and age, gender and IQ matched control participants. The participants played the mini-Ultimatum Game (Figure 1) while being scanned in a magnetic resonance imaging (MRI) scanner. The mini-Ultimatum Game is a two-choice modified version of the Ultimatum Game (Güth et al., 1982) aimed at elucidating the role of intentions in fairness considerations. It has been shown that assessment of fairness is strongly modulated by the ascription of intentions: people react less negatively to unfair offers when they feel the inequity was unintentional (Blount, 1995; Güroğlu et al., 2009, 2010, 2011).

During the scanning session we set out to investigate the behavioral and neural responses to unfair offers. Based on previous studies we expected that participants would reject unfair offers more often when the alternative was fair, compared to when the alternative was also unfair (i.e. 'no alternative'), indicating they take the intentions of the proposer into account (Falk et al., 2008; Güroğlu et al., 2009, 2010, 2011). We expected that juvenile delinquents deviate in behavior particularly in the 'no alternative' condition, when the intentions of the proposer are ambiguous. Our imaging analyses were aimed at exploring group differences in neural activity in the brain regions associated with the processes underlying fairness considerations.

Finally, recent studies have also highlighted that antisocial youth represents a very heterogeneous group (Sebastian et al., 2012; Viding et al., 2012). One approach to capture the heterogeneity associated with antisocial behavior has been to map callous and unemotional (CU) traits (for review, see Viding et al., 2012). Thus, to take into account the heterogeneity of the delinquent group we explored the relation between brain activity, behavior and CU traits in the delinquent group.

METHODS

Participants

Thirty-four male participants aged 15 to 21 years took part in the study, consisting of a group of juvenile delinquents showing severely antisocial behavior ($N = 17$) and an age- and IQ-matched control group ($N = 17$). The juvenile delinquents were selected from a juvenile detention center ($N = 7$), and a forensic treatment center ($N = 10$).

Criteria for inclusion for the juvenile delinquents were at least one account of violent behavior and/or multiple accounts of non-violent behavior (see Supplementary Data for full description on assessment of antisocial behavior of both groups). Furthermore, participants had to

be right-handed, were required to have a full IQ of 80 or higher and the ability to speak Dutch at primary school level. Reason for exclusion was current use of psychotropic medications that could not be stopped for the scanning session. All participants from the control group were healthy and right-handed volunteers who reported no neurological or psychiatric impairments. The control group was matched for mean age and IQ (for more details on assessment and matching procedures see Supplementary Data).

All participants provided informed consent (parents provided consent for participants younger than 18 years), and all procedures were approved by the medical ethical committee of the Leiden University Medical Center. A radiologist reviewed all anatomical scans; no anomalies were found.

Task description

Participants played the role of the responder in the mini-Ultimatum Game. This is a two-choice Ultimatum Game where one unfair offer is presented together with an alternative offer. This modification of the Ultimatum Game includes experimental manipulations that enable us to study intention considerations. There were three conditions that were labeled depending on the alternative offer pitted against a fixed unfair 8/2 offer: (i) 5/5 offer (fair alternative), (ii) 8/2 offer (no alternative) and (iii) 2/8 offer (hyperfair alternative).

Before the scanning session participants practiced the task on a computer until they fully understood it and subsequently they played 168 trials in the scanner. In these 168 trials participants received 126 unfair offers and 42 fair offers (filler trials). The unfair offers were divided over the three conditions (42 fair alternative, 42 no alternative and 42 hyperfair alternative).

The trials were presented in three blocks of 56 trials lasting 8.3 min each. For the purposes of our study the analyses are focused on unfair offers, where we made the comparison of unfair offers in the context of fair alternatives and no alternatives. The hyperfair alternatives were not included in the analyses.

Each trial was played with a different age and gender matched anonymous proposer to avoid reputation effects. Participants were explained that the offers of the proposers had been obtained in an earlier part of the study. They were also explained that at the end of the session the computer would randomly select 10 trials to determine their total earnings, which would be added to the standard compensation for their participation (cf. Güroğlu et al., 2011). Furthermore, participants were told that the proposers' earnings would be contingent upon their decisions. In reality, the offers presented to the participants were computer simulated, but were based on behavior

reported in prior experiments (Güroğlu *et al.*, 2010). None of the participants expressed doubt about the set up of the task. The control group was debriefed directly after the experiment. In order to prevent participants recruited from the detention and treatment centers from informing each other about the true set up, they were not immediately debriefed.

Participants were given a variable reward between 3.50 and 5.50 euros. The standard compensation for participating was 20 euros, except for detainees, for whom payment was limited by government regulations. Detained participants were provided with 10 euros in telephone cards, which was considered a fitting and attractive incentive by the institution psychologists. The reason for using different pay offs was practical, but prior results showed that rejection rates on the Ultimatum Game are relatively robust against variations in magnitude of payment (Falk *et al.*, 2008; Amir *et al.*, 2012).

MRI data acquisition

The scanning session was carried out at the Leiden University Medical Center using a 3.0T Philips Achieva. The scanning sessions consisted of three types of scans in the following order: (i) localizer scan, (ii) T2*-weighted echo-planar imaging (EPI) sequence measuring the bold-oxygen-level-dependent (BOLD) signal [TR = 2.2 s, TE = 30 ms, slice-matrix = 80 × 80, slice-thickness = 2.75 mm, slice gap = 0.28 mm gap, field of view (FOV) = 220 mm], (iii) high-resolution T1-weighted anatomical scan. The groups did not show difference in mean head displacement during the scanning session [$F(1, 32) < 1$, $P = .82$], and none of the participants showed a displacement that was larger than the maximum allowed threshold of 3 mm.

