



Published in final edited form as:

*J Pers Disord.* 2013 February ; 27(1): 99–112. doi:10.1521/pedi.2013.27.1.99.

## CALLOUS-UNEMOTIONAL TRAITS MODULATE THE NEURAL RESPONSE ASSOCIATED WITH PUNISHING ANOTHER INDIVIDUAL DURING SOCIAL EXCHANGE: A PRELIMINARY INVESTIGATION

Stuart F. White, PhD, Sarah J. Brislin, BA, Harma Meffert, PhD, Stephen Sinclair, PhD, and R. James R. Blair, PhD

National Institute of Mental Health

### Abstract

The current study examined whether Callous-Unemotional (CU) traits, a core component of psychopathy, modulate neural responses of participants engaged in a social exchange game. In this task, participants were offered an allocation of money and then given the chance to punish the offerer. Twenty youth participated and responses to both offers and the participant's punishment (or not) of these offers were examined. Increasingly unfair offers were associated with increased dorsal anterior cingulate cortex (dACC) activity but this responsiveness was not modulated by CU traits. Increasing punishment of unfair offers was associated with increased dACC and anterior insula activity and this activity *was* modulated by CU traits. Higher CU trait participants showed a weaker association between activity and punishment level. These data suggest that CU traits are associated with appropriate expectations of other individual's normative behavior but weaker representations of such information when guiding behavior of the self.

---

Humans regularly cooperate and the evolutionary mechanisms for such behavior have been of interest for some time (Boyd, Gintis, Bowles, & Richerson, 2003). Over the past decade, considerable progress has been made on the cognitive neuroscience of cooperation (de Quervain et al., 2004; Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003). This has led to work investigating whether selective disruptions in the mechanisms mediating cooperation might be associated with specific forms of pathology (e.g., borderline personality disorder; King-Casas et al., 2008).

Social neuroscience work has considered several cooperation tasks (e.g., the Ultimatum and Trust games). In such games, a proposer suggests an allocation of resources and typically the subject decides whether or not to accept this allocation and/or punish the proposer for the unfairness of his offer (e.g., de Quervain et al., 2004; King-Casas et al., 2008). Within this literature, unfair offers by proposers have been found to elicit activity in subjects within both the anterior insula cortex (AIC) and dorsal anterior cingulate cortex (dACC; King-Casas et al., 2008; Rilling et al., 2008; Sanfey et al., 2003). There have been suggestions that activity within these regions reflects anger elicited by unfairness to the self (Sanfey et al., 2003) or that they play a critical role in detecting and reacting to social norm violations (King-Casas et al., 2008; Spitzer, Fischbacher, Herrnberger, Gron, & Fehr, 2007). Indeed, it has been argued that AIC/inferior frontal cortexes (iFC) respond to anger/expectations of anger

(including in response to norm violations) and organize changes in behavior (Blair & Cipolotti, 2000).

Decisions to not accept the proposer's unfair offers are typically also associated with increased activity within AIC and dACC (e.g., King-Casas et al., 2008). In addition, there have been suggestions on the basis of increased striatal activity when punishing another that participants find punishing individuals rewarding (de Quervain et al., 2004), though it should be noted that increased activity when punishing is not only seen within striatum but also within AIC, anterior and posterior cingulate cortex, and dorsolateral prefrontal cortex (dlPFC; Strobel et al., 2011). While the ventromedial prefrontal cortex (vmPFC) has been less frequently seen following unfair offers or the rejection/punishment of such offers, patients with vmPFC damage show an increased propensity to reject unfair offers (Koenigs & Tranel, 2007).

One goal of the current study was to determine the neural correlates of performance on a social exchange game in youth. Only two previous fMRI studies have examined youth (Guroglu, van den Bos, van Dijk, Rombouts, & Crone, 2011; Sharp, Burton, & Ha, 2011). The results of Guroglu et al. (2011) are difficult to interpret as the study produced data inconsistent with most of the previous literature (decisions to accept, rather than reject, the proposer's unfair offers were associated with increased activity within AIC and dACC). However, the results of Sharp et al. (2011) are of interest in that this study contrasted youth with externalizing behavior who showed reduced differential responses, relative to typically developing youth, within AIC and caudate to the offers of a neutral relative to a mean or kind partner.

However, the main goal of the current study was to determine the modulation by callous-unemotional (CU: reduced guilt and empathy) traits on the neural responses associated with performance of a social exchange game. CU traits in youth with Conduct Disorder are associated with decreased amygdala responses to fearful expressions (Marsh et al., 2008; White et al., 2012) and disrupted vmPFC and caudate functioning (Finger et al., 2011; Finger et al., 2008). CU traits are thought to exist on a continuum and have been associated with theoretically consistent personality characteristics in community samples (Essau, Sasagawa, & Frick, 2006). Thus by examining the interaction of CU traits with BOLD responses in a healthy population, we can determine their interaction with neural responses in a population unconfounded by significant pathology.

Considerable work suggests that youth with Disruptive Behavior Disorders (DBD), particularly those with CU traits, show difficulties in decision-making (Blair, Colledge, & Mitchell, 2001; Budhani, Richell, & Blair, 2006). This appears to relate to deficits in prediction error signaling (differences between expected and received reward/punishment) and the representation of reinforcement value mediated by striatum and vmPFC (Finger et al., 2011; Finger et al., 2008). Recent work has also indicated reduced modulation of negative reinforcement expectancies within the AIC and dACC in youth with DBD; this information is less used by youth with DBD, particularly DBD with elevated CU traits, to guide the individual away from less than optimal decisions (White et al., in press). In short, regions implicated in appropriate social exchange behavior have been found to be dysfunctional in individuals with elevated CU traits.

