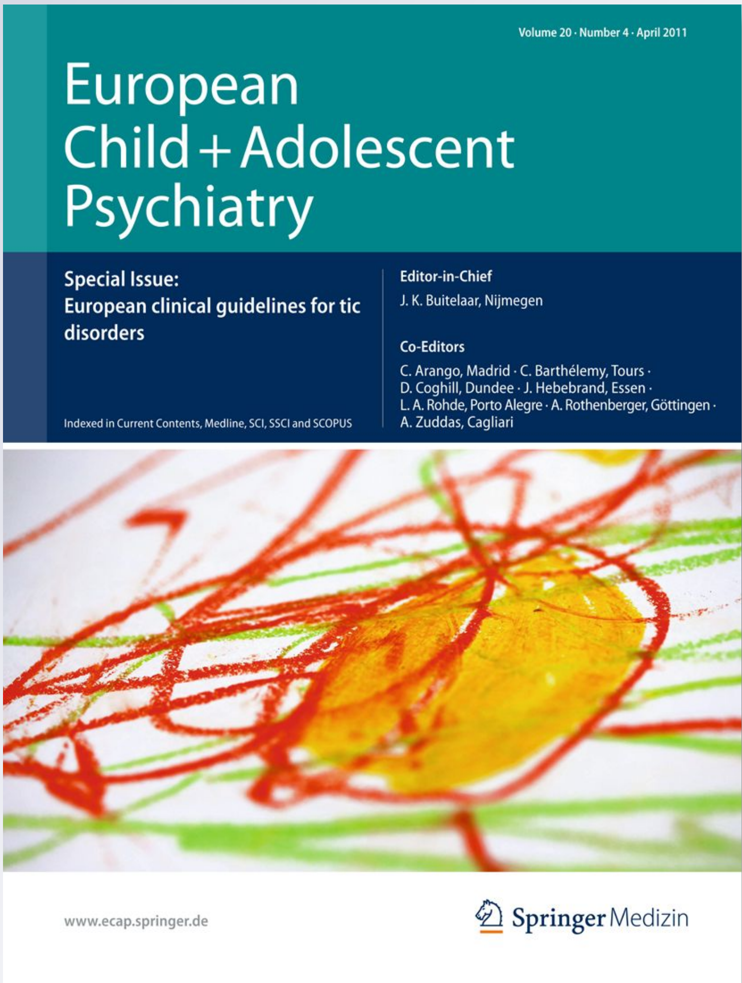
“Better the devil you know”: a preliminary study of the differential modulating effects of reputation on reward processing for boys with and without externalizing behavior problems

Carla Sharp, Philip C. Burton & Carolyn Ha

European Child & Adolescent Psychiatry

ISSN 1018-8827
Volume 20
Combined 11-12

Eur Child Adolesc Psychiatry (2011)
20:581-592
DOI 10.1007/s00787-011-0225-x



Eur Child Adolesc Psychiatry (2011) 20:581–592
DOI 10.1007/s00787-011-0225-x

ORIGINAL CONTRIBUTION

"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

Carla Sharp · Philip C. Burton · Carolyn Ha

Received: 4 February 2011 / Accepted: 8 October 2011 / Published online: 27 October 2011

Abstract Very little is known about the neurobiological correlates of reward processing during social decision-making in the developing brain and whether prior social and moral information (reputations) modulates reward responses in youth as has been demonstrated in adults. Moreover, although externalizing behavior problems in youth are associated with deficits in reward processing and social cognition, a real-life social interaction paradigm using functional neuroimaging (fMRI) has not yet been applied to probe reward processing in such youth. Functional neuroimaging was used to examine the neural correlates of reward-related decision-making during a trust task in two samples of age-matched 11 to 16-year-old boys: with ($n = 10$) and without ($n = 10$) externalizing behavior problems. The task required subjects to decide whether to share or keep monetary rewards from partners they themselves identified during a real-life peer sociometric procedure as interpersonally aggressive or kind (vs. neutral). Results supported the notion that prior social and moral information (reputations) modulated reward responses in the adolescent brain. Moreover, boys with externalizing problems showed differential activation in the bilateral insula during the decision phase of the game as well as the caudate and anterior insula during the outcome phase of the game. Similar activation in adolescents in response to reward related stimuli as found in adults suggests some developmental continuity in corticostriatal circuits. Group differences are interpreted with caution given the small group sizes in the current study. Notwithstanding this limitation, the study provides preliminary evidence for anomalous reward responses in boys with externalizing behavior problems, thereby providing a possible biological correlate of well-established social-cognitive and reward-related theories of externalizing behavior disorders.

Keywords fMRI · Reward · Social decision-making · Externalizing behavior problems · Adolescents

Introduction

Externalizing behavior problems refer to a broad range of disruptive antisocial behaviors as captured by the diagnoses of conduct disorder and oppositional defiant disorder [3]. One of the hallmark features of externalizing problems is interpersonal difficulties with peers [34]. Social-cognitive (e.g. [13] and reward-processing (see [14] approaches provide useful frameworks for understanding the relation between externalizing problems and interpersonal difficulties. Despite known neurobiological correlates for social cognition and reward processing, very little neuroimaging work has been conducted to investigate the neurobiological correlates of reward-related decision-making in social contexts as it relates to externalizing behavior problems in youth. The handful of functional neuroimaging (fMRI) studies focusing on externalizing problems in youth have mostly examined the neural correlates of emotional arousal. Given that fMRI studies of antisocial behavior in adults (the long-term outcome of externalizing behavior problems in youth) have indicated deficits in both social cognition

C. Sharp (✉) · C. Ha
Department of Psychology, University of Houston, 126 Heyne Building, Houston, TX 77204, USA
e-mail: csharp2@uh.edu

P. C. Burton
University of Minnesota, Minneapolis, USA

and reward processing [27], there is a need to study reward-related decision-making in social contexts in children with and without externalizing problems.

The emerging interdisciplinary field of neuroeconomics offers a new approach to operationalizing reward-related decision-making in real-time, real-life, emotionally charged interaction through economic exchange games in controlled settings [21, 29]. Reward-related decision-making in social contexts, as defined in neuroeconomics, assumes that the basic building blocks of decision-making that underlie the process of learning and valuation also are important for decision-making in social contexts [23]. However, greater uncertainty may accompany decision-making in social contexts given the fact that such decision-making affects and is affected by the behaviors of relatively unpredictable conspecifics, resulting in potentially more sophisticated learning algorithms in social decision-making [9]. The meaning of "reward" in this context relates to its meaning in economics, where the subjective desirability of a particular choice is quantified by its utility function. As such, our understanding of the meaning of rewards in this context is guided by reinforcement learning theory, which is the dominant paradigm for studying the neural basis of decision-making in social and non-social contexts. It is, however, important to note that within the context of social decision-making, the utility function of a particular action may be valued not only in terms of monetary gain, but also in terms of social or psychological gains (for instance, feelings of well-being, moral obligation, status, self-interest, fairness, punishment or altruism). Whether valued in terms of monetary or social/psychological gain, rewards are encoded in the mesolimbic dominergic system that underlies reinforcement learning, most notably the striatum, insula and orbital frontal cortex [23].

