Neural correlates of retaliatory and prosocial reactions to social exclusion: Associations with chronic peer rejection

Geert-Jan Will, Eveline A. Crone, Pol A.C. van Lier, Berna Güroğlu

1. Introduction

From infancy to old age, humans have a fundamental need to form and maintain lasting positive relationships with others (Baumeister and Leary, 1995). Social exclusion frustrates this need and can lead to retaliation toward the sources of exclusion (Twenge et al., 2007), but may also lead to prosocial responses aimed at reconnection; both toward potential new sources of affiliation (Maner et al., 2007) as well as the peers responsible for exclusion (Will et al., 2015). The way people react to social exclusion has been hypothesized to be shaped through exposure to prolonged rejection by close others, such as parents (Feldman and Downey, 1994), or peers (London et al., 2007). Indeed, children and adolescents with a history of chronic peer rejection become increasingly more likely to defensively expect, readily perceive, and overreact to social rejection (London et al., 2007) and show a heightened neural reactivity to social exclusion (Will et al., 2016). Yet, how the neural responses underlying behavioral reactions to exclusion vary as a function of a history of chronic peer rejection remains to be investigated. Therefore, we examined neural processes involved in retaliatory (i.e. punishing) and prosocial (i.e. forgiving) reactions to social exclusion in adolescents with a history of chronic peer rejection and tested how they differed from adolescents with a history of stable high levels of peer acceptance.

Peer rejection reflects the collective valence of negative sentiments in a group toward a specific individual in that group, which is most commonly assessed through asking group members who they like most (positive) and who they like least (negative) (Bukowski et al., 2000; Coie et al., 1982; Newcomb and Bukowski, 1983). Children who receive many negative nominations and very few positive nominations are classified as rejected and develop widespread impairments in daily life, ranging from conduct problems (Sturaro et al., 2011) to delinquency (Kupersmidt et al., 1995) and dropping out of school (Hyymel et al., 1996). Transactional developmental models posit that such impairments arise out of a sustained pattern of reciprocal interactions between peers expressing dislike toward a rejected group member and the rejected member’s reactions to being disliked (Coie, 1990; Sandstrom and Coie, 1999). Social exclusion – defined as excluding someone from a group or
activity—i.e., the most common methods adolescents use to express dislike toward rejected peers (Coie, 1990). Transactional models predict that adolescents who react to exclusion with retaliatory vengeance might be more likely to elicit further rejection than those who show behavior aimed at reconnecting after exclusion. Preliminary support for this hypothesis comes from studies showing that adolescents with a rejected status report using more aggressive coping styles in response to social exclusion in a hypothetical scenario (Sandstrom, 2004). Elucidating the neurocognitive mechanisms underlying behavioral reactions to exclusion can further our understanding why some adolescents become trapped in a vicious cycle of chronic rejection and exclusion.

Neuroimaging studies have demonstrated that a history of peer rejection and accompanying intra- and interpersonal vulnerability factors are associated with enhanced neural responses to social exclusion in adolescence. Activity in the dorsal Anterior Cingulate Cortex (ACC)—a region involved in detecting and generating exclusion-related distress—during exclusion is enhanced in adolescents who were chronically rejected by peers during childhood (Will et al., 2016), who spent less time with friends (Masten et al., 2012), and those who anxiously and angrily expect exclusion (Masten et al., 2009). Extending our scope to the neural processes underlying behavioral reactions to social exclusion, and how they vary as a function of exposure to peer rejection, can increase our understanding of the processes underlying adolescents’ decisions to seek revenge or reconnection after exclusion.

Neural processes underlying retaliatory and prosocial reactions to social exclusion have proven to be reliably examined by giving people the opportunity to distribute money between themselves and those who previously either included or excluded them (i.e., the includers and excluders respectively). People selectively punish the excluders by decreasing their monetary outcomes while treating the includers fairly. This form of punishment has been associated with increased activity in the pre-supplementary motor area (pre-SMA)/ACC and anterior insula (Al) (Gunther Moor et al., 2012; Will et al., 2015). Refraining from punishment and acting prosocial toward the excluders (i.e., forgiveness) through sharing a sum of money equally with them has been associated with increased activation in neural circuitry supporting social cognition (i.e., the temporo-parietal junction [TPJ] and the dorsomedial prefrontal cortex [dMPFC] and executive control (i.e., lateral prefrontal cortex [IPFC]) (Gunther Moor et al., 2012; Will et al., 2015).