In line with this, previous behavioral work in clinical and subclinical samples with elevated CU (or psychopathic) traits on social exchange has documented reduced cooperation (Mokros et al., 2008; Rilling et al., 2007). However, no previous work has examined responses to unfair offers or punishment of such offers as a function of CU traits. In this study we used a modified Ultimatum Game to study neural responses to unfair behavior in

healthy children, who varied in their level of reported CU traits. We made three predictions: First, given data suggesting reduced detection of social contract violations in psychopathic individuals (Ermer & Kiehl, 2010), increased CU traits would be associated with reduced modulation of AIC and dACC activity by offer unfairness. Second, given data that patients with vmPFC lesions show greater rejection/punishment of unfair offers (Koenigs & Tranel, 2007) and suggestions that vmPFC dysfunction is seen in youth with CU traits (Blair, 2007), we predicted that CU traits would be associated with increased punishment of unfair offers. Third, given data suggesting reduced modulation of activity within AIC and dACC by negative reinforcement expectancies during decision-making (White et al., in press), increased CU traits would be associated with reduced AIC and dACC to unfair offers.

## METHODS

### PARTICIPANTS

Twenty healthy youth aged 11 to 17. Youth were recruited from the community through advertizing. A statement of informed assent and consent was obtained from participating children and parents. This study was approved by the Institutional Review Board of the National Institutes of Mental Health.

All youth and parents completed Kiddie Schedule for Affective Disorders and Schizophrenia (K-SADS; Kaufman et al., 1997) assessments conducted by a doctoral-level clinician as part of a comprehensive psychiatric and psychological assessment. The K-SADS has demonstrated good validity and inter-rater reliability ( $\kappa > 0.75$  for all diagnoses; Kaufman et al., 1997). By exclusion criteria, subjects did not meet criteria for any K-SAD diagnosis (including Oppositional Defiant Disorder, Conduct Disorder, or Attention Deficit Hyperactivity Disorder). For additional exclusion criteria see White et al. (2012). IQ was measured by the Wechsler Abbreviated Scale of Intelligence (two-subtest form). CU traits were measured by the Inventory of Callous-Unemotional Traits (ICU; Frick, 2004), a 24-item self-report scale of proven construct validity in community and juvenile justice samples (Essau et al., 2006).

The sample contained 13 males and 7 females, who did not significantly differ by age (mean age = 14.15;  $SD = 2.29$ ), IQ (mean IQ = 106.25,  $SD = 14.83$ ), racial breakdown (60% European-American, 25% African-American, 10% Hispanic, 5% Asian-American) or ICU score (mean ICU score = 18.45;  $SD = 6.05$ ; range = 9–30).

**The Social Fairness Game**—Participants were presented with a version of the Ultimatum game (Sanfey et al., 2003); the Social Fairness game (Figure 1). In this game, participants were told that they are playing with a series of partners (while not explicitly stated most participants assumed that these partners were computer based, not human). For each trial, the partners had \$20 and the participant had \$3. During the offer-phase (3000 ms), the partner offered to split the money with the participant either fairly (e.g., \$10 to participant; \$10 to partner) or unfairly (e.g., \$6, \$4, \$2 to participant; \$14, \$16, \$18 to partner). Participants were exposed to 12 partners, who were indicated by name to reduce the probability of learning associations with specific names. The offer-phase was followed by a 500–3500 ms randomly jittered interval. During the decision-phase (4000 ms), the participants indicated, via button press, whether they wished to accept the partners offer and give up their \$3 in exchange (e.g., receive \$10 from the fair offer) or if they wished to spend some of their \$3 and mete out punishment to the partner. Participants could spend \$1 and cost the partner \$7, \$2 and cost the partner \$14, or \$3 and cost the partner \$21. During the outcome-phase (3000 ms), the results, based on the participants' choices, were displayed (e.g., partner gets -\$1, you get \$0). A second 500–3500 ms randomly jittered interval preceded the next trial. Participants were exposed to 120 trials (60 fair trials, 24 \$14/\$6 trials, 24

\$16/\$4 trials, and 12 \$18/\$2 trials) over four 6-minute 52-second runs (see Figure 1). In the offer-phase, activity was modulated by the level of unfairness of the offer (\$10/\$10 = 0, \$14/\$6 = 1, 16/\$4 = 2, \$18/\$2 = 3). In the decision-phase, activity was modulated by the level of punishment meted out by the participant (accept = 0, spend \$1 = 1, \$2 = 2, \$3 = 3).

## MRI PARAMETERS

Participants were scanned using a 3T GE Signa scanner. A total of 165 functional images per run were taken with a gradient echo planar imaging (EPI) sequence (repetition time = 2560 milliseconds; echo time = 27 milliseconds;  $64 \times 64$  matrix;  $90^\circ$  flip angle; 24 cm field of view). Whole-brain coverage was obtained with 46 axial slices (thickness, 2.5mm; .5mm spacing; in-plane resolution,  $3.75 \times 3.75$  mm). A high-resolution anatomical scan (3-dimensional spoiled gradient recalled acquisition in a steady state; repetition time = 7 milliseconds; echo time = 2.984 milliseconds; 24 cm field of view;  $12^\circ$  flip angle; 128 axial slices; thickness, 1.2 mm;  $256 \times 256$  matrix) in register with the EPI data set was obtained covering the whole brain.