MRI data analysis

SPM5 software (www.fil.ion.ucl.ac.uk/spm/) was used for image pre-processing and analyses. Slice-time correction, realignment, spatial normalization to EPI templates and spatial smoothing using a 6 mm full-width half-maximum 3D Gaussian kernel were carried out. The functional time series were modeled by a series of events convolved with a canonical hemodynamic response function (HRF). The moment of stimulus presentation with null duration was used to model the data. The unfair offers (8/2 offers) were modeled separately based on context (two levels: fair- or no alternative) and response (two levels: accept or reject). For the purposes of the current experiment the unfair trials with a hyperfair alternative and the fair offer trials (filler trials) were modeled as events of no interest. Contrast images for each individual were used in the second-level random effects model to run full-factorial analysis of variance and one-tailed *post hoc* t-tests. We conducted regression analyses to test for brain–behavior relations. For whole-brain analyses a significance threshold of $P < 0.05$ FWE corrected for multiple comparisons was calculated with AlphaSim, resulting in an uncorrected threshold of $P < 0.001$, requiring a minimum of 24 voxels in a cluster (Forman *et al.*, 1995). By iterating the process of random image generation, spatial correlation of voxels, thresholding and cluster identification, the program provides an estimate of the overall significance level achieved for various combinations of individual voxel probability threshold and cluster size threshold that is equal to a FWE corrected threshold of $P < 0.05$ (Forman *et al.*, 1995; Poline *et al.*, 1997 and see Bennet *et al.*, 2009 for comparison other methods).

Regions of interest (ROIs) analyses were based on functional masks of the group level whole-brain choice contrast, based on all participants. For these analyses mean parameter estimates were extracted for each ROI for each individual. Effects were considered significant at an α of 0.0083, based on Bonferroni correction for multiple comparisons. The multiple comparisons were based on six ROIs which were based on the general contrast 'accept unfair offer' *vs* 'reject unfair offer',

$P = 0.05/6$ ROIs [rTPJ, rIFG, dACC, posterior cingulate cortex (PCC), anterior insula and ventral striatum]. For the ROI analyses we performed additional robust regression analyses with the Huber weighting function to account for possible effects driven by outliers. Given that these analyses did not change the results (all significant results meet $P < 0.0083$ threshold) we have decided to report the results of the linear correlation analyses. Results are reported in the MNI305 stereotaxic space.

RESULTS

Rejection rates

Given that free choice patters are not normally distributed, non-parametric tests were used. The analysis of behavioral responses to unfair offers revealed that participants rejected unfair offers more often in the 'fair alternative' condition (Median = 88%) compared with the 'no alternative' condition (Median = 52%, Wilcoxon signed rank test, $P < 0.0001$, Figure 2A; left panel). To test for group differences we computed difference scores in rejection rates between the 'fair alternative' and 'no alternative' condition. As can be seen in Figure 2A (right panel), the juvenile delinquents showed a smaller difference between the 'fair alternative' and 'no alternative' condition than the control group (Kruskal–Wallis test non-normality of the data, $H_{(1)} = 6.13$, $P < 0.04$). *Post hoc* test revealed that the behavioral difference was driven by significantly higher rejection rates in the 'no alternative' condition in juvenile delinquents [Kruskal–Wallis test, $H_{(1)} = 7.41$, $P < 0.02$], whereas there was no group difference in the 'fair alternative' condition [Kruskal–Wallis test, $H_{(1)} = 3.6$, $P = 0.24$].

fMRI results

To identify the brain regions involved in deciding to accept or reject unfair offers we performed whole-brain contrasts [accept$_{unfair}$ *vs* reject$_{unfair}$] across both experimental conditions. For all participants, accepting unfair offers was associated with increased activity in the right IFG, right TPJ, bilateral ventral striatum, PCC and the network comprising the dACC and anterior insula (Table 1 and Figure 2B). No areas were more active when rejecting compared with accepting unfair offers. The areas that showed increased activity were used as ROIs in subsequent analyses. Finally, we have compared the results of the whole-brain contrast [accept > reject] between groups to investigate whether there were brain areas that were engaged in one group but not the other. None of these contrasts (Delinquent$_{[accept>reject]}$ − Control$_{[accept>reject]}$ and Delinquent$_{[reject>accept]}$ − Control$_{[reject>accept]}$) revealed any significant activation differences, even at the more liberal threshold of $P < 0.005$ uncorrected. This result suggests that both groups rely on the same general network of brain areas when deciding whether to accept or reject the ultimatum game offers.

ROI analyses

To further investigate the effect of the experimental conditions and groups on activity in the ROIs identified in the whole-brain contrast, we conducted 2 × 2 × 2 ANOVAs with response (accept and reject) and condition ('fair alternative' and 'no alternative') as the within subject factors, and group (juvenile delinquents and controls) as the between subjects factor. These analyses did not yield a three-way interaction for any ROI, but there were group × condition interactions in the rTPJ [$F_{(1, 24)} = 6.72$, $P < 0.006$] and the rIFG [$F_{(1, 26)} = 7.73$, $P < 0.005$], and condition × response interactions in the dACC [$F_{(1, 26)} = 6.18$, $P < 0.002$] and anterior insula [$F_{(1, 26)} = 3.83$, $P < 0.007$]. No interactions with group or condition were found in the PCC and bilateral ventral striatum.