Well-established behavioral and cognitive signatures of a rejected (vs. an accepted) status inform our hypotheses about how neural processes underlying behavioral reactions to exclusion vary as a function of peer status history. In comparison with their accepted classmates, children with a rejected status are more likely to deal with interpersonal anger in aggressive ways that instigate further conflict (Fabes and Eisenberg, 1992; Rabiner et al., 1990). Furthermore, they exhibit deficits in social cognition (e.g., less sophisticated perspective-taking skills) and executive control (e.g., problems in impulse control and emotion regulation) skills (Dodge et al., 2003b; Eisenberg et al., 1997; Fink et al., 2014), which have been shown to be crucially involved in refraining from punishment and forgiving excluders. To be specific, lower levels of behavioral control are associated with higher levels of aggression toward excluders (Chester et al., 2013) and higher levels of perspective taking are associated with higher levels of forgiveness in the form of refraining from punishment and sharing a sum of money with excluders (Will et al., 2015). Based on these findings, we hypothesized that chronically rejected adolescents would show: 1) higher levels of punishment (and thus lower levels of forgiveness) toward excluders; 2) lower levels of perspective-taking and higher levels of executive control problems and 3) differential recruitment of neural circuitry supporting social cognition (e.g., dMPFC and TPJ) and executive control (e.g., IPFC) during forgiveness.

To test these hypotheses, we recruited participants whose acceptance and rejection among peers was assessed annually across six elementary school grades as part of a large-scale longitudinal study (Sturaro et al., 2011; van Lier and Koot, 2010). Using strict selection criteria, we invited a group of adolescents who were chronically rejected by peers and a group of adolescents who had a stable accepted status among peers to participate in the current study. While undergoing functional Magnetic Resonance Imaging (fMRI), they were first included and then excluded by two unknown excluders in a virtual ball-tossing game called Cyberball (Williams et al., 2000). Subsequently, they played an economic game, previously validated in adults, in which they could either punish or forgive the excluders (Will et al., 2015). Results on the neural correlates of exclusion in Cyberball are reported elsewhere (see Ref. Will et al., 2016).

We anticipated that punishment of excluders would be associated with increased activity in the pre-SMA/ACC and Al (Sanfey et al., 2003; Strobel et al., 2011). In contrast, forgiveness was expected to be associated with increased activity in the dmPFC, TPJ and IPFC (Brüne et al., 2013; Will et al., 2015). With respect to individual differences, we expected that adolescents with a history of chronic peer rejection, relative to adolescents with a history of stable peer acceptance, would show enhanced recruitment of brain regions implicated in social cognition (e.g., dmPFC, TPJ) and executive control (e.g., IPFC) during forgiveness of excluders, consistent with findings demonstrating that adults who showed less forgiveness behavior activated these networks to a greater extent when they did decide to forgive (Will et al., 2015). To further explore how individual differences in social cognition (i.e., perspective taking) and executive control (i.e., behavioral regulation) were associated with punishment and forgiveness behavior and neural activity during forgiveness, we tested for correlations with self-reported perspective-taking and parent-reported behavioral regulation skills.

2. Material and methods

2.1. Participants and recruitment procedure

Participants were recruited from a longitudinal study (N = 1189) investigating the impact of social experiences on behavioral, emotional, and academic outcomes between the ages of 6 and 12 (annually from first to sixth grade of elementary school). Each year, participants were asked to nominate the peers in their class whom they liked most and liked least (unlimited nominations). Using those nominations, an average social preference score (liked most—liked least nominations) across the six waves was calculated. Participants were identified as chronically rejected if they were in the lower 10th percentile of that 6-year average social preference and as stably highly accepted if they were in the upper 10th percentile. Using a 10% threshold insured that none of the chronically rejected adolescents were ever classified as sociometrically popular and none of the stably highly accepted adolescents were ever classified as rejected in any of the six waves. Correlations between social preference scores of adjacent years (rs 0.67–0.70, all ps < 0.001) were comparable to those reported in prior work (Salmivalli and Isaacs, 2005; Vitaro et al., 2007).