## IMAGING DATA PREPROCESSING

Data were analyzed within the framework of the general linear model using Analysis of Functional Neuroimages (AFNI; Cox, 1996). Both individual and group-level analyses were conducted. The first five volumes in each scan series, collected before equilibrium magnetization was reached, were discarded. Motion correction was performed by registering all volumes in the EPI dataset to a volume collected close to acquisition of the high-resolution anatomical dataset.

The EPI datasets for each subject were spatially smoothed (isotropic 6 mm kernel) to reduce variability among individuals and generate group maps. Next, the time series data were normalized by dividing the signal intensity of a voxel at each time point by the mean signal intensity of that voxel for each run and multiplying the result by 100, producing regression coefficients representing percent-signal change.

## GENERAL LINEAR MODEL (GLM) ANALYSIS

The model involved six motion regressors and the following three task regressors: (1) a parametric modulator function, constructed by multiplying an indicator function by offer unfairness (ranging from -2 fair: \$10-\$10 to 2 unfair: \$18-\$2); (2) a parametric modulator function, constructed by multiplying an indicator function by each subject's level of punishment, for each choice component (ranging from -2 no punishment: \$0 response to +2 maximum punishment: \$3 response); (3) a non-modulated regressor modeling the outcome-phase to avoid baseline contamination. Thus, this GLM detected regions modulated by the subject's response to the unfairness of the offer and their punishment level of the partner's offer. All regressors were convolved with a canonical hemodynamic response function (HRF) to account for the slow hemodynamic response.

Linear regression modeling was performed using the 3 regressors described earlier, plus regressors to model a first-order baseline drift function. This produced  $b$  coefficients and associated  $t$  statistics for each voxel and regressor. In accordance with findings that normalization of brain volumes from age 7-8 years onward does not introduce major age-related distortions in localization or time course of the blood-oxygen-level-dependent (BOLD) signal in event-related fMRI (Burgund et al., 2002; Kang, Burgund, Lugar, Petersen, & Schlaggar, 2003), the participants' anatomical scans were individually registered to the Talairach and Tournoux atlas (Talairach & Tournoux, 1988). The individuals' functional EPI data were then registered to their Talairach anatomical scan within AFNI.

## FMRI DATA ANALYSIS

The group analysis of the BOLD data was then performed on regression coefficients from individual subject analyses using a *t*-test versus baseline for both the offer-phase of the task and the decision-phase of the task. Initial thresholding was set at  $p < 0.005$  with an extent threshold of 10 voxels, a combination that has been demonstrated to produce a desirable balance between Type I and Type II error rates (Lieberman & Cunningham, 2009). After observing hypothesized group differences, post-hoc analyses were performed to facilitate interpretations. For these analyses, average percent signal change was measured across all voxels within each region of interest (ROI) generated from the functional masks, and data were analyzed using appropriate follow-up tests within SPSS.

## RESULTS

### BEHAVIORAL RESULTS

An initial repeated measures four-level (offer unfairness: \$10/\$10, \$14/\$6, \$16/\$4, \$18/\$2) ANCOVA was conducted on the choice data, with ICU scores as the covariate. This revealed a significant main effect of offer unfairness,  $F(3, 16) = 16.12, p < 0.001$ ; participants spent more money to punish the partner as offer fairness increased, \$10/\$10  $M = .03$  ( $SD = .06$ ), \$14/\$6  $M = .75$  (.49), \$16/\$4  $M = 1.52$  (.77), \$18/\$2  $M = 2.46$  (.76), and the amount of punishment was significantly greater from one level of unfairness to the next ( $ts -6.81$ — $16.89, ps < .001$ ). No significant effect was observed for the offer unfairness-by-CU interaction,  $F(3, 16) = .513, p = .675$ . A second four-level ANCOVA was conducted on the RT data. This revealed no significant effects for either the main effect of offer unfairness,  $F(3, 16) = .291, p = .832$  or the offer unfairness-by-CU interaction,  $F(3, 16) = .107, p = .747$ .

### FMRI RESULTS

Our initial analyses examined regions showing parametric modulation by offer unfairness and regions showing parametric modulation by decided punishment level.

#### MODULATION BY OFFER UNFAIRNESS

Regions showing significant modulation by offer unfairness included dorsomedial prefrontal cortex (dmPFC), bilateral inferior parietal lobules and right IFC. All of these regions, except right IFC showed increased activity as a function of increased offer unfairness (see Table 1).

#### MODULATION BY PUNISHMENT LEVEL

Regions showing significant modulation by punishment level included bilateral AIC, dACC/dmPFC, bilateral iFC, bilateral dIPFC, ventromedial prefrontal cortex/rostral anterior cingulate (vmPFC/rACC) and posterior cingulate (PCC) cortices (see Table 1). All of these regions, except vmPFC/rACC, showed positive modulation by punishment level; i.e., activity in these regions increased when participants decided to punish the partner more. In contrast, activity within vmPFC/rACC decreased as a function of the participant's decision to increase punishment.