To further explore the group × condition interactions in brain activity, we performed post-hoc analyses of activity patterns in the
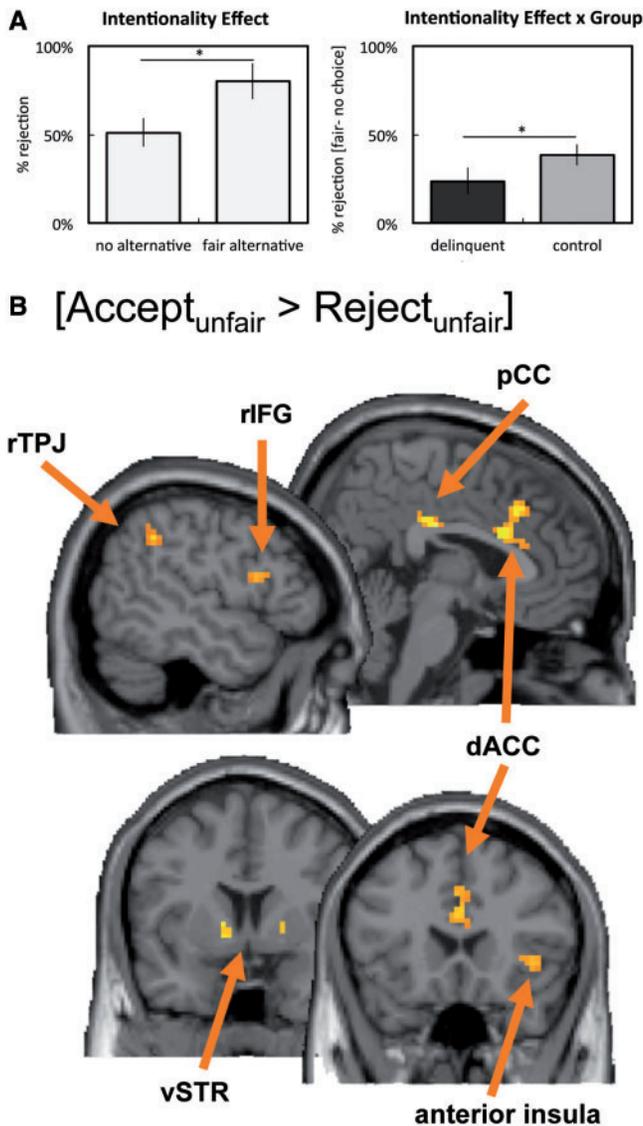
A   Intentionality Effect   Intentionality Effect x Group

B [Accept_unfair > Reject_unfair]

Fig. 2 (**A**) Display of means and standard errors of rejection rates of unfair offers in the 'fair alternative' and 'no alternative' conditions collapsed over groups, and difference scores (fair alternative–no alternative) in rejection rates for each group separately. (**B**) Network of brain regions that was more active for accepting than rejecting unfair offers: (rTPJ: right temporal parietal junction [54, −45, 41], rIFG: right inferior frontal gyrus [54, 12, 18], dACC: dorsal anterior cingulate cortex [3, 18, 27], PCC [0, −33, 30], anterior insula: anterior insula [36, 15, 19], Vstr: bilateral ventral striatum [14, 9, −4] and [−12, 9, −4]. Cluster corrected threshold: $P < 0.001$ and $k > 24$ voxels.

Table 1 Brain regions revealed by whole-brain contrast accept unfair offer–reject unfair offer

Anatomical region	L/R	Z	MNI coordinates		
			x	y	z
Accept_unfair > Reject_unfair					
TPJ	R	5.61	54	−45	41
IFG	R	4.88	54	12	18
anterior insula	R	4.51	45	37	21
dACC	R	4.60	57	−48	27
PCC	R	4.55	54	−57	39
Ventral Striatum*					
	R	4.36	12	7	−5
	L	4.36	−17	9	−2

MNI coordinators for main effects, peak voxels reported at $P < 0.001$, at least 24 contiguous voxels.
*Striatal ROIs < 16 voxels, collapsed to one ROI in analyses.

rTPJ and rIFG ROIs. These analyses revealed that the control group showed increased rTPJ activity in the 'no alternative' condition relative to the juvenile delinquents [$F_{(1, 24)} = 11.01$, $P < 0.002$, Figure 3A], whereas the groups did not differ in activation in the 'fair alternative' condition [$F_{(1, 24)} < 1$, $P = 0.53$].

In contrast, there was increased activity in the rIFG in the 'fair alternative' condition for the control group compared with the juvenile delinquents [$F_{(1, 26)} = 6.15$, $P < 0.001$, Figure 3B], whereas there was no difference in activity in the 'no alternative' condition [$F_{(1, 26)} < 1$, $P = 0.63$]. Taken together, these results suggest that the rIFG and rTPJ are both more active for control participants than for the juvenile delinquents, but this difference was dependent on the context of the unfair offer (intention consideration vs fairness judgments). These separate patterns were confirmed statistically by a significant three-way interaction between condition × group × region [$F_{(1,32)} = 29.19$, $P < 0.001$].

Finally, we examined the condition × response interaction in the dACC and anterior insula. These post hoc analyses showed that the activation in both the dACC and right anterior insula were higher for accepting than rejecting unfair offers in the 'fair alternative' condition [dACC: $t_{(1,33)} = 2.89$, $P < 0.008$; anterior insula: $t_{(1,33)} = 3.43$, $P < 0.002$; Figure 4]. In contrast, activation in the anterior insula was higher for rejecting than accepting unfair offers in the 'no alternative' condition [$t_{(1,26)} = −2.71$, $P < 0.01$]. Consistent with previous studies these results suggest that, in the context of the mini Ultimatum Game, the dACC and anterior insula are associated with the detection of personal norm violations (e.g. accepting an unfair offer when the proposer had a fair alternative, see also Supplemental Figure S1). (Montague and Lohrenz, 2007; van den Bos et al., 2009; Guroglu et al., 2010).