A neuroeconomic game which has been used in relation to psychopathology [20] is the trust task [8]. The trust task is based on an evolutionary model that draws on game theory principles. One player (the Investor) is endowed with a certain amount of money (or points as proxies for money). The Investor can keep all the money or decide to 'invest' some amount with the partner (the Trustee). The amount invested is tripled in value as it is sent to the Trustee, who then decides what portion to return to the Investor. Recently, we used the trust task to examine differences in trust (Investor's initial offer) and reciprocity (Trustee's relative return offer) between boys with and without externalizing behavior problems at a behavioral level [29]. A trust game was played under two conditions: an anonymous version where the identity of the trust game partner was not known and a "known identity" version where identities were revealed prior to the game. Results showed that whereas the known identity condition of the task increased reciprocity for normally functioning boys, the opposite was true for boys with externalizing behavior. Moreover, reduced reciprocity was associated with hostile intentions, but not reflective of a general theory of mind deficit. Our findings for normally functioning boys fit with research in normally functioning adults which shows that knowing the identity of partners attenuates aggressive acts to promote trust and reciprocity during the trust game [9]. It also demonstrated for the first time reduced reciprocity in boys with externalizing behavior problems.

In the current study, we build on these findings by exploring the possibility that anomalies in trust behavior in externalizing boys especially with known partners may be explained by differential modulating effects of the identity of partners, here operationalized as reputation. To this end, we used a trust task adapted for fMRI by Delgado et al. [12] who examined the modulating role of prior moral and social information (reputations) on reward responses in normally functioning adults in a multi-round trust task. Subjects played an iterated version of the trust task with three hypothetical partners, each of a different moral character (good, bad, neutral). Fictitious scripts about the moral character were provided to subjects prior to the game that were consistent with each moral character. In each round, the subject was given \$1 and had to decide whether to give or keep the \$1 from their hypothetical partners. If the subject decided to send \$1 to their partner, the \$1 was tripled and the partner received \$3. The partner then had the opportunity to return \$1.50 to the subject. Unbeknownst to subjects, the game was rigged so that all partners, regardless of their moral character, would receive return offers only 50% of the time. Behavioral results showed that despite the knowledge that game partners did not actually make real decisions during the game, the experimental manipulation was effective such that trust ratings for the morally good partner were significantly higher than for the morally neutral character, followed by the morally bad character. Brain results showed that the task activated the corticostriatal loops often associated with neuroeconomic games, including the anterior cingulate cortex (ACC), caudate and the insula. Moreover, differential brain activation was observed in these areas depending on the type of partner, thereby showing that prior social and moral information about others (reputation) modulated reward responses in the brain. More specifically, for the outcome phase of the game, repeated-measures ANOVA using mean beta weights extracted from the caudate nucleus ROI showed a significant interaction between moral character (good and bad vs. neutral) and outcome (positive vs. negative feedback). For the decision phase of the game, activation in the anterior cingulate and insula was associated with bias-incongruent decisions (i.e., share with the bad partner and keep with the good partner), while riskier decision-making (with neutral partners vs. good/bad partners) was associated with activation in the striatum. The fact that stronger activation was associated with the neutral partners for both outcome and decision phases of the game highlights an important feature of reward processing that fits with other research on learning and valuation in social contexts [23]: when playing with neutral partners, there is a greater degree of uncertainty (and risk) involved in social decision-making which associates with stronger learning signals in the brain, compared to playing with partners with known reputations (bad or good partners).

We adapted the above task for use in a sample of boys since boys are more affected by externalizing behavior problems than girls, with a boy–girl ratio of 3:1 to 5:1 increasing in favor of girls with development [7]. We purposefully recruited adolescents from community youth groups (Boy Scouts) to exploit the fact that children recruited from scout troops have already established reputations (models) of each other. Instead of hypothetical descriptions of good, bad and neutral partners, we made use of standard peer nomination methodology in developmental psychology [4] whereby peers nominated a mean and kind peer in their scout troops. As such, the current study aimed to offer a more ecologically valid experimental approach, thereby being the first study in social neuroscience to make use of actual reputations versus hypothetical reputations.

The original Delgado paradigm was furthermore adapted by lengthening the decision and outcome phases to allow for slower processing speed in youngsters. Finally, the original paradigm was developmentally adapted by increasing the return offer from partners from \$1.5 to \$2. This was done to maximize the probability that subjects would in fact share/invest with their partners, which was necessary to create comparable experimental contrasts. Concerns about boys not sharing at all were based on the studies demonstrating age-related changes in sharing with development. For instance, Van den Bos et al. [32] examined the role of perspective taking in trust game behavior by defining perspective taking as the ability to consider the intentions of others as well as the consequences of subjects' own trust game behavior for others. By varying these outcomes (risk vs. benefit) across different age groups (9–25 years), the authors demonstrated age-related changes in sensitivity to outcome for partners as indexed by subjects' behavioral choices. Specifically, they found age-related changes in sensitivity to the benefit of the other player in trust decisions.

First, on the basis of previous fMRI work in children and adolescents reviewed by Fareri et al. [15] suggesting continuity in the neural correlates of reward processing through development, we expected that the same corticostriatal loops would be activated by the decision and outcome phases of the game as found by Delgado et al. [12] in adults. That is, increased BOLD responses would be observed in the ACC and insula during the decision-making phase of the game when playing the neutral partner relative to the kind and mean partners due to the uncertainty involved in "the unknown". Similarly, increased BOLD responses were expected in the caudate nucleus in trials with the neutral (vs. known reputation) partners for the outcome phase of the game.

Second, we expected that reputations (task conditions) would have differential effects on reward responses for boys with and without externalizing behavior problems. This expectation was based on prior work demonstrating general anomalies in trust behavior in these boys [29], as well as findings demonstrating that children with externalizing problems have deficits in moral reasoning [22], social cognition [13] and reward processing, using both behavioral [14] and neural paradigms [17]. Our focus was on brain areas known to be associated with reward-related decision-making in social contexts: the insular cortex [12, 30] and the striatum [12, 16]. More specifically, because children with externalizing disorders have been shown to be less prosocial [18], we hypothesized that externalizing boys would show increased activity in brain areas associated with aversive experiences (insula) when making share decisions during the game. For the outcome phase of the task, we expected externalizing boys to be less sensitive to the reputation of partners. While we expected modulation of reward responses by reputations for normal controls, we expected absence of modulation of reward responses in the anterior insula for externalizing boys. This expectation was based on prior work showing reduced insula activation in borderline personality disorder that may indicate general insensitivity in integrating social information for optimal social decision-making associated with psychiatric disorders [20].

Methods

Participants

Subjects ($n = 20$; ages 11–16) were selected from a larger study of social cognition and externalizing behavior problems [29]. Subjects in the larger study ($n = 171$) of male youth were recruited from community youth groups (Boy Scouts). Groups met weekly for activities and boys had known each other for an average of 2.3 years (SD = 1.4 years). Boy Scouts were purposefully recruited to exploit the fact that these children have already established reputations (models) of each other. Positive consent and assent were obtained from parents and children, respectively. Measures of externalizing behavior problems, information on fMRI exclusion criteria (left-handedness, metal, substance use problems, psychotropic medication) and demographic information were obtained individually for the full sample. This information was then used to select and scan two age- and SES-matched groups (total $n = 29$) who differed on measures of externalizing behavior disorders. Nine boys were excluded for excessive movement, resulting in a final sample with an overall mean age of 12.70 (SD = 1.71): those with externalizing problems ($n = 10$) and without externalizing problems ($n = 10$).