#### THE RELATIONSHIP OF CU TRAITS TO THE RESPONSE TO OFFER UNFAIRNESS AND PUNISHMENT LEVEL

To address our core hypotheses that increased CU traits would be associated with reduced modulation of AIC and dACC activity by offer unfairness and when punishing others, we examined the relationship of CU traits to the average modulation of activity per voxel within these regions, as defined by our functional ROIs. In addition, we examined the relationship



CU traits to modulation of activity to several other regions, as defined by functional ROIs, either because of their importance to models of CU traits (caudate, vmPFC/rACC, PCC) or previous work on social exchange tasks (AIC/iFC). Notably, CU traits were significantly and inversely related to modulated activation as a function of punishment level in bilateral AIC (albeit left anterior insula:  $r = -.441$ ,  $p = .051$ ), dACC, bilateral iFC and right caudate (see Table 2). However, there was only a modest positive trend ( $r = .353$ ,  $p = .127$ ) for the association between CU traits and vmPFC/ rACC activity modulated by punishment level (see Figure 2).

There was no significant relationship between CU traits and modulation by offer unfairness within dmPFC. Indeed, the relationship between CU traits and dmPFC parameter responsiveness (and right AIC and, at trend levels, right caudate) was significantly greater for punishment level than for offer unfairness, Hotelling's  $t(17) = 2.12, 2.26, 2.05$ ;  $p < .05, 0.05$ , and  $0.1$ , respectively.

## DISCUSSION

The current study investigated the neural systems mediating social exchange in youth and critically the extent to which activation within these regions is modulated by level of CU traits. There were two main findings: First, as the unfairness of an offer increased, activity within dACC/dmPFC increased. However, the extent of this increase in activity was not modulated by level of CU traits. Second, as the participant increased punishment level, there were increases in activity within AIC, dACC/dmPFC, bilateral iFC, bilateral dlPFC, and posterior cingulate (PCC) cortices and decreases in activity in vmPFC/rACC. Moreover, the extent of increases within most of these regions as a function of punishment level was modulated by level of CU traits.

The finding that as the unfairness of the offer increased, there was increased activity within dmPFC was consistent with previous research (King-Casas et al., 2008). Up until recently, models of dmPFC activity have stressed its role in conflict resolution or error detection (Botvinick, Cohen, & Carter, 2004). However, these models have faced empirical challenges and recently a new theory, the predicted-outcome model has been proposed (Alexander & Brown, 2011). The model suggests that dmPFC neurons generate signals reflecting predictions of the probability and timing of the various possible outcomes of an action. These predictions are inhibited when the corresponding predicted outcome actually occurs and therefore dmPFC activity is maximal when an expected outcome fails to occur. This new model is particularly attractive here as it overlaps with positions on the response to unfair offers; i.e., that they represent error signals to the social norm that allocations of resources should be equitable—the activity is seen when the expected norm is failed to be adhered to (Montague & Lohrenz, 2007). Suggestions concerning dmPFC and AIC/IFG functioning can be integrated, such that dmPFC is thought to be sensitive to unexpected outcomes (cf. Alexander & Brown, 2011) and AIC/IFG orchestrates potentially necessary changes in behavioral response (cf. Blair & Cipolotti, 2000; Budhani, Marsh, Pine, & Blair, 2007).

Given data that patients with vmPFC lesions show greater rejection/punishment of unfair offers (Koenigs & Tranel, 2007) and suggestions that vmPFC dysfunction is seen in youth with CU traits (Blair, 2007), it was hypothesized that CU traits would be associated with a greater propensity to reject/punish unfair offers. However, this was not seen—perhaps reflecting limited power (see Sharp et al., 2011; for a similar absence of behavioral differences in externalizing youth). Within the BOLD responses, there was significant negative modulation by punishment level in vmPFC/rACC. Considerable work implicates vmPFC in the representation of outcome value (Levy & Glimcher, 2011; O'Doherty, 2004).

On this basis, we suggest that the vmPFC activity in the current task represents the reward value of the chosen option—higher punishments in response to unfair offers represent less reward for the participant. It should be noted though that CU traits showed a very weak relationship with vmPFC activity to punishment level (see above). While this may represent an issue of power, power analyses indicated that an  $N$  of 48 will yield a power of 0.8 given the moderate effect size ( $=0.353$ ). It appears that in the current study CU traits were more closely associated with activity in other regions such as dmPFC, AIC, and caudate as a function of punishment level.

CU traits were not significantly associated with the dmPFC response to unfair offers. This suggests that CU traits are not associated with a diminished sensitivity to the inappropriateness of fairness allocations when they are the victims of this unfair allocation; i.e., youth with CU traits appropriately represent the expectancy that the fairness norm will be adhered to (cf. Alexander & Brown, 2011). However, CU traits were significantly associated with the response within dmPFC, AIC, and caudate (at trend) as a function of administered punishment. This suggests that CU traits are associated with reduced sensitivity to either: (1) the decision going against the individual's financial interest (consistent with previous findings of reduced reward sensitivity in youth with elevated CU traits; Finger et al., 2011; Finger et al., 2008; White et al., in press); or (2) the expectancy that participants should not typically engage in aggressive, punitive responses to others; or (3) both.

Three caveats should be considered with respect to the current results. First, the results should be considered preliminary given the small sample size. While recent neuroimaging studies investigating associations between personality traits and neural activity have also involved  $N$ s of 20 or even less (Sharot, Riccardi, Raio, & Phelps, 2007), it would be advisable for future work to involve larger  $N$ s. Second, the current study cut across a relatively broad developmental range, but did not address issues of relative development or puberty. While CU traits have been shown to be relatively stable across the developmental range of the current sample (Lynam, Caspi, Moffitt, Loeber, & Stouthamer-Loeber, 2007), it will be important to determine whether adolescence has specific effects on the neural expression of CU traits. Third, the participants in this study were all typically developing with no current psychiatric diagnoses. It will be important to see if the current results replicate in a sample of youth with CD and elevated CU traits.