Brain–behavior correlations

Given that the rTPJ and rIFG were previously suggested to be instrumental in accepting or rejecting unfair offers, we performed exploratory analyses on the relation between rejection rates and brain activity in these areas. In addition, we conducted exploratory analyses focusing on the individual differences in CU traits within the delinquent group.

Based on our initial ROI analyses we hypothesized that rTPJ activity would be related to the relative decrease in rejection rates in the 'no alternative' condition, whereas rIFG activity was expected to be related to rejection rates in the 'fair alternative' condition. First, we explored the relation between difference scores in rejection rates ['fair alternative'–'no alternative' condition] and the activity in the rTPJ for the ['no alternative' vs 'fair alternative'] contrast. These analyses revealed a significant positive correlation between difference scores and rTPJ activity for the control group ($r = 0.76$, $P < 0.004$), and a significant positive correlation for the juvenile delinquents, although the latter did not survive Bonferroni correction ($r = 0.49$, $P = 0.04$; Figure 3C). Post hoc comparison of differences in regression slopes between difference scores and rTPJ activity was significantly greater for the control group than for the juvenile delinquents ($t = 2.41$, $P < 0.02$). Thus the more control participants showed increased rTPJ activity in the no alternative condition relative to the fair alternative condition, the more they accepted unfair offers in the no alternative condition relative to the fair alternative condition.

Second, we investigated the relation between rejection rates in the 'fair alternative' condition and the activity in the rIFG for the ['fair alternative' vs 'fixation'] contrast. This analysis revealed a significant negative correlation between rejection rates and rIFG activity for the control group ($r = −0.75$, $P < 0.001$), but not for the juvenile delinquents ($r = −0.01$, $P = 0.94$, Figure 3D). Again, post hoc comparison of differences in regression slopes revealed that the correlation between
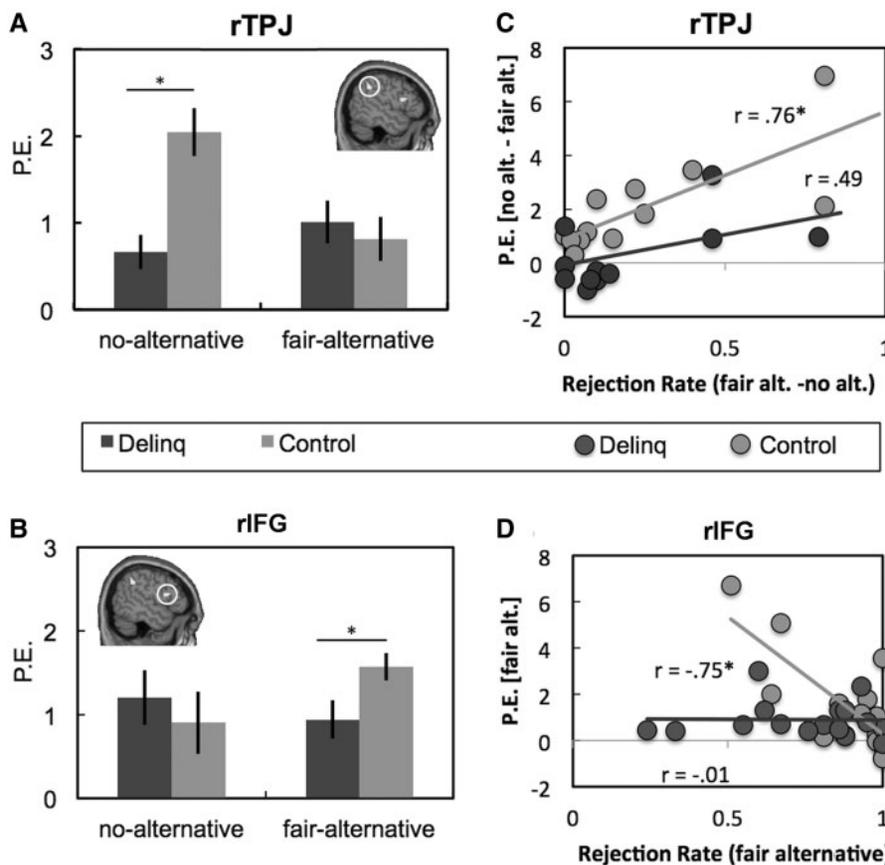
Fig. 3 Contrast values in (**A**) right TPJ and (**B**) right IFG for juvenile delinquent and control participants in the 'fair alternative' and 'no alternative' conditions. (**C**) Activation in the right TPJ for the [no alternative–fair alternative] contrast correlated positively with the difference scores [fair alternative–no alternative] in rejection rates for the control group, but not for the juvenile delinquents. (**D**) Activation in the right IFG in the [fair alternative–fixation] condition correlated positively with the rejection rates in the fair condition for the control group, but not significantly for the juvenile delinquents. ROI results are considered significant at a Bonferroni corrected α of $P < 0.008$.
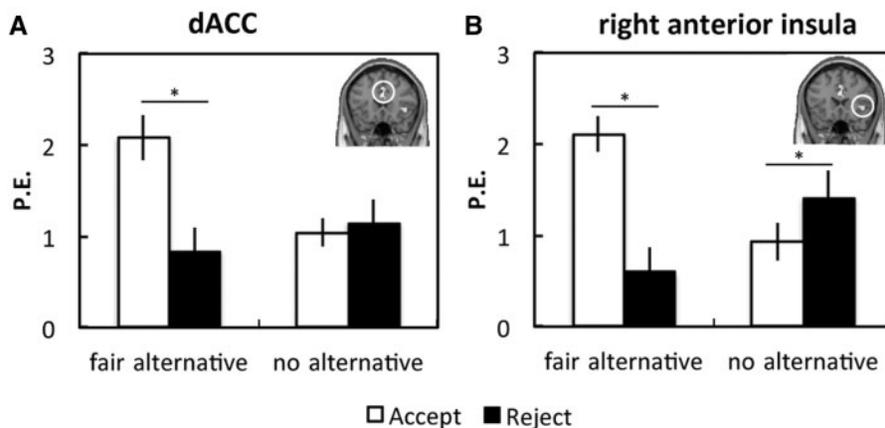