Measures

Externalizing behavior problems

It is well known that different informants may validly contribute different information regarding psychopathology [33]. To this end, we combined three measures of externalizing behavior to include youth self-report, parent-report, and peer nominations.

Youth self-report and parent-report The Youth Self Report (YSR; [1] and Child Behavior Checklist (CBCL; [1] are well established evidence-based assessment instruments [19] that assess global and specific psychopathology among youth aged 6–18 years during the past six months. For the current study, we used the recommended *T*-score of 65 on the Externalizing subscale to identify boys in the externalizing group. Prior research indicates this threshold discriminates well between clinical and non-clinical populations [1].

Peer nomination as relationally aggressive or pro-social A peer nomination instrument developed by Werner and Crick [35] was used to assess relational aggression and pro-social behavior. The measure consists of 24 items of which seven tap into a relational aggression subscale, shown to have high reliability (Cronbach's alpha = 0.87). Nine items tap into a pro-social behavior subscale and have also been shown to be highly reliable (Cronbach's alpha = 0.91; [35]. The peer nomination instrument was administered at the same time as the collection of the youth self-report and parent-report data on the full sample ($N = 171$) from which subjects for the fMRI study was selected. Participants were provided with a group membership roster and were instructed to nominate up to five peers who best fit each description. The number of nominations each participant received from his or her peers was then summed for each item and totaled for each subscale. To aid in the identification of boys who were relationally aggressive versus those who were prosocial, we used the 50th percentile to identify boys above and below the median for relationally aggressive and pro-social nominations. As is often the case with sociometric studies [10], many boys were nominated as both pro-social and relationally aggressive, so several boys fell above the 50th percentile for both subscales. We were interested in identifying boys who were perceived as 'extreme' indicated by being nominated as "only pro-social" or "only interpersonally aggressive". Of the $N = 171$ boys, 38 were identified as "pro-social only" and 23 were identified as "relationally aggressive only." The mean number of nominations as relationally aggressive was 1.40 (SD = 1.97) with the maximum number of nominations being 13 and the minimum 0.

Combined measure of externalizing behavior problems Since normality assumptions for externalizing behavior disorder variables were violated (possibly due to the fact that the sample was drawn from the community), all variables were submitted to a normalizing transformation (*z*-scores). A boy was considered to meet criteria for externalizing behavior problems if he was above cut-off for parent- and self-report externalizing problems, and described as "only relationally aggressive." This variable in combination with fMRI exclusion criteria was used to identify boys in the externalizing group. As expected, means for both self-reported ($t = -2.39$; $df = 18$; $p = 0.02$) and parent-reported ($t = -2.46$; $df = 18$; $p = 0.02$) externalizing behavior problems were significantly higher in the externalizing group versus the non-externalizing group. On the day of scanning, the CBCL was administered again to confirm stability of group status. Differences between means were even more significant on the day of scanning ($t = -4.224$; $df = 18$; $p = 0.001$).

Experimental design

The fMRI task was adapted from Delgado and colleagues [12] and is represented visually in Fig. 1. Boys were asked to imagine that they were playing with three types of partners. Each of the three types of partner was represented by a photograph of the partner shown on trials during which play with that partner took place. Two of the partners were known to the participant. One 'known' partner was a boy that the participant identified during the peer nomination session as the most "mean" (relationally aggressive) child. The other 'known' partner was the peer that the participant identified during peer nomination as the most "kind" (prosocial) child. The third partner was a stranger with no known social or moral information attached to the child (neutral). On the day of the scanning, subjects were reminded of their peer nominations and the reasons for nominating peers as either kind or mean.

Participants played 60 trials randomly ordered per reputation type. Trials were divided into a decision phase and an outcome phase. During the decision phase (4 s), a child viewed the name and face of the partner and the options to

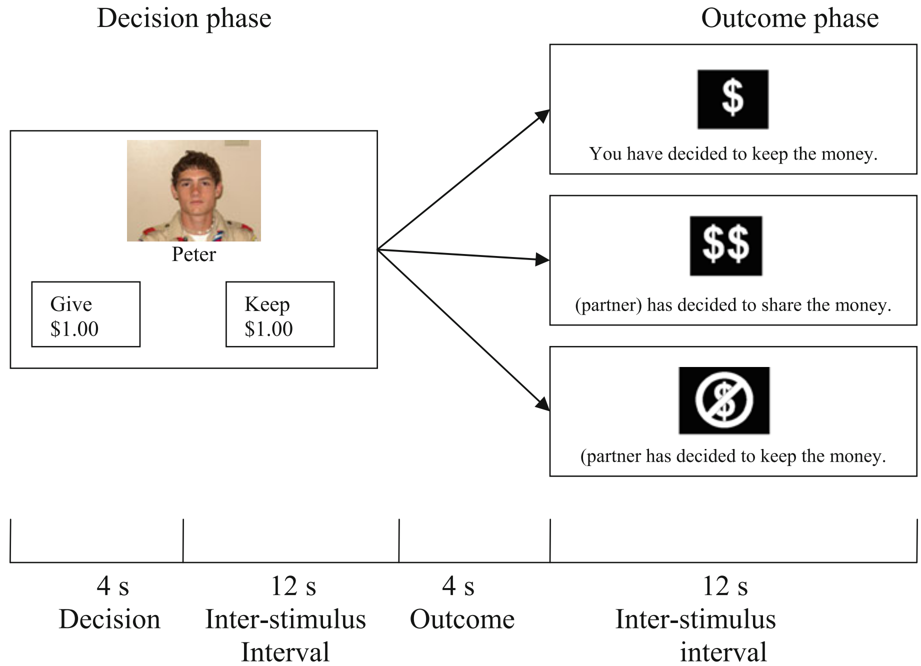


Fig. 1 Experimental task

keep or share $1. Participants knew that an investment of $1 could be reciprocated by a $2 return from their partner. A 12-s interval followed. During this interval, the screen was blank with a single fixation point. During the interval phase, the partner "made a decision" as whether to send money back ($2) to the participant. During the outcome phase (4 s), three child friendly outcomes of the trial were displayed to the participant denoted by symbols that they learned during the behavioral trial run before the actual experiment. The partner's 'behavior' was, in actuality, manipulated so that he shared 50% of the time regardless of his status as mean, kind or neutral. Prior to the experiment, participants were given individualized instructions, and a trial run was played on the computer outside of the scanner. Participants were aware that they were not playing a 'live' partner, but they were not aware that return offers had been rigged so that only 50% of share offers would be returned.

The instructions were as follows: For the next 30 min, you are going to imagine that you play a game with three other kids. Two of these kids (Peter and John) will be known to you because they were in the same scout troop as you. The third one (Kevin) will be a stranger to you. Therefore, you do not know Kevin. You nominated Peter as being kind to others. The reason you gave for nominating him as kind is that he always helps others. You nominated John as being mean to others. The reason you gave for nominating him as mean is that he always fights with other children. You do not know Kevin. So you do not know whether he is mean or kind to others. This game has many rounds. In each round, a computer screen will show you a photograph of the boy you are playing with in that round. His name will be printed below his photograph. You will be asked whether you would like to give $1 to the boy you are playing with or keep $1. You have only 4 s to decide. Press the button in your right hand for KEEP, or press the button in your left hand to GIVE. Once the screen goes blank you cannot make your decision anymore. So make your decision as soon as possible after you get the option to KEEP or GIVE. If you decide to give the dollar to the other boy, it will be tripled and he will receive $3. He will then be given the chance to give back $2 to you, or keep the $3 he received. You will see the decisions made in that round on the next computer screen for only 4 s.