In short, the current data support suggestions of reduced use of expectancy information (either reduced sensitivity to reward or norm expectations of the self's behavior or both) in individuals with elevated CU traits (Blair, 2007). However, they extend these previous hypotheses in two ways. First, by suggesting this expectancy insensitivity is also shown in social decision making; i.e., during social exchange, a core aspect of social cognition. Second, by suggesting that this expectancy insensitivity can be seen in dmPFC and AIC (i.e., regions beyond the amygdala, vmPFC, and caudate; see also White et al., in press). A reduction in the use of expectancy information likely contributes to the profoundly disturbed decision making in this population and their increased use of aggression.

## Acknowledgments

This work was supported by the Intramural Research Program at the National Institute of Mental Health.

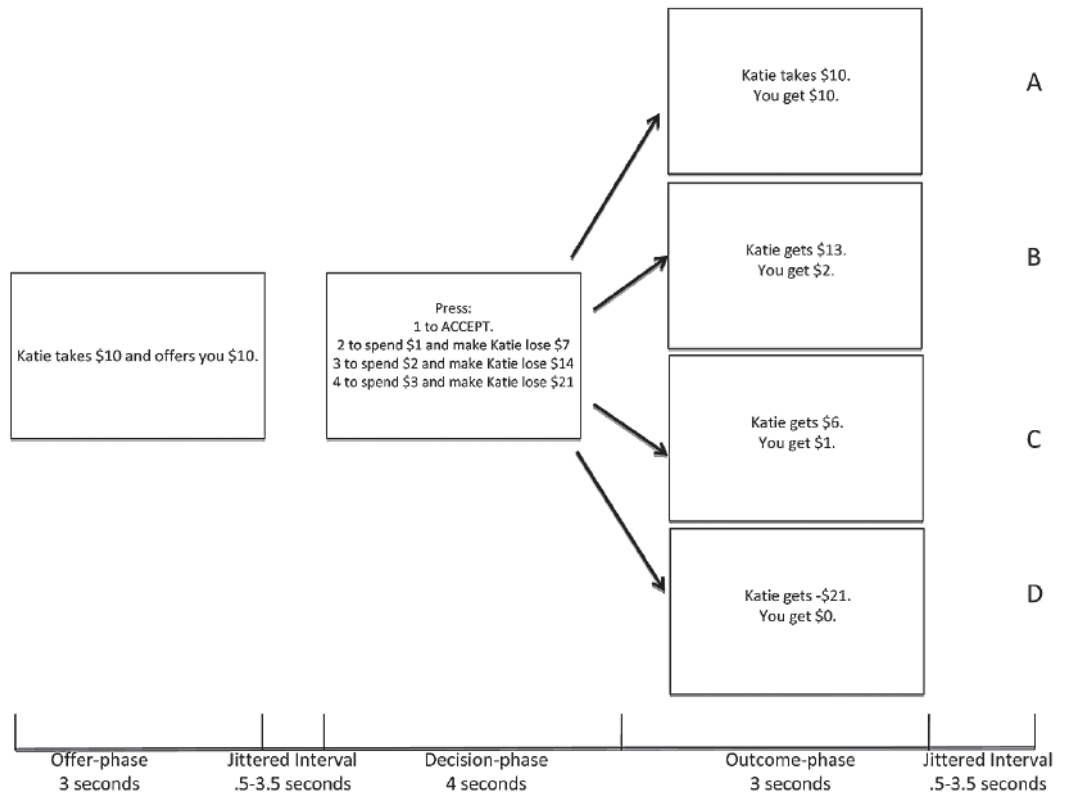
## References

Alexander WH, Brown JW. Medial prefrontal cortex as an action-outcome predictor. *Nature Neuroscience*. 2011; 14(10):1338–1344.