Fig. 4 Contrast values in (**A**) dACC and (**B**) right anterior insula for accepting and rejecting unfair offers in the 'fair alternative' and 'no alternative' conditions collapsed across all participants. ROI results are considered significant at a Bonferroni corrected α of $P < 0.008$.

rejection rates and rIFG activity was significantly greater for the control group than for the juvenile delinquents ($t = 6.47$, $P < 0.001$).

Individual differences within the delinquent population

To take into account the heterogeneity of the severely antisocial group we investigated the relation between behavior and brain activity with the scores on the callous–unemotional (CU) dimension, and its underlying constructs. We have also explored whether drug and alcohol use had any relation to the behavior or brain activity within the delinquent population, but this did not yield any significant result (see Supplementary Table S1 for more detail).

Based on several studies that have suggested the factors that underlie CU may be independent and differentially associated with aggression,
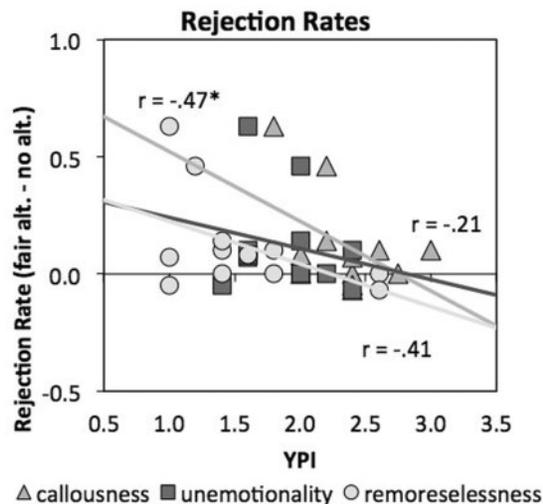
Fig. 5 Correlation with callousness, unemotionality and remorselessness with difference scores (intentionality effect: fair alternative–no alternative) in rejection rates. Delinquent data only.

delinquency and emotional reactivity (see Kimonis et al., 2008 for a large adolescent sample) we investigated the individual factors as well as the usual composite score of the three CU dimensions. First, we found that the higher the score on callousness the smaller the intentionality effect was in terms of rejection rates ($r = -0.47$, $P < 0.05$, Figure 5). Thus, those with a low callousness score rejected less when the other had no choice compared to when the other had a fair alternative, whereas those with a high callousness score did not show a difference between the conditions. We did not find a significant correlation with rejection rates and unemotionality ($r = -0.21$, $P = 0.51$) or remorselessness ($r = -0.41$, $P = 0.08$).

The neuroimaging data showed that in the TPJ and the IFG there was a decrease in activity related to increased callousness, but these effects failed to reach significance ($r = -0.42$, $P = 0.11$ and $r = -0.37$, $P = 0.15$, respectively). We did not find any significant correlation for unemotionality, remorselessness or with the composite score of the three CU dimensions (all P's > 0.2). Finally, further exploration of correlations with the CU dimension or factors in the affective ACC/Insula network did not reveal any significant results or trends (all P's > 0.3).

DISCUSSION

The goal of this study was to gain a better understanding of the neural processes underlying social decision-making of juvenile delinquents who show severely antisocial behavior. Behavioral analyses indicated that all participants showed lower rejection rates in the 'no alternative' than in the 'fair alternative' condition, suggesting they take the social context of the proposals into account (Falk et al., 2008; Güroğlu et al., 2009, 2011). However, juvenile delinquents showed less acceptance of the no alternative condition compared with the control group. This suggests that they react more strongly to the unfairness of the offer, or are less concerned about the intentions behind the offer.

The imaging results revealed that these differences in behavior were accompanied by group differences in a specific subset of brain areas (rTPJ and rIFG). First, the severely antisocial adolescents showed reduced levels of rTPJ activity in the 'no alternative' condition (i.e. the condition which required intention consideration) compared with the control group. This finding can be interpreted in the context of a recent meta-analysis (Carter and Huettel, 2013) that suggests that the TPJ is a convergence zone for different functions (memory, attention,

social processing). This convergence enables the TPJ to have a higher order role in the creation of a social context for behavior. The locus of activation of the current study falls right in the area where activation is associated with both attention and social cognitive processes (ToM). This suggests that the delinquent group may be less focused on the social context of the ultimatum offer. This interpretation is supported by our exploratory analyses of individual differences within the delinquent group that showed that rejection rates were related with callousness. However, future studies should further investigate the relation between attention and social context of Ultimatum Game rejections by making the outcomes for the other even more salient or explicitly train the participants to focus on them (for instance, see Sebastian et al., 2013).