Neuroimaging procedures and data analyses

Each participant's imaging data were acquired using one of two identical Siemens Allegra 3T scanners (software version Syngo MR 2002B, Erlangen, Germany). The scanners undergo extensive Quality Assurance procedures by in-house fMRI physicists and engineers as well as Siemens to ensure equivalence of data obtained from the scanners. Foam padding was used to minimize participants' head movement. High-resolution structural images (MPRage pulse sequence; TR = 1,200 ms, TE = 2.93 ms, flip angle = 12°, FOV = 214.4 mm, image matrix = 448 × 512, voxel size = 0.48 × 0.48 × 1 mm, 192 axial slices were acquired at the beginning of each session. Functional images were obtained using an echo-planar imaging (EPI) sequence (TR = 2,000 ms, TE = 40 ms, flip angle = 90°, FOV = 220 mm, matrix = 64 × 64, voxel size 3.4 × 3.4 × 4 mm, 26 axial slices, angled to align the anterior and posterior commissures), providing whole brain coverage. Each participant completed one functional run, with a 32:04 duration (two "dummy" volumes were acquired at the beginning of the run to allow steady-state equilibrium to be reached, followed by 960 volumes of data acquisition while the task was performed).

All image preprocessing and data analyses were conducted in AFNI [11]. To correct for slice acquisition time differences within each functional volume, voxels within each slice were temporally shifted (using Fourier interpolation) into alignment in time with the average of all time points within the TR. To correct for head movement in each participant, all volumes within the functional scan were spatially aligned with the functional volume acquired nearest in time to the structural scan (the first volume of the functional scan for most participants) using an iterated linearized least squares algorithm. Data were spatially smoothed by applying a 4 mm full width at half maximum (FWHM) Gaussian blur to each EPI volume.

Data from each participant were analyzed within the framework of a general linear model using AFNI [11]. In addition to the 12 experimental conditions, the six translation and six rotation parameters from the motion correction procedure were included in the model as regressors of no interest to reduce further the impact of participant head movement on the analyzed data. Deconvolution (Glover 1997) was used to estimate the impulse response function (IRF) over a seven-TR (12 s) window following stimulus onset for each condition at each voxel for each participant, with no assumptions about the shape of the response function. Thus, a beta weight was estimated at 2-s intervals for each of seven time points beginning at the onset of the each phase of the trial (Decision/Outcome) for each decision-partner combination (i.e., Keep/Share, Kind/Mean/Neutral) for the decision phase and for each outcome-partner combination (i.e., Positive/Negative, Kind/Mean/Neutral) for the outcome phase.

The mean of the third time point (i.e., 6 s following decision or outcome onset, roughly the point at which the typical hemodynamic response function peaks) and the two flanking time points were averaged for each of the time courses associated with both the Decision and Outcome phases. These means and the time courses from which they were derived were then transformed to Talairach space [31] using a 12-point affine transformation and used as the dependent variable in group analyses.

Results

Behavioral results

The effect of reputation on pre-and post-scan trust ratings

On the day of scanning, subjects were reminded of their peer nominations and the reasons for nominating peers as either kind or mean. Next, subjects completed a pre-scanning questionnaire during which they provided trustworthiness ratings for each partner on a scale of 1–7. For instance, a subject would be asked "How much do you trust Peter to give you back $2 when you give him $1?". A similar question was asked post-scan, e.g. "On the basis of what happened in the game today, how much do you trust Peter now?".

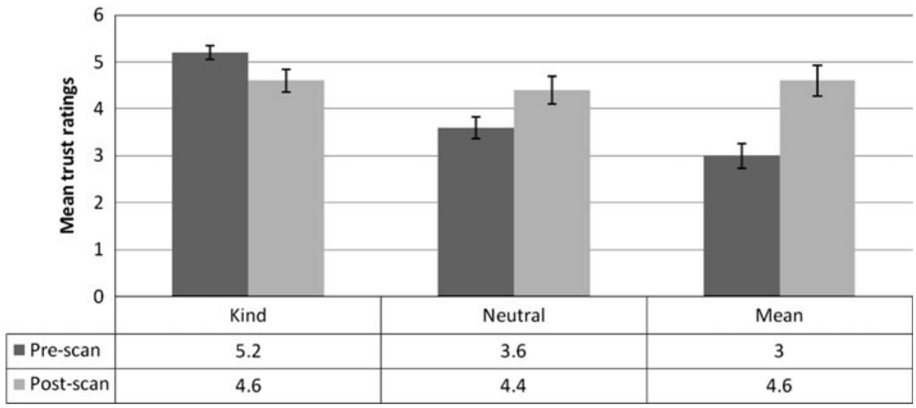
A two-way repeated measures ANOVA with Reputation (kind, neutral, mean) and Time (pre-scan, post-scan) as within-subjects factors and Externalizing problems as between-subjects factor showed main effects for Reputation ($F = 8.17$, $df = 2{,}18$, $p = 0.001$) and Time ($F = 33.90$, $df = 2{,}18$, $p < 0.001$) as well as an interaction effect between Reputation and Time ($F = 10.13$, $df = 2{,}18$, $p < 0.001$). No effects were found for Reputation*Externalizing problems ($F = 0.02$, $df = 18$, $p = 0.98$) or Time*Externalizing problems ($F = 0.42$, $df = 18$, $p = 0.52$). For pre-scan trust ratings within-subjects, trust ratings differed significantly for Kind versus Mean partners ($t = 7.93$, $df = 19$, $p < 0.001$), and kind versus neutral ($t = -5.007$, $df = 19$, $p < 0.001$). The difference for Neutral versus Mean partners was not significant although means were in the expected direction with trust ratings ($M = 3.00$ vs. $M = 3.60$; Mean vs. Neutral). These findings suggest that reputations mattered at the beginning of the game with subjects trusting in the expected directions. However, the fact that a main effect was found for Time, suggests that as the game progressed, subjects learned that partners were not necessarily behaving in a way consistent with their reputations (Fig. 2). Paired sample t tests revealed that post-scan ratings were significantly decreased for the Kind partner ($t = 2.35, df = 19, p = 0.03$) and significantly increased for the Mean partner ($t = -5.44$, $df = 19$, $p < 0.001$). Trust ratings also increased significantly for Neutral partners ($t = -2.37, df = 19, p = 0.03$). These findings show that by and large, the experimental manipulation affected the perception of trustworthiness.

The fact that the externalizing problems variable did not show any interaction effects with the independent variables suggest that boys with externalizing problems did not trust less than those without externalizing problems. Also, reputation had the same effect on trust ratings of boys with externalizing problems.

The effect of reputation on share decisions

Figure 3 shows the proportion of share decisions for each partner type broken down by high and low externalizing and early or late trials (Task Phase). Subjects shared more with kind and neutral partners early in the experiment and their rate of sharing decreased later in the experiment. The opposite pattern was observed for mean partners. A repeated measures ANOVA with Reputation and Task Phase (first 20 trials vs.

Fig. 2 The effect of reputation on pre-and post-scan trust ratings across both groups. Pre-scan question "How much do you trust X to give you back \$2 when you give him \$1?". Post-scan question: "On the basis of what happened in the game today, how much do you trust X now?"