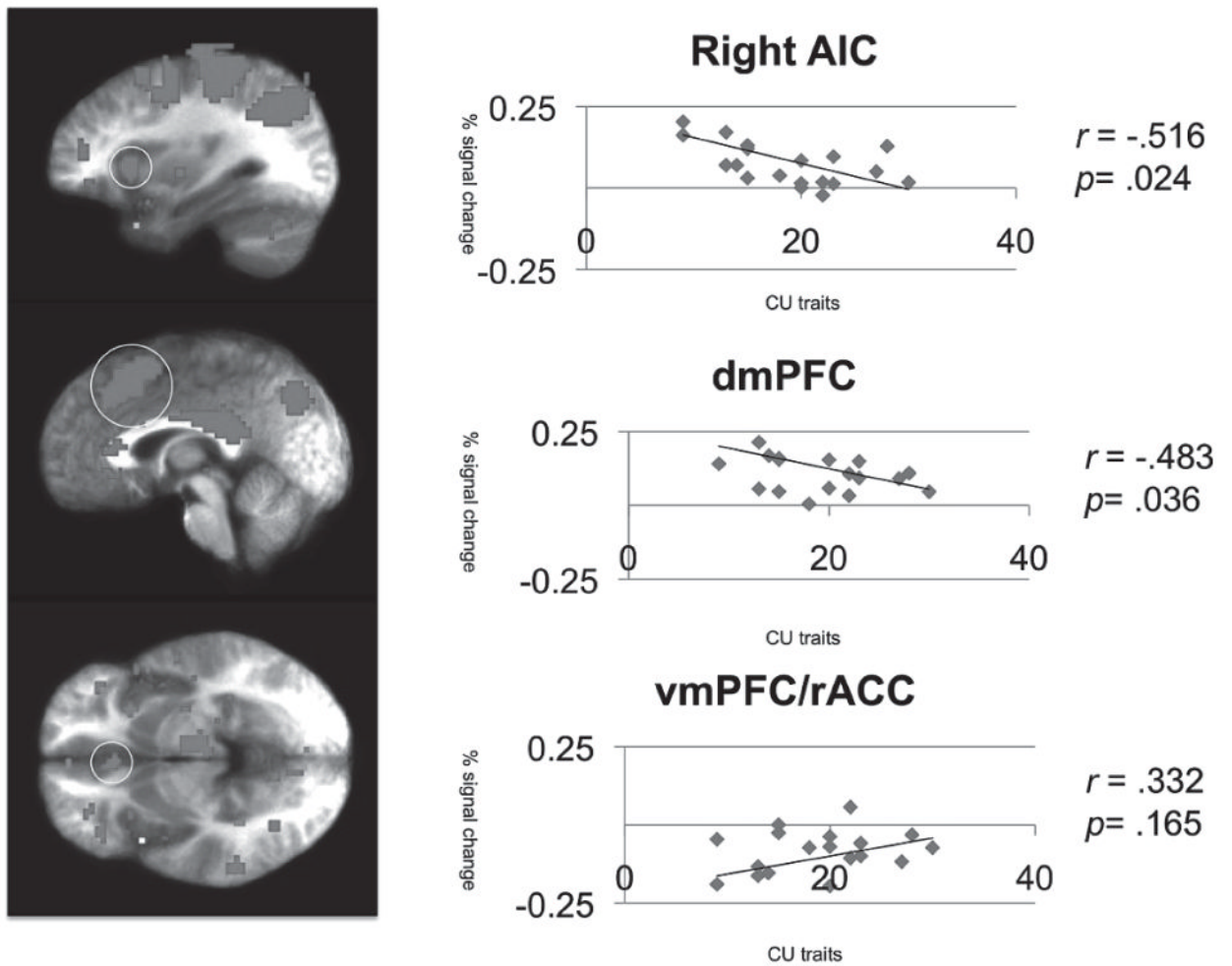
- Blair RJR. The amygdala and ventromedial prefrontal cortex in morality and psychopathy. *Trends in Cognitive Sciences*. 2007; 11:387–392. [PubMed: 17707682]
- Blair RJ, Cipolotti L. Impaired social response reversal. A case of ‘acquired sociopathy’. *Brain*. 2000; 123(Pt 6):1122–1141. [PubMed: 10825352]
- Blair RJR, Colledge E, Mitchell DGV. Somatic markers and response reversal: Is there orbitofrontal cortex dysfunction in boys with psychopathic tendencies? *Journal of Abnormal Child Psychology*. 2001; 29(6):499–511. [PubMed: 11761284]
- Botvinick MM, Cohen JD, Carter CS. Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Science*. 2004; 8(12):539–546.
- Boyd R, Gintis H, Bowles S, Richerson PJ. The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences of the United States of America*. 2003; 100(6):3531–3535. [PubMed: 12631700]
- Budhani S, Marsh AA, Pine DS, Blair RJR. Neural correlates of response reversal: Considering acquisition. *Neuroimage*. 2007; 34(4):1754–1765. [PubMed: 17188518]
- Budhani S, Richell RA, Blair RJ. Impaired reversal but intact acquisition: Probabilistic response reversal deficits in adult individuals with psychopathy. *Journal of Abnormal Psychology*. 2006; 115(3):552–558. [PubMed: 16866595]
- Burgund ED, Kang HC, Kelly JE, Buckner RL, Snyder AZ, Petersen SE, et al. The feasibility of a common stereotactic space for children and adults in fMRI studies of development. *Neuroimage*. 2002; 17(1):184–200. [PubMed: 12482076]
- Cox RW. AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*. 1996; 29(3):162–173. [PubMed: 8812068]
- de Quervain DJ, Fischbacher U, Treyer V, Schellhammer M, Schnyder U, Buck A, et al. The neural basis of altruistic punishment. *Science*. 2004; 305(5688):1254–1258. [PubMed: 15333831]
- Ermer E, Kiehl KA. Psychopaths are impaired in social exchange and precautionary reasoning. *Psychological Science*. 2010; 21(10):1399–1405. [PubMed: 20855897]
- Essau CA, Sasagawa S, Frick PJ. Callous-unemotional traits in a community sample of adolescents. *Assessment*. 2006; 13(4):454–469. [PubMed: 17050915]
- Finger EC, Marsh AA, Blair KS, Reid ME, Sims C, Ng P, et al. Disrupted reinforcement signaling in the orbitofrontal cortex and caudate in youths with conduct disorder or oppositional defiant disorder and a high level of psychopathic traits. *American Journal of Psychiatry*. 2011; 168(2):152–162. [PubMed: 21078707]
- Finger EC, Marsh AA, Mitchell DG, Reid ME, Sims C, Budhani S, et al. Abnormal ventromedial prefrontal cortex function in children with psychopathic traits during reversal learning. *Archives of General Psychiatry*. 2008; 65(5):586–594. [PubMed: 18458210]
- Frick. The inventory of callous/unemotional traits. Unpublished Ratings Scale; University of New Orleans: 2004.
- Guroglu B, van den Bos W, van Dijk E, Rombouts SA, Crone EA. Dissociable brain networks involved in development of fairness considerations: Understanding intentionality behind unfairness. *Neuroimage*. 2011; 57(2):634–641. [PubMed: 21554961]
- Kang HC, Burgund ED, Lugar HM, Petersen SE, Schlaggar BL. Comparison of functional activation foci in children and adults using a common stereotactic space. *Neuroimage*. 2003; 19(1):16–28. [PubMed: 12781724]
- Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, et al. Schedule for Affective Disorders and Schizophrenia for School-Age Children–Present and Lifetime version (K-SADS-PL): Initial reliability and validity data. *Journal of the American Academy of Child & Adolescent Psychiatry*. 1997; 36(7):980–988. [PubMed: 9204677]
- King-Casas B, Sharp C, Lomax-Bream L, Lohrenz T, Fonagy P, Montague PR. The rupture and repair of cooperation in borderline personality disorder. *Science*. 2008; 321(5890):806–810. [PubMed: 18687957]
- Koenigs M, Tranel D. Irrational economic decision-making after ventromedial prefrontal damage: Evidence from the ultimatum game. *Journal of Neuroscience*. 2007; 27(4):951–956. [PubMed: 17251437]