Second, severely antisocial adolescents differed from the control group by showing decreased rIFG activity in the 'fair alternative' condition, and not showing a correlation between rIFG activity and behavioral responses to unfair offers. These results are in line with previous studies that have suggested that juvenile antisocial behavior is related to difficulties to engage the regulatory processes associated with the frontal cortex (Fairchild et al., 2009b; Sterzer and Stadler, 2009). More specifically, the rIFG is often associated with both attention (selection/switching) and response inhibition (Aron et al., 2004; Hampshire et al., 2010). In context of the mini-Ultimatum Game reduced activity can thus be interpreted as reduced attentional processing or reduced inhibition of prepotent responses. Based on our current design we cannot conclude whether reduced activation was associated with less attention to the social context of the unfair offer, or failure of the inhibition of the prepotent response (reject unfair offers).

Finally, there where no group differences in how norm violations were processed in the dACC and right anterior insula. A recent study used resting state connectivity patterns to define three functionally distinct networks in the insula (posterior, ventral anterior and dorsal anterior; Chang et al., 2012). The function of these networks was further specified by large-scale reverse inference based on the Neurosynth database (Yarkoni et al., 2011). The dorsal anterior network, which overlaps with our functional activation, showed high connectivity with the dACC and is thought to be involved in processing conflict and errors in a diverse set of tasks, whereas the ventral posterior network is thought to be associated with emotion processing and anxiety. These findings suggest that severely antisocial adolescents and control participants may both be equally aware of behaving against their personal norms.

The findings of this study suggest that there are disturbances of processes associated with rTPJ and rIFG activity underlying aberrant social behavior in juvenile delinquents. Earlier studies have shown reduced LPFC engagement in antisocial groups (Fairchild et al., 2009b; Sterzer and Stadler, 2009); the current results extend these findings by showing that also the TPJ, an area associated with social cognition, shows reduced levels of activation in juvenile delinquents with severely antisocial behavior. These results support the hypothesis that in social interactions, severely antisocial adolescents do not have the tendency to take the social context fully into account. However, the correlation between TPJ activity and rejection rates suggests that when the severely antisocial adolescents take the perspective of the other they are more willing to accept an unfair offer, just as the control participants. This finding is in line with studies that suggest that cognitive perspective taking skills are not necessarily deficient in antisocial populations, but they may not spontaneously engage them (Jones et al., 2010, Schwenck et al., 2011; Sebastian et al., 2013).

Finally, the results of this study should be interpreted in the light of several important limitations. First, the results of the current study do not reveal the causes of atypical processing in rIFG and rTPJ. Recent

work has demonstrated that both environmental (Dodge *et al.*, 2006) and genetic factors are related to the development of antisocial behavior (Wallace *et al.*, 2007; Cesarini *et al.*, 2008). In the current study we did not have any explicit measures of environmental variables, such as socioeconomic status, or genetic make-up. Longitudinal neuroimaging studies are needed to elucidate how environmental and genetic factors interact to give rise to the changes in the brain that are related to antisocial behavior. With a larger sample the current design is promising in providing a better understanding of such developmental trajectories across adolescence. Second, our sample contained boys only, so we do not know whether our results are generalizable to the less studied group of girls who show severely antisocial behavior.

In sum, the current results offer empirical support for aberrant social decision-making in severely antisocial adolescents, and provide a template for the development of quantitative measures that may be useful for the understanding of the development and prognosis of antisocial behavior. The results emphasize the importance of understanding the social aspects of antisocial behavior. Adolescence is a period that is characterized by a unique set of physical, social and neurological changes (Spitzer *et al.*, 2007; Casey *et al.*, 2008). It is hypothesized that these changes contribute to both typical adolescent aberrant behavior, but also provide a unique window of opportunity for re-directing behavior in case development goes astray (for review, see Crone and Dahl, 2012). In the future, neuroimaging may provide useful additional prognostic information, or biomarkers, for treatment (Popma *et al.*, 2006).

SUPPLEMENTARY DATA

Supplementary data are available at *SCAN* online.

Conflict of Interest

None declared.

REFERENCES

Achenbach, T.M. (1991). *Integrative Guide for the 1991 CBCL/4-18, YSR, and TRF Profiles.* Burlington: Department of Psychiatry, University of Vermont.

Amir, O., Rand, D.G., Gal, Y.K. (2012). Economic games on the internet: the effect of $1 stakes. *PloS One*, 7, e31461.

Andershed, H.A., Kerr, M., Stattin, H., Levander, S. (2002). Psychopathic traits in non-referred youths, A new assessment tool. In: Blauuw, E., Sheridan, L., editors. *Psychopaths: Current International Perspectives* The Hague: Elsevier, pp. 131–58.

Aron, A.R., Behrens, T.E., Smith, S., Frank, M.J., Poldrack, R.A. (2007). Triangulating a cognitive control network using diffusion-weighted magnetic resonance (MRI) and functional MRI. *Journal of Neuroscience*, 27, 3743–52.

Arsenio, W.F., Lemerise, E.A. (2004). Aggression and moral development: integrating social information processing and moral domain models. *Child Development*, 75, 987–1002.

Blount, S. (1995). When social outcomes aren't fair, the effect of causal attributions on preferences. *Organizanizational Behavior and Human Decision Procceses*, 63, 131–44.