Based on these criteria, suitability for participation in the fMRI study and availability of recent contact information, 131 adolescents were asked to participate in the fMRI study. Twenty adolescents were excluded because they were either left-handed (n = 4), had an autism spectrum disorder (n = 1) or had braces (n = 15). Seven adolescents could not be reached. Of the remaining 104 candidate participants, 47 adolescents and their parents agreed to participate in the current fMRI study.
chose not to participate in the fMRI study (n = 57) did not differ from those who did participate in terms of average social preference, age, or gender (all ps > 0.19).

All participants indicated to be healthy and reported no contraindications for MRI (e.g. no head injuries, no history of neurological or psychiatric disorders), except for four participants with a history of rejection who were diagnosed with Attention-Deficit Hyperactivity Disorder (ADHD). Of those, three participants with ADHD were on a stable dose of methylphenidates, but were medication-free on the day of scanning and the preceding day. A radiologist reviewed all anatomical scans, and one participant was excluded from the analyses due to an anomaly. Three participants were excluded from neuroimaging analyses because their head movement exceeded 1 voxel (3 mm) in at least one direction.

The final sample consisted of 43 adolescents, including 25 adolescents with a history of stable peer acceptance (M [SD] social preference = 1.17 [0.18]; M [SD] age = 14.0 [0.78]; 13 males) and 18 adolescents with a history of chronic peer rejection (M [SD] social preference = −1.59 [0.52]; M [SD] age = 14.1 [0.57]; 13 males). Stably highly accepted and chronically rejected adolescents did not differ in age, pubertal status, gender, race or IQ (all ps > 0.12; see Table 1).

This study was conducted in accordance with the ethical standards of the American Psychological Association as expressed in the Declaration of Helsinki. All participants and their parents gave informed consent for the study. The recruitment procedure was blind, such that experimenters were not informed about individual participants' peer status history. Participants and their parents were not informed about their (child's) peer status history and were told that a wide range of participants from the longitudinal study was invited to participate in the fMRI study. Both the longitudinal study and the fMRI study were approved by the medical ethical committees of the respective universities.

2.2. Experimental procedure

Participants were first familiarized with the scanner environment with a mock scanner. After receiving instructions, participants carried out the following tasks in the scanner: (i) Cyberball inclusion with two anonymous peers, (ii) Cyberball exclusion by two novel anonymous peers, and (iii) A Dictator game in which participants distributed money between themselves and one of the players from the previous Cyberball games (i.e., one of the includers or the excluders; see Fig. 1A). Participants could see the stimuli on a screen located at the head of the scanner bore via a mirror mounted on the head coil. Head movement was restricted through the use of foam inserts inside the coil. After scanning, participants filled out a battery of questionnaires and were debriefed. Participants received a monetary compensation for participation and small gifts.

2.3. fMRI tasks

2.3.1. Cyberball

Participants played two rounds of a virtual ball-tossing game called Cyberball (Williams et al., 2000). Participants were led to believe that the other players in the game were other participants in the experiment who were connected with them via the Internet. In reality, the other players' behavior was preprogrammed. First, participants played Cyberball with two unfamiliar peers (i.e. the includers; two cartoon figures accompanied by a girl's and a boy's name) who included them in a game where each player received the ball an equal amount of times (10/30 throws). Subsequently, they played another round of Cyberball with two novel unfamiliar players (i.e. the excluders; with a new boy's and a new girl's name) who threw the ball once to the participants at the start of the game, but further excluded the participants by not throwing the ball to them for the remainder of the game (28/30 throws).

2.3.2. Dictator game

Following Cyberball, participants played a modified Dictator Game in which they could distribute coins between themselves and a recipient (Will et al., 2015). The participants were instructed that the recipient was one of the players from either the first or the second Cyberball game (i.e., Team 1 and Team 2, respectively). It was explained that they would encounter all four players from both Cyberball games over successive rounds of the Dictator Game and that the recipient on each individual trial would be one person (e.g. one player from team 2; see Fig. 1A). To ensure that the participants were aware of this notion, they were asked during which of the two ball game they encountered the players on each team. Participants were told that the coins represented real money and that their decisions determined how much money they and the recipients would receive at the end of the experiment. In reality, each participant received the same fixed amount of money as compensation.