- Levy DJ, Glimcher PW. Comparing apples and oranges: Using reward-specific and reward-general subjective value representation in the brain. *Journal of Neuroscience*. 2011; 31(41):14693–14707. [PubMed: 21994386]
- Lieberman MD, Cunningham WA. Type I and type II error concerns in fMRI research: Re-balancing the scale. *Social Cognitive and Affective Neuroscience*. 2009; 4(4):423–428. [PubMed: 20035017]
- Lynam DR, Caspi A, Moffitt TE, Loeber R, Stouthamer-Loeber M. Longitudinal evidence that psychopathy scores in early adolescence predict adult psychopathy. *Journal of Abnormal Psychology*. 2007; 116(1):155–165. [PubMed: 17324026]
- Marsh AA, Finger EC, Mitchell DG, Reid ME, Sims C, Kosson DS, et al. Reduced amygdala response to fearful expressions in children and adolescents with callous/unemotional traits and disruptive behavior disorders. *American Journal of Psychiatry*. 2008; 165(6):712–720. [PubMed: 18281412]
- Mokros A, Menner B, Eisenbarth H, Alpers GW, Lange KW, Osterheider M. Diminished cooperativeness of psychopaths in a prisoner's dilemma game yields higher rewards. *Journal of Abnormal Psychology*. 2008; 117(2):406–413. [PubMed: 18489216]
- Montague PR, Lohrenz T. To detect and correct: Norm violations and their enforcement. *Neuron*. 2007; 56(1):14–18. [PubMed: 17920011]
- O'Doherty J. Reward representations and reward-related learning in the human brain: Insights from neuroimaging. *Current Opinion in Neurobiology*. 2004; 14(6):769–776. [PubMed: 15582382]
- Rilling JK, Glenn AL, Jairam MR, Pagnoni G, Goldsmith DR, Elfenbein HA, et al. Neural correlates of social cooperation and non-cooperation as a function of psychopathy. *Biological Psychiatry*. 2007; 61(11):1260–1271. [PubMed: 17046722]
- Rilling JK, Goldsmith DR, Glenn AL, Jairam MR, Elfenbein HA, Dagenais JE, et al. The neural correlates of the affective response to unreciprocated cooperation. *Neuropsychologia*. 2008; 46(5):1256–1266. [PubMed: 18206189]
- Sanfey AG, Rilling JK, Aronson JA, Nystrom LE, Cohen JD. The neural basis of economic decisionmaking in the ultimatum game. *Science*. 2003; 300(5626):1755–1758. [PubMed: 12805551]
- Sharot T, Riccardi AM, Raio CM, Phelps EA. Neural mechanisms mediating optimism bias. *Nature*. 2007; 450(7166):102–105. [PubMed: 17960136]
- Sharp C, Burton PC, Ha C. "Better the devil you know": A preliminary study of the differential modulating effects of reputation on reward processing for boys with and without externalizing behavior problems. *European Child & Adolescent Psychiatry*. 2011; 20(11–12):581–592. [PubMed: 22038344]
- Spitzer M, Fischbacher U, Herrnberger B, Gron G, Fehr E. The neural signature of social norm compliance. *Neuron*. 2007; 56(1):185–196. [PubMed: 17920024]
- Strobel A, Zimmermann J, Schmitz A, Reuter M, Lis S, Windmann S, et al. Beyond revenge: Neural and genetic bases of altruistic punishment. *Neuroimage*. 2011; 54(1):671–680. [PubMed: 20673803]
- Talairach, J.; Tournoux, P. Co-planar stereotaxic atlas of the human brain. Stuttgart: Thieme; 1988.
- White S, Pope K, Sinclair S, Fowler K, Brislin S, Williams WC, et al. Disrupted prediction error signaling during a passive avoidance task in youth with disruptive behavior disorders and psychopathic traits. *American Journal of Psychiatry*. in press.
- White SF, Marsh AA, Fowler KA, Schechter JC, Adalio C, Pope K, et al. Reduced amygdala response in youths with disruptive behavior disorders and psychopathic traits: Decreased emotional response versus increased top-down attention to nonemotional features. *American Journal of Psychiatry*. 2012; 169(7):750–758.