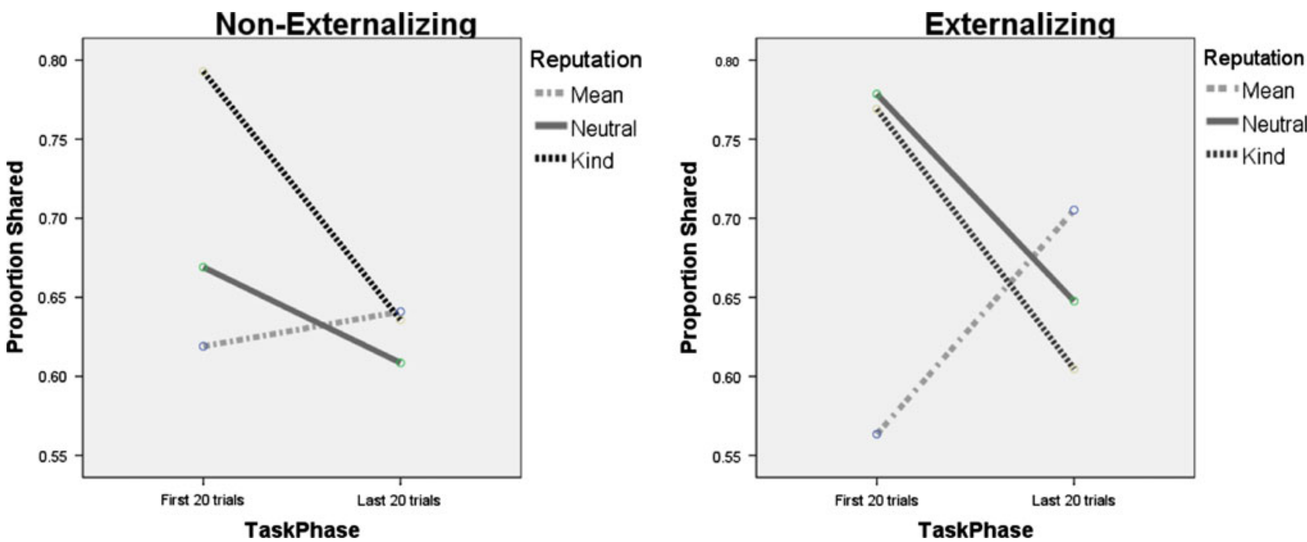


Fig. 3 The proportion of share decisions for each partner type, broken down by high and low externalizing and early (first 20) or late (last 20) trials (Task Phase)

last 20 trials) as within-subjects factors and Externalizing problems as a between-subjects factor revealed no main effect for Reputation type on the decision to share $F(2,72) = 1.97$, $p = 0.15$ and no interaction with Externalizing behavior $F(2,72) = 0.23$, $p = 0.64$. The interaction between Reputation and Task Phase however, was significant, $F(2,72) = 6.48, p = 0.003$, with both groups showing a pattern of sharing more with the mean partner toward the end of the experiment relative to early trials, and with the opposite pattern for kind and neutral partners. An ANOVA to test for any group differences in reaction times revealed no differences between boys with and without externalizing problems across all conditions [$F(1, 1172) = 0.70, p = 0.40$].

Brain results

Decision phase

Data (i.e., mean of peak time points) were initially submitted to a 3 (Kind vs. Mean vs. Neutral partner) × 2 (Keep vs. Share decision) ANOVA, with subjects as a random factor, to explore differences among condition means at each voxel. This analysis did not yield meaningful clusters of activation demonstrating differences among condition means. However, because subjects' brains responded robustly to performing the task, we explored the data further by identifying responsive regions of interest (ROIs) and averaging across voxels within regions. Our rationale was that averaging across voxels that were behaving similarly might provide power sufficient to detect differences that went undetected in individual voxels. To determine which voxels were active during the decision phase of the task, all decisions (i.e., both keep and share decisions for both groups of subjects) were collectively compared to fixation baseline (0) using a t test. Clusters larger than or equal to 100 μL of voxels significant at $t > 3.925$, $p < 0.0009$, for a false discovery rate $q < 0.01$, are listed in Table 1. Of interest in the present investigation are the bilateral insula regions (ROIs 8 and 10 in Table 1) and the anterior cingulate region (ROI 2 in Table 1), which were treated as regions of interest (ROIs).

Table 1 Clusters of activation for decision versus baseline fixation

Cluster number	Cluster size (μL)	Mean % change	Max % change	Talairach coordinates			Structures in vicinity of most active voxel in cluster
				x	y	z	
1	88,289	3.4278	8.0213	−30	−56	−18	Left fusiform gyrus
2	4,047	1.8369	3.3022	0	12	40	Left cingulate gyrus
3	840	1.6395	2.604	−10	−31	−2	Left parahippocampul gyrus
4	804	1.7229	5.264	−47	−38	57	Left inferior parietal lobule
5	727	1.7015	2.5922	0	−30	24	Left posterior cingulate
6	718	0.9265	−1.7534	−24	23	54	Left middle frontal gyrus
7	715	1.2549	1.737	37	1	30	Right inferior frontal gyrus
8	623	1.5058	1.931	45	19	0	Right anterior insula
9	502	2.8687	7.1048	34	−21	66	Right precentral gyrus
10	478	1.2415	−1.7519	44	−8	4	Right insula
11	330	1.9177	2.7075	−28	−56	42	Left superior parietal lobule
12	292	3.2719	5.8081	38	−33	64	Right postcentral gyrus
13	190	3.5299	5.9427	−10	0	70	Left superior frontal gyrus
14	154	1.2242	−1.6062	63	−30	27	Right inferior parietal lobule
15	149	1.3967	1.8009	−29	25	9	Left insula
16	130	2.0392	2.566	15	3	−7	Caudate nucleus
17	108	1.4961	1.9074	−39	17	4	Left insula

For each region, data for each time point in the 12-s window following the onset of the decision phase for each condition were averaged across all voxels within each ROI. The average of the peak time points (i.e., the typical peak at the 3rd time point and the two flanking time points) were used as the dependent variable in ROI analyses. Initial analyses performed with three levels of the Reputation factor did not yield significant results. However, following Delgado et al.'s [12] strategy, these data were submitted to a 2 (partner Reputation, with Kind and Mean combined vs. Neutral) × 2 (Share vs. Keep decision) × 2 (Group, Externalizing vs. Non-externalizing) ANOVA, yielding a main effect for Reputation in the ACC, $F(1,18) = 10.324$, $p = 0.005$, so that increased ACC activity was associated with decisions about neutral partners compared to making decisions about kind/mean partners. There were no interaction effects for Group*Reputation, $F(1,18) = 0.092$, $p = 0.766$ or Group*Decision, $F(1,18) = 0.004$, $p = 0.95$.

For the bilateral insula regions (ROIs 8 and 10), there was no main effect for Reputation, $F(1,18) = 0.009$, $p = 0.925$, but a marginally significant interaction effect for Group*Decision, such that increased insula activity was associated with share decisions for the externalizing group, $F(1,18) = 4.045$, $p = 0.06$. Figure 4 shows percent signal change averaged across all voxels within the bilateral insula for externalizing versus non-externalizing boys, depicting greater response to share decisions for externalizing boys across all reputation types. There were no interaction effects for Group*Reputation $F(1,18) = 0.056$, $p = 0.815$.