Participants were given a dichotomous choice to either share an amount of money equally or unequally (see Fig. 1B). The equal distribution (which always took the form of 5 coins for self/5 for the recipient) was pitted against an unequal alternative, which varied across three conditions: (i) beneficial inequality (8 coins for the participant/2 for the recipient), (ii) non-costly inequality (5 coins for the participant/2 for the recipient), (iii) prosocial inequality (4 coins for the participant/6 for the recipient). In line with previous work (Will et al., 2015), the prosocial inequality condition was added as a filler condition. By including a filler condition in which the unequal option increased the recipient's outcome (as opposed to decreasing their outcomes in the punishment conditions), we aimed to prevent automaticity in responding. Behavioral analyses confirmed that behavior in the prosocial inequality condition was indeed unrelated to behavior in the other conditions (see Section 3.1). Consequently, the prosocial inequality condition was not included in our neuroimaging analyses of the neural correlates of punishment and forgiveness (see Section 2.7).

The Dictator Game consisted of 120 trials (20 trials per condition; 3 inequality conditions x 2 recipients) and was administered in two runs of 207 vol each, lasting about 15 min in total. After a screen with fixation cross which had a jittered duration (M = 1540 ms; SD = 1083 ms; min = 550 ms; max = 4950 ms; optimized with OptSeq2; Dale, 1999), participants were presented with a decision screen that showed: (i) the two distributions they could choose from and (ii) whether the recipient was a player from Team 1 (i.e. an includer) or Team 2 (i.e. an excluder) (see Fig. 1A). Participants could choose one of two distributions of money by pressing a button with the index or middle finger of their right hand. After a button was pressed, a red rectangle appeared around the chosen distribution until 6 s after trial onset. Failing to respond within 5 s resulted in the presentation of a screen with “Too late!” with the duration of 1 s. Trials without a response consisted of less than 1% of all trials and were excluded from further analyses.

2.4. Questionnaires

2.4.1. Executive control

To assess executive control the participants' parents filled out the 'Behavioral regulation index' scale of the 'Behavior Rating Inventory of Executive Function' questionnaire (BRIEF; Gioia et al., 2000; Huizenga and Smidt, 2011). The Behavioral regulation index represents the ability to shift cognitive sets and to modulate behavior and emotions. The scale comprises of three subscales: inhibition (the capacity to suppress impulses, e.g. “Blurts things out”), shifting (the capacity to flexibly adjust behavior to changing demands of a given situation, e.g. “Becomes upset by new situations”), and
emotional control (the ability to regulate emotional responses, e.g. “Has explosive, angry outbursts”). All items were rated on a scale consisting of 1 (never), 2 (sometimes) to 3 (often) and summed. Higher scores on the Behavioral regulation index reflect increased difficulty with behavioral regulation.

2.4.2. Perspective taking

To assess the capacity to adopt another person’s point of view participants filled out the perspective-taking subscale of the Interpersonal Reactivity Index (IRI; Davis, 1983) (e.g., “Before criticizing somebody, I try to imagine how I would feel if I were in their place.”). All items were rated on a scale from 1 (not at all) to 5 (very much) and averaged to a mean score of trait perspective-taking.

2.5. fMRI data acquisition

Scans were acquired using a 3T Philips Achieva MRI system at the Leiden University Medical Center. After obtaining a localizer scan, T2*-weighted Echo-Planar Images (EPI) were acquired (repetition time (TR) = 2.2 s, echo time (TE) = 30 ms, slice matrix = 80 × 80 matrix, slice thickness = 2.75 mm, slice gap = 0.28 mm gap, field of view (FOV) = 220 × 220 × 114.68 mm) during two functional runs of 207 vol each. The first two volumes in each functional run were discarded to allow for equilibration of T1 saturation effects. High-resolution T1-weighted and T2-weighted anatomical scans (TR = 9.760 ms; TE = 4.59 ms, 140 slices, 0.875 × 0.875 × 1.2 mm voxels, field of view = 224 × 168 × 177 mm) were acquired for anatomical reference.

2.6. fMRI data analysis

MRI data were preprocessed and analyzed using SPM8 statistical parametric mapping imaging analysis software (Wellcome Trust Centre for Neuroimaging, University College London). Functional images were slice-time corrected, realigned, co-registered to individual structural T1 scans, normalized to a T1 template, and spatially smoothed using an 8-mm FWHM Gaussian filter. The normalization algorithm, resampled the volumes to 3 mm cubic voxels using a 12-parameter affine transformation and a nonlinear transformation involving cosine basic functions. All results are reported in MNI305 stereotactic space.