**FIGURE 1.**  
 The Social Fairness Game: The participant was offered a portion of a \$20 pot and then was required to indicate, via button press, whether he or she would accept the offer (A) or punish the individual making the offer. For every \$1 punishment money spent, the individual making the offer lost \$7 (B/C/D).

**FIGURE 2.**

Regions showing modulation by level of punishment versus baseline and the relationship of this with CU traits. CU = Callous-Unemotional; AIC = anterior insula cortex; vmPFC/rACC = ventromedial prefrontal cortex/rostral anterior cingulate cortex.

Table 1

Brain Regions Demonstrating Differential Modulation of BOLD Responses by Chosen Punishment Level During the Offer-Phase Versus Baseline and the Decision-Phase Versus Baseline

Region <sup>a</sup>	Coordinates of Peak activation <sup>b</sup>							
	Left/Right	BA	x	y	z	F (df = 1, 19)	p	Voxels
Offer-Phase								
Dorsomedial frontal cortex	Right	6	7.5	10.5	44.5	4.650	<.001	21
Inferior frontal	Right	45/47	49.5	28.5	2.5	4.441	<.001	11
Medial frontal	Left	6	-16.5	4.5	53.5	4.382	<.001	15
Posterior cingulate cortex	Right	31	16.5	-61.5	20.5	3.863	<.001	11
Superior temporal gyrus	Left	22	-52.5	-13.5	-3.5	5.240	<.001	31
Superior temporal gyrus	Right	38	43.5	10.5	-15.5	4.102	<.001	10
Inferior parietal lobule	Left	39/40	-46.5	-43.5	50.5	5.562	<.001	341
Inferior parietal lobule	Right	40	46.5	-46.5	50.5	5.911	<.001	105
Middle insula	Right	13	37.5	-10.5	11.5	3.830	<.005	27
Decision-Phase								
Dorsomedial frontal cortex	Right	32	4.5	7.5	47.5	7.791	<.001	398
Dorsolateral frontal cortex	Right	10	43.5	16.5	35.5	6.888	<.001	575
Dorsolateral frontal cortex	Left	46	-40.5	1.5	29.5	6.773	<.001	375
Orbital frontal cortex	Left	46/10	-28.5	43.5	11.5	4.999	<.001	63
Orbital frontal cortex	Right	11/47	40.5	34.5	-3.5	4.419	<.001	17
Inferior frontal cortex	Left	47/11	-37.5	34.5	-3.5	3.867	<.001	13
Rostral anterior cingulate cortex	Left	32	-13.5	31.5	2.5	4.426	<.001	39
Posterior cingulate cortex	Right	23	1.5	-28.5	26.5	9.566	<.001	273
Anterior insula	Left	13	-43.5	13.5	2.5	5.542	<.001	101
Anterior insula	Right	47	28.5	19.5	2.5	5.696	<.001	55
Caudate	Right		10.5	10.5	8.5	4.921	<.001	70
Thalamus/caudate/periaqueductal gray	Left		-13.5	-22.5	5.5	10.20	<.001	397
Postcentral gyrus/insula/cingulate	Right	3/13/31	7.5	-25.5	41.5	8.749	<.001	1196
Unsula	Left	13	-37.5	-25.5	20.5	4.609	<.001	16
Middle temporal gyrus	Left	21	-55.5	-31.5	-6.5	3.899	<.001	11
Parahippocampal gyrus	Left	36	-37.5	-28.5	-9.5	3.481	<.005	10

Region <sup>a</sup>	Coordinates of Peak activation <sup>b</sup>						Voxels	
	Left/Right	BA	x	y	z	F (df = 1, 19)		P
Parietal cortices/precuneus	Left & Right	40/39/7	-37.5	-55.5	47.5	9.163	<.001	2747
Cerebellar tonsil	Right		22.5	-46.5	-21.5	6.644	<.001	122
Culmen	Left		-13.5	-52.5	-15.5	6.150	<.001	95
Cuneus	Right	17	22.5	-85.8	14.5	4.956	<.001	45
Cuneus	Left	17	-19.5	-91.5	17.5	4.971	<.001	43

Notes.

<sup>a</sup> According to the Talairach Daemon Atlas (<http://www.nitrc.org/projects/tal-daemon/>).

<sup>b</sup> Based on the standard brain template of the Montreal Neurological Institute (MNI).

BA = Brodmann's Area, *df* = degrees of freedom.



**Table 2**

Bivariate Correlations Between Callous-Unemotional Traits Scores and Activity During Decision-Making in Regions Showing Modulation of Activity by Amount of Punishment During Decision-Making Chosen

<b>Region</b>	<b><i>r</i></b>	<b><i>p</i></b>
Offer-phase		
Dorsomedial frontal cortex	.180	.447
Decision-phase		
Dorsomedial frontal cortex	-.453	.045
Right anterior insula cortex	-.539	.014
Left anterior insula cortex	-.441	.051
Right inferior frontal cortex	-.570	.009
Left inferior frontal cortex	-.557	.011
Right caudate	-.487	.029
Left thalamus/caudate	-.366	.113
Posterior cingulate cortex	-.318	.171
RostralACC/ventromedial prefrontal cortex	.353	.127