Outcome phase

To determine which voxels were active during the outcome phase of the task, all outcomes were collectively compared to baseline (0) using a *t* test. Clusters of voxels with $t > 2.979$ (False Discovery Rate $q < 0.05$) larger than or equal to 100 μL are listed in Table 2. Of interest in the present investigation were the caudate (ROI 10 in Table 2) and left and right insula clusters (ROIs 6 and 5, respectively, in Table 2), which were treated as ROIs.

For each region, data for each time point in the 12-s window following the onset of the decision phase for each condition were averaged across all voxels within each ROI. The average of the peak time points were used as the dependent variable in ROI analyses. As with the decision phase, initial analyses performed with three levels of the Reputation factor did not yield significant results so kind and mean conditions were averaged to create a Reputation factor with two levels. A 2 (Kind and Mean averaged vs. Neutral) × 2 (Positive vs. Negative outcome) × 2 (Externalizing vs. Non-externalizing group) ANOVA revealed a significant Reputation*Group interaction in the caudate $F(1,18) = 14.442, p = 0.002$. There was also a significant interaction effect for Reputation*Outcome $F(1,18) = 4.606$, $p = 0.048$. Thus, increased BOLD responses in the caudate were associated with outcomes from Neutral partner for Non-externalizing boys, while the caudate responded more strongly to outcomes from Kind/Mean partners for Externalizing boys. No main effect was found for Reputation in caudate $F(1,18) = 0.244$, $p = 0.628$.

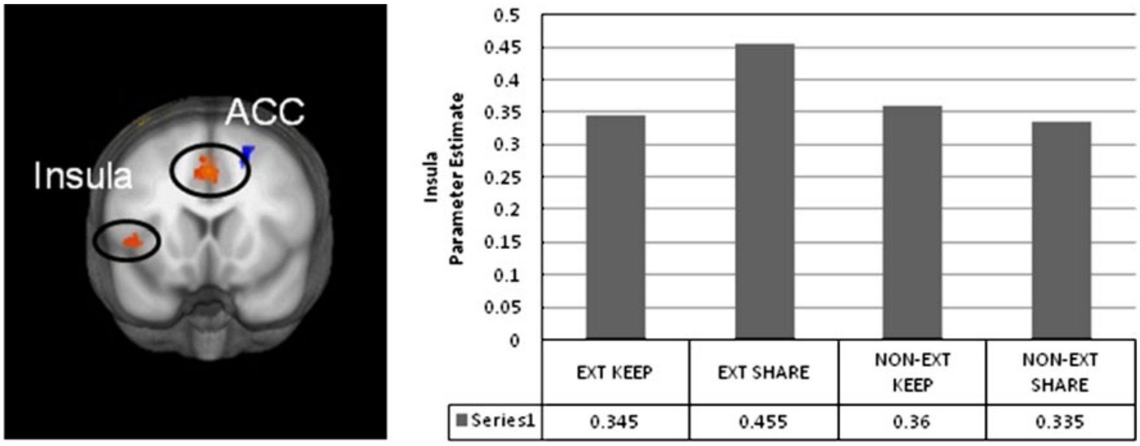


Fig. 4 Brain results for the decision phase of the game. Increased activity in the ACC was associated with decisions about neutral partners versus kind/mean partners for the full sample, with no interaction effects for Group*Reputation. Increased activity was also observed in the insula associated with share decisions for boys with externalizing behavior problems across all reputation types

Table 2 Clusters of activation for outcome versus baseline fixation

Cluster number	Cluster size (µL)	Mean % change	Max % change	Talairach coordinates			Structures in vicinity of most active voxel in cluster
				x	*y*	*z*	
1	17,4870	2.7493	7.6426	24	−98	8	Large cluster spanning multiple occipital regions in both hemispheres
2	8,473	1.1218	2.4483	0	6	46	Anterior cingulate/Medial frontal cluster, both hemispheres
3	7,351	1.2255	2.1184	37	−3	32	Right precentral gyrus
4	3,302	1.3239	3.0587	0	−26	25	Posterior cingulated (both hemispheres)
5	2,911	1.344	3.3482	54	13	−12	Right anterior insula
6	1,916	1.1616	3.1732	−51	19	−11	Left anterior insula
7	1,574	1.387	4.6787	−44	−9	58	Left precentral gyrus
8	981	1.1953	2.3507	62	−55	13	Right superior temporal gyrus
9	981	1.0154	1.4607	−41	4	33	Left inferior frontal gyrus
10	587	0.956	1.6563	7	4	12	Right caudate
11	241	0.8961	1.0852	−52	−25	0	Left superior temporal gyrus

In the right insula, there was a main effect for Reputation $F(1,18) = 4.606$, $p = 0.04$ with increased activity associated with Neutral reputations versus Kind/Mean. No interaction effects were found for Reputation*Group $F(1,18) = 1.437$, $p = 0.246$ or Outcome*Group $F(1,18) = 0.371$, $p = 0.55$. However, a Group*Reputation interaction was found for the left insula, such that boys with externalizing behavior problems showed greater activation for both types of outcome $F(1,18) = 4.606$, $p = 0.04$. There was no main effect for Reputation $F(1,18) = 1.182$, $p = 0.291$ or interaction effect for Outcome*Group $F(1,18) = 1.136$, $p = 0.301$. In the bilateral insula (Fig. 5), a marginal main effect for Reputation $F(1,18) = 3.182$, $p = 0.09$ was found with greater responses associated with Neutral reputations versus Kind/Mean reputations, except for boys with externalizing problems, as evidenced by an interaction effect for Reputation*Group $F(1,18) = 2.863$, $p = 0.108$ whose bilateral insula responses did not differ by reputation type. There was no Outcome*Group interaction effect in the bilateral insula $F(1,18) = 0.662$, $p = 0.426$.

Discussion

To our knowledge, this is the first study to employ an economic exchange game to investigate the neural correlates of reward-related decision-making in a social context in relation to externalizing behavior problems. Using an economic exchange game, it is possible to study the influence of social factors (e.g. reputations) on decision-making and associated corticostriatal circuitry as it relates