Carter, R.M., Huettel, S.a. (2013). A nexus model of the temporal-parietal junction. *Trends in Cognitive Sciences*, 17, 328–36.

Casey, B.J., Jones, R.M., Hare, T.A. (2008). The adolescent brain. *Annals of the New York Academy of Sciences*, 1124, 111–26.

Cesarini, D., Dawes, C.T., Fowler, J.H., Johannesson, M., Lichtenstein, P., Wallace, B. (2008). Heritability of cooperative behavior in the trust game. *Proceedings of the National Academy of Sciences*, 105, 3721–6.

Chang, L.J., Yarkoni, T., Khaw, M.W., Sanfey, A.G. (2013). Decoding the role of the insula in human cognition: functional parcellation and large-scale reverse inference. *Cerebral Cortex*, 23(3), 739–49.

Crick, N.R., Dodge, K.A. (1994). A review and reformulation of social information-processing mechanisms in childrens social-adjustment. *Psychological Bulletin*, 115, 74–101.

Crone, E.A., Dahl, R.E. (2012). Understanding adolescence as a period of social & affective engagement and goal flexibility. *Nature Reviews Neuroscience*, 13, 636–50.

Decety, J., Michalska, K.J., Akitsuki, Y., Lahey, B.B. (2009). Atypical empathic responses in adolescents with aggressive conduct disorder, a functional MRI investigation. *Bioloical Psychology*, 80, 203–11.

Dodge, K.A., Frame, C.L. (1982). Social cognitive biases and deficits in aggressive boys. *Child Development*, 53, 620–35.

Dodge, K.A., Coie, J., Lynam, D. (2006). Aggression and antisocial behavior in youth. In: Eisenberg, N., editor. *Handbook of Child Psychology, Vol 3, Social, Emotional, and Personality development.* New York: Wiley, pp. 719–88.

Dodge, K.A. (2010). Social information processing patterns as mediators of the interaction between genetic factors and life experiences in the development of aggressive behavior. In: Shaver, M.M.P.R., editor. *Understanding and Reducing Aggression, Violence, and their Consequences* Washington: American Psychological Association, pp. 165–85.

Fairchild, G., Van Goozen, S., Stollery, S., Goodyer, I. (2009a). Deficits in facial expression recognition in male adolescents with early-onset or adolescence-onset conduct disorder. *Journal of Child Psychology and Psychiatry*, 50, 627–36.

Fairchild, G., Van Goozen, S., Stollery, S., Aitken, M., Savage, J., Moore, S., Goodyer, I. (2009b). Decision making and executive function in male adolescents with early-onset or adolescence-onset conduct disorder and control subjects. *Biological Psychiatry*, 66, 162–8.

Falk, A., Fehr, E., Fischbacher, U. (2008). Testing theories of fairness—intentions matter. *Games and Economic Behavior*, 62, 287–303.

Forman, S.D., Cohen, J.D., Fitzgerald, M., Eddy, W.F., Mintun, M.A., Noll, D.C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI) use of a clustersize threshold. *Magnetic Resonance and Medicine*, 33, 636–47.

Graham, S., Hudley, C., Williams, E. (1992). Attributional and emotional determinants of aggression among African-American and Latino young adolescents. *Developmental Psychology*, 28, 731–40.

Güroğlu, B., van den Bos, W., Crone, E.A. (2009). Fairness considerations, increasing understanding of intentionality during adolescence. *Journal of Experimental Child Psychology*, 104, 398–409.

Güroğlu, B., van den Bos, W., Rombouts, S.A.R.B., Crone, E.A. (2010). Unfair? It depends, neural correlates of fairness in social context. *Social Cognitive Affective Neuroscience*, 5, 414–23.

Güroğlu, B., van den Bos, W., van Dijk, E., Rombouts, S.A.R.B., Crone, E.A. (2011). Dissociable brain networks involved in development of fairness considerations, understanding intentionality behind unfairness. *NeuroImage*, 57, 634–41.

Güth, W., Schmittberger, R., Schwarze, B. (1982). An experimental analyses of ultimatum barganing. *Journal of Economic Behavior & Organization*, 3, 367–88.

Ha, C., Sharp, C., Goodyer, I. (2011). The role of child and parental mentalizing for the development of conduct problems over time. *European Child & Adolescent Psychiatry*, 20, 291–300.

Hampshire, A., Chamberlain, S.R., Monti, M.M., Duncan, J., Owen, A.M. (2010). The role of the right inferior frontal gyrus: inhibition and attentional control. *NeuroImage*, 50, 1313–9.

Happé, F., Frith, U. (1996). Theory of mind and social impairment in children with conduct disorder. *British Journal of Developmental Psychology*, 14, 385–98.

Herpertz, S.C., Huebner, T., Marx, I., Vloet, T.D., Fink, G.R. (2008). Emotional processing in male adolescents with childhood-onset conduct disorder. *Journal of Child Psychology and Psychiatry*, 49, 781–91.

Jones, A.P., Happé, F., Gilbert, F., Burnett, S., Viding, E. (2010). Feeling, caring, knowing, different types of empathy deficit in boys with psychopathic tendencies and autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, 51, 1188–97.

Kimonis, E.R., Frick, P.J., Skeem, J.L., Marsee, M.A. (2008). Assessing callous-unemotional traits in adolescent offenders: validation of the inventory of callous-unemotional traits. *Journal of the International Association of Psychiatry and Law*, 31, 241–51.

Kishida, K.T., King-Casas, B., Montague, P.R. (2010). Neuroeconomic approaches to mental disorders. *Neuron*, 67, 543–54.

Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., Fehr, E. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science*, 314, 829–32.

Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S.F., Baker, C.I. (2009). Circular analysis in systems neuroscience, the dangers of double dipping. *Nature Neuroscience*, 12, 535–40.

Loeber, R., Burke, J.D., Pardini, D.A. (2009). Development and etiology of disruptive and delinquent behavior. *Annual Review Clinical Psychology*, 5, 291–310.

Loeber, R., Burke, J.D., Lahey, B.B., Winters, A., Zera, M. (2000). Oppositional defiant and conduct disorder, a review of the past 10 years, Part I. *Journal of the American Academy of Child & Adolescent Psychiatry*, 39, 1468–84.

Marsh, A.A., Blair, R. (2008). Deficits in facial affect recognition among antisocial populations, a meta-analysis. *Neuroscience Biobehavioral Review*, 32, 454–65.

Maughan, B., Rutter, M. (2001). Antisocial children grown up. Conduct disorders in childhood and adolescence. In: Hill, J., Maughan, B., editors. *Conduct Disorders in Childhood and Adolescence* New York: Cambridge University Press, pp. 507–52.

Montague, P.R., Lohrenz, T. (2007). To detect and correct, norm violations and their enforcement. *Neuron*, 56, 14.

Moffitt, T.E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: a developmental taxonomy. *Psychological Review*, 100, 674–701.

Moffitt, T.E., Scott, S. (2009). Conduct disorders of childhood and adolescence. In: Rutters, et al. editors. *Rutter's Child and Adolescent Psychiatry*. Malden MA, USA: Blackwell Publishing Ltd, pp. 543–64.

Mitchell, J.P. (2008). Activity in right temporo-parietal junction is not selective for theory-of-mind. *Cerebral Cortex*, 18, 262–71.

Passamonti, L., Fairchild, G., Goodyer, I.M., et al. (2010). Neural abnormalities in early-onset and adolescence-onset conduct disorder. *Archives of General Psychiatry*, *67*, 729–38.

Patterson, G.R., DeBaryshe, B.D., Ramsey, E. (1989). A developmental perspective on antisocial behavior. *American Psychologist*, *44*, 329–35.

Rilling, J.K., Sanfey, A.G. (2011). The neuroscience of social decision-making. *Annual Review Psychology*, *62*, 23–48.

Sanfey, A.G. (2007). Social decision-making, insights from game theory and neuroscience. *Science*, *318*, 598–602.

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*, 1755–8.

Saxe, R., Wexler, A. (2005). Making sense of another mind, The role of the right temporo-parietal junction. *Neuropsychologia*, *42*, 1391–9.

Sebastian, C.L., McCrory, E.J., Dadds, M.R., et al. (2014). Neural responses to fearful eyes in children with conduct problems and varying levels of callous-unemotional traits. *Psychological Medicine*, *44*, 99–109.

Sebastian, C.L., Fontaine, N.M.G., Bird, G., et al. (2012). Neural processing associated with cognitive and affective theory of mind in adolescents and adults. *Social Cognitive and Affective Neuroscience*, *7*, 53–63.

Schwenck, C., Mergenthaler, J., Keller, K., et al. (2011). Empathy in children with autism and conduct disorder, group-specific profiles and developmental aspects. *Journal of Child Psychology and Psychiatry*, *53*, 651–59.

Sharp, C. (2008). Theory of mind and conduct problems in children, deficits in reading the 'emotions of the eyes'. *Cognition and Emotion*, *22*, 1149–58.

Steinberg, L. (2010). A dual systems model of adolescent risk-taking. *Developmental Psychobiology*, *52*, 216–24.

Sterzer, P., Stadler, C. (2009). Neuroimaging of aggressive and violent behaviour in children and adolescents. *Frontiers in Behavioral Neuroscience*, *3*, 35.

Sterzer, P., Stadler, C., Krebs, A., Kleinschmidt, A. (2005). Abnormal neural responses to emotional visual stimuli in adolescents with conduct disorder. *Biolocial Psychiatry*, *57*, 7–15.

Sterzer, P., Stadler, C., Poustka, F., Kleinschmidt, A. (2007). A structural neural deficit in adolescents with conduct disorder and its association with lack of empathy. *NeuroImage*, *37*, 335–42.

Spitzer, M., Fischbacher, U., Herrnberger, B., Grön, G., Fehr, E. (2007). The neural signature of social norm compliance. *Neuron*, *56*, 185–96.

Popma, A., Raine, A. (2006). Will future forensic assessment be neurobiologic? *Child & Adolescent Psychiatric Clinics of North America*, *15*, 429–44.

Van Goozen, S.H.M., Fairchild, G. (2008). How can the study of biological processes help design new interventions for children with severe antisocial behavior? *Developmental Psychopathology*, *20*, 941–73.

van den Bos, W., van Dijk, E., Westenberg, P.M., Rombouts, S.A.R.B., Crone, E.A. (2009). What motivates repayment? Neural correlates of reciprocity in the Trust Game. *Social Cognitive and Affective Neuroscience*, *4*(3), 294–304.

Viding, E., Fontaine, N.M.G., McCrory, E.J. (2012). Antisocial behaviour in children with and without callous-unemotional traits. *Journal of the Royal Society of Medicine*, *105*, 195–200.

Wallace, B., Cesarini, D., Lichtenstein, P. (2007). Heritability of ultimatum game responder behavior. *Proceedings of the National Academy of Sciences USA*, *104*, 15631–4.