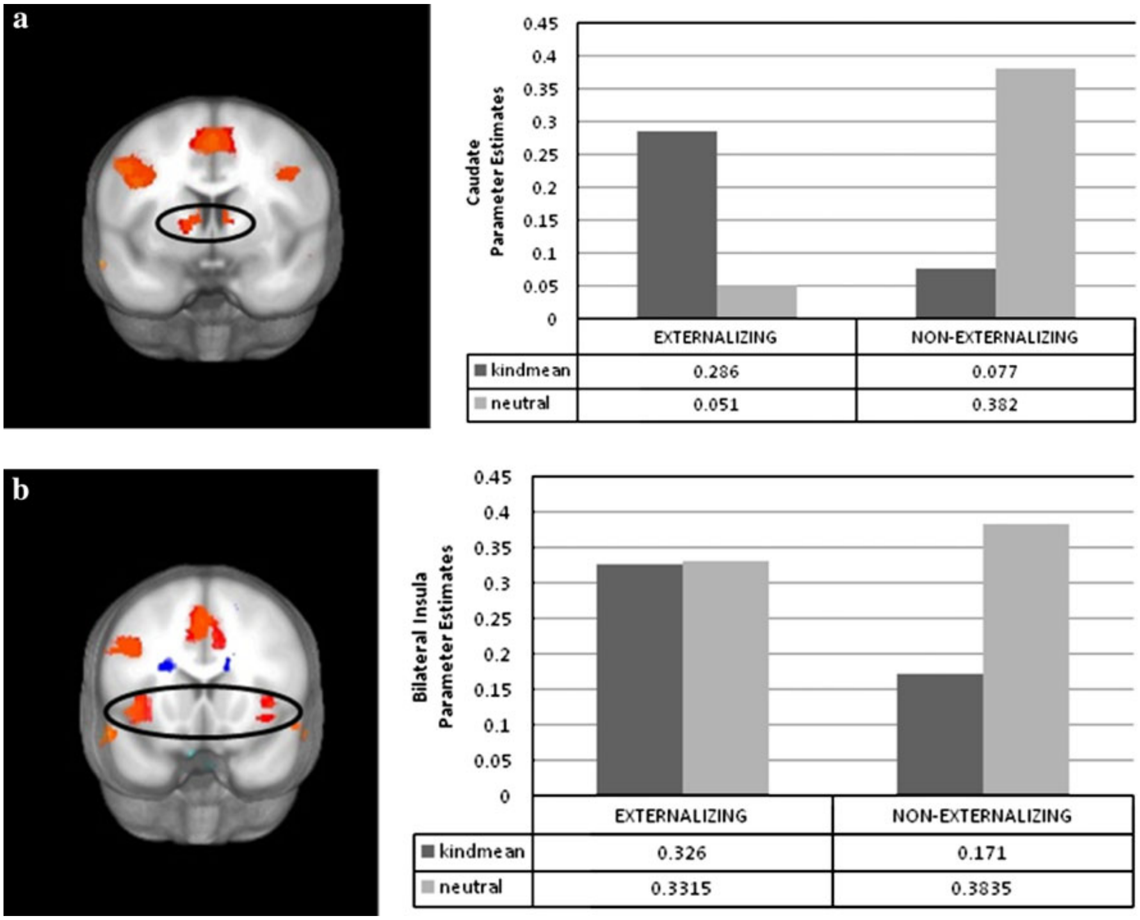


Fig. 5 Brain results for the outcome phase of the task. **a** increased BOLD responses in the caudate were associated with outcomes from neutral partner for non-externalizing boys, while the caudate responded more strongly to outcomes from kind/mean partners for externalizing boys. **b** increased responses in the bilateral insula associated with outcomes received from neutral partners for non-externalizing boys only. For externalizing boys, reputation did not affect responses in the bilateral insula

to psychopathology. In this context, our study is particularly novel in that for the first time in social neuroscience research actual real-life reputations (as opposed to hypothetical vignettes) were used. While useful in generating new findings, the results discussed below are interpreted with caution given the small sample size in the current study. We emphasize the need for replication of this preliminary work in larger groups of adolescents with and without externalizing problems.

Our findings generally support the notion that reputation modulates reward responses in the brain in similar ways to adults [12]. This conclusion was clearly supported during the decision phase of the game where a significant main effect for reputation was demonstrated for both externalizing and non-externalizing boys in the ACC. The ACC plays a key role in conflict monitoring [6]. In the context of neuroeconomic games, activity in the ACC can be seen as consistent with the existence of a tradeoff between self-interest and pro-social motives (integrating emotional feelings about costs vs. benefits, [16]. Generally then, we can conclude that more conflict between self-interest and pro-social motives was experienced in sharing with unknown partners (neutral) compared to the known (kind/mean) partners across all subjects regardless of externalizing status.

A second, but marginally significant finding, was observed for the decision phase of the game: increased bilateral insula activity was associated with share decisions for the externalizing group regardless of reputation. Delgado et al. [12], who used the same task, demonstrated insula activation in normally functioning adults when bias-incongruent decisions were made (share with the bad partner and keep with the good partner). Using this as a framework for the interpretation of our results, it is possible that insula activation in boys with externalizing disorder during share decisions (regardless of reputation) indicate incongruent decision-making (that is, reluctance to share). Given that the bilateral insula is often related to emotional resentment [16], we may argue that boys with externalizing problems experience greater emotional resentment when sharing.

Reward responses were also differentially modulated by reputation during the outcome phase of the game for boys with and without externalizing problems. Increased BOLD responses in the caudate were associated with interactions with neutral partner for non-externalizing boys, while the caudate responded more strongly to kind/mean partners for externalizing boys. The result in normally functioning boys is similar to Delgado et al's [12] study where during the outcome phase the caudate activated more strongly for repayment outcomes from the neutral partner, but not from the other partners, presumably because the neutral partner represents unpredictable outcomes and there is more to learn. In other words, outcomes from the neutral partner were more "surprising to the brain" in non-externalizing boys. In externalizing boys, the opposite pattern was found, suggesting "the devil you know" to be more unsettling for these boys.

Differential modulating effects of reputation were also significant for externalizing versus non-externalizing boys in the anterior insula. A trend-level finding suggested that while the bilateral insula was more activated in response to outcomes from neutral reputations in non-externalizing boys, bilateral insula responses in boys with externalizing problems did not distinguish between outcomes from different reputation types. In the context of neuroeconomic games, the bilateral insula is associated with negative or aversive evaluation of perceived or planned action from another [24]. Thus, studies have demonstrated increased anterior insula activation in response to aversive or uncomfortable occurrences in social interactions, such as unfairness [28], risky choices, frustration, or impending loss of social status [36]. Of particular relevance to the current study, the anterior insula has also been found to respond to the intentions and emotional state of others [2] and to violations in responses to social norms [20]. Against this background, the "insensitivity" of the bilateral insula in externalizing boys to the reputation of their partner in judging the fairness or unfairness of outcome, may indicate insensitivity in interpreting social cues from others associated with psychiatric disorder in general; or it may suggest that these boys expect unfair offers from the outset and, therefore, the reputation of their partners in judging offers as fair or unfair is irrelevant to them.

Taken together, and despite no differences at the behavioral level (probably due to reduced power), our results provide preliminary evidence for what may be brain-based anomalies in key aspects of social decision-making in boys with externalizing behavior problems. There are several limitations to this study, most notably the use of a community sample. While care was taken to identify boys who are deserving of referral for externalizing problems, the current study aimed to elucidate potential disease mechanisms for externalizing behavior problems and should, therefore, also be replicated in a sample of clinically referred youth. Second, while two groups of $n = 10$ afforded us enough power to detect differences, a larger sample size (combined with clinical characteristics) will result in larger effect sizes, thereby more strongly establishing the role of reward-related social decision-making in externalizing behavior problems. Third, due to the small sample size and the narrow age band in the current study, developmental issues were not considered. These are important for future studies because it is possible that additional striatal deficits may be observed in younger samples [17]. Fourth, the use of a stranger in the "neutral" condition of the task opens up the possibility that findings in the current study is related to a familiar versus unfamiliar contrast rather than moral character associated with each condition. An interesting follow-up study could make use of children with "neutral reputations" as stimuli instead of strangers. Finally, future researchers using this task may consider carrying out the experiment over two sessions as 32 min in one session can lead to fatigue especially in younger subjects. Future studies should also pay more careful attention to the possible confounding effects of using two scanners without explicitly modeling scanner effects.

Setting these limitations aside, and fully acknowledging the preliminary nature of the study, the study is important in that it lends biological plausibility to well-established social-cognitive and reward-based theories of externalizing behavior disorders, as well as treatment approaches that focus on reward-related decision-making in social contexts. This may be especially true in early pubertal adolescence given the protracted development of brain circuits underlying reward processing [15] and social cognition [5] during adolescence. Although the sample size in the current study does not allow for developmental analyses, behavioral studies (e.g. Sutter & Kocher, 2007) as well as neuroimaging studies with larger sample sizes have demonstrated asynchronous developmental patterns during adolescence in areas involved in social decision making. For example, using a trust task in 12–22 years old, Van den Bos et al. [32] demonstrated increased activation of brain areas involved in perspective taking and control of selfish actions with age, with mid-adolescence identified as an important transition period for intention detection and reciprocal behavior. Relatedly, Blakemore [5] talks about a "pubertal dip" that occurs early in adolescence in the performance of social-cognitive tasks which recovers by late adolescence. These normative maturational changes in the adolescent brain are thought to be responsible for increases in risky and poor decision-making in normally functioning adolescents [25], and, as demonstrated in the current study, may also form the basis of the risk for developing psychopathology [26].

Acknowledgments The study was funded by Departmental Seed Funding to Dr. Carla Sharp while she was faculty at Baylor College of Medicine. We wish to thank the youth and their families who participated in this study. We also wish to thank Dr. Michael Beauchamp for useful comments on the task design, and Dr. Manuel de la Cruz Gutierrez for comments on an earlier version of the paper.

Conflict of interest None.

References

1. Achenbach TM, Rescorla LA (2001) Manual for ASEBA school-age forms and profiles. University of Vermont, Research Center for Children, Youth and Families, Burlington
2. Adolphs R, Damasio H, Tranel D, Cooper G, Damasio AR (2000) A role for somatosensory cortices in the visual recognition of emotion as revealed by three-dimensional lesion mapping. J Neurosci 20(7):2683–2690
3. American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders*. 4th edn. Text Revision. APA (DSM-IV-TR), Washington, DC
4. Asher SR, Hymel S (1981) Children's social competence in peer relations: sociometric and behavioral assessment. In: Wine JD, Smye MD (eds) Social competence. Guilford, New York, pp 125–157
5. Blakemore SJ (2008) The social brain in adolescence. Nat Rev Neurosci 9:267–277
6. Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD (2001) Conflict monitoring and cognitive control. Psychol Rev 108(3):624–652
7. Boyle MH, Offord DR, Racine YA, Szatmari P, Fleming JE, Links PS (1992) Predicting substance use in late adolescence: results from the Ontario Child Health Study follow-up. Am J Psychiatry 149(6):761–767
8. Camerer CF, Weigelt K (1988) Experimental tests of the sequential equilibrium reputation model. Econometrica 56:1–36
9. Camerer C (2003) Behavioral game theory: experiments in strategic interaction. Princeton University Press, Princeton
10. Cillessen AHN, Bukowski WM (2000) Recent advances in the measurement of acceptance and rejection in the peer system. Jossey-Bass, San Fransisco
11. Cox RW (1996) Software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res 29:162–173
12. Delgado MR, Frank RH, Phelps EA (2005) Perceptions of moral character modulate the neural systems of reward during the trust game. Nat Neurosci 8(11):1611–1618
13. Dodge KA (1986) A social information processing model of social competence in children. In: Perlmutter M (ed) Minnesota symposium on child psychology. Erlbaum, Hillsdale, pp 77–125
14. Fairchild G, van Goozen SH, Stollery SJ, Aitken MR, Savage J, Moore SC et al (2009) Decision making and executive function in male adolescents with early-onset or adolescence-onset conduct disorder and control subjects. Biol Psychiatry 66(2):162–168
15. Fareri DS, Martin LN, Delgado MR (2008) Reward-related processing in the human brain: developmental considerations. Dev Psychopathol 20(4):1191–1211
16. Fehr E, Camerer C (2007) Social neuroeconomics: The neural circuitry of social preferences. Trends in Cognitive Sciences 11(10):419–427
17. Gatzke-Kopp LM, Beauchaine TP, Shannon KE, Chipman J, Fleming AP, Crowell SE, Liang O, Johnson C, Aylward E (2009) Neurological correlates of reward responding in adolescents with and without extenralizing behavior disorders. J Abnorm Psychol 118(1):203–213
18. Hill J (2002) Biological, psychological and social processes in the conduct disorders. J Child Psychol Psychiatry 43(1):133–164
19. Holmbeck GN, Thill AW, Bachanas P, Garber J, Miller KB, Abad M et al (2008) Evidence-based assessment in pediatric psychology: measures of psychosocial adjustment and psychopathology. J Pediatr Psychol 33(9):958–980 discussion 981-952
20. King-Casas B, Sharp C, Lomax L, Lohrenz T, Fonagy P, Montague R (2008) The rupture and repair of cooperation in borderline personality disorder. Science 321:806–810
21. Kishida KT, King-Casas B, Montague PR (2010) Neuroeconomic approaches to mental disorders. Neuron 67(4):543–554
22. Kochanska G (1991) Socialization and temperament in the development of guilt and conscience. Child Dev 62(6):1379–1392
23. Lee D (2008) Game theory and neural basis of social decision making. Nat Neurosci 11(4):404–409
24. Meyer-Lindenberg A (2008) Trust me on this. Science 321:778–780
25. Nelson EE, Leibenluft E, McClure EB, Pine DS (2005) The social re-orientation of adolescence: a neuroscience perspective on the process and its relation to psychopathology. Psychol Med 35:163–174
26. Paus T, Keshavan M, Giedd JN (2008) Why do many psychiatric disorders emerge during adolescence? Nat Rev Neurosci 9:947–957
27. Popma A, Raine A (2006) Will future forensic assessment be neurobiologic? Child Adolesc Psychiatr Clin N Am 15(2): 429–444, ix
28. Sanfey AG, Rilling JK, Aronson JA, Nystrom LE, Cohen JD (2003) The neural basis of economic decision-making in the Ultimatum Game. Science 300(5626):1755–1758
29. Sharp C, Ha C, Fonagy P (2011) Get them before they get you: Trust, trustworthiness and social cognition in boys with and without externalizing behavior problems. Dev Psychopathol 23:647–658
30. Singer T, Critchley HD, Preuschoff K (2009) A common role of insula in feelings, empathy and uncertainty. Trends Cogn Sci 13(8):334–340
31. Talairach J, Tournoux P (1988) Co-planar stereotaxic atlas of the human brain. Thieme, Stuttgart
32. Van den Bos W, Westenberg M, Van Dijk E, Crone EA (2010) Development of trust and reciprocity in adolescence. Cogn Dev 25(1):90–102
33. Verhulst FC, Van der Ende J (1992) Agreement between parents' reports and adolescents' self-reports of problem behavior. J Child Psychol Psychiatry 33(6):1011–1023
34. Vitaro F, Tremblay RE, Bykowski WM (2001) Friends, friendships and conduct disorders. In: Hill J, Maughan B (eds) Conduct disorders in childhood and adolescence. Cambridge University Press, Cambridge, pp 346–376
35. Werner NE, Crick NR (1999) Relational aggression and social-psychological adjustment in a college sample. J Abnorm Psychol 108(4):615–623
36. Zink CF, Tong Y, Chen Q, Bassett DS, Stein JL, Meyer-Lindenberg A (2008) Know your place: neural processing of social hierarchy in humans. Neuron 58(2):273–283