A first-level general linear model was defined for each participant’s functional run that included a boxcar regressor for each epoch of interest (e.g., decision phase) and convolved with a canonical hemodynamic response function (HRF). The duration of epochs in which participants submitted a response was modeled using the participant’s reaction time. Regressors were defined separately for choosing equal and unequal distributions in each of the three inequality conditions and analyzed separately for includers and excluders. This model consisted of 12 decision-related regressors (i.e., inequality condition [3] × recipient [2] × choice [2]), a regressor indicating missed trials, and a covariate for each run to control for run effects (3), resulting in a General Linear Model (GLM) with a total of 16 predictors and contained a basic set of cosine functions that high-pass-filtered the data. The least-squares parameter estimates of the height of the best-fitting canonical HRF for each condition separately were used in pair-wise contrasts at the subject level. There was no minimum amount of trials for the decision-related regressors. The resulting contrast images were submitted to group analyses where participants were treated as a random effect. Subsequently, we performed one-tailed t-tests of which results were considered significant at an uncorrected threshold of p < 0.001 with a minimum cluster size of 10 contiguous voxels to balance between Type 1 and Type 2 errors (Lieberman and Cunningham, 2009). We also report which clusters are significant using Family wise Error (FWE) cluster-correction at p < 0.05 with a cluster forming threshold of p < 0.001. We used the Marsbar toolbox (Brett et al., 2002; http://marsbar.sourceforge.net/) to extract and plot activity in functional regions of interest derived from the whole-brain contrasts.

2.7. Operationalization of constructs and plan of analyses

Punishment and its counterpart forgiveness are both behavioral reactions to an offense or a violation of a social norm (McCullough et al., 1997). Consequently, punishment and forgiveness can only take place in interactions with the excluders. To get a clean measure of punishment/forgiveness baseline preferences for equality have to be accounted for. Therefore, punishment was operationalized as unequal treatment of excluders (i.e. decreasing the recipient’s monetary outcomes) while controlling for baseline treatment of includers (difference score: excluders unequal—includes unequal treatment) in both our behavioral and neuroimaging analyses (see Fig. 1C). Consistent with prior work (Brüne et al., 2013; McCullough et al., 1997; van der Wal et al., 2014; Will et al., 2015), forgiveness was operationalized as refraining from punishment and acting prosocial toward the offenders (i.e. excluders), i.e. equal treatment of excluders while controlling for baseline fairness preferences. Importantly, only treatment in the beneficial inequality and non-costly inequality conditions was used for these analyses, because the prosocial inequality condition does not offer participants the opportunity to punish (because excluders’ outcomes could not be reduced) neither to forgive (since there is no possibility to refrain from punishment) (Will et al., 2015). To substantiate these conceptual considerations by our data, we examined allocation behavior toward includers and excluders in all conditions and subsequently tested the hypothesized correlations between treatment of includers and excluders in each of the conditions separately. To insure that group differences in BOLD responses could not be explained by differences in reaction times, we also tested for differences in reaction times between groups and conditions.
Fig. 1. (A) Experimental procedure: Participants carried out the following tasks in the scanner: (i) Cyberball inclusion with two anonymous peers (includers: green), (ii) Cyberball exclusion by two novel anonymous peers (excluders: red), and (iii) Dictator game in which participants distributed money between themselves (depicted with red coins) and the players from the Cyberball games (i.e., includers and excluders; depicted with blue coins). (B) In every trial, participants were given a dichotomous choice between either an equal distribution of money and an unequal distribution, of which the latter varied depending on the inequality condition. (C) Punishment was operationalized as the difference between unequal treatment of excluders – unequal treatment of includers in the punishment (i.e. beneficial inequality and non-costly inequality) conditions.

Given that our conceptual operationalization of punishment and forgiveness was supported by the data (see analysis of behavioral results: 3.1), we collapsed across beneficial and non-costly punishment choices for our neuroimaging analyses to increase power. To examine the neural correlates of punishment, we contrasted unequal treatment of excluders in the beneficial inequality and non-costly inequality condition with unequal treatment in those conditions for includers (Unequal treatment excluders > Unequal treatment includers; median amount of trials unequal treatment excluders = 30; range = 3–40; median amount of trials unequal treatment includers = 13; range = 1–40). To examine the neural correlates of forgiveness, we contrasted equal treatment of excluders in the beneficial inequality and non-costly inequality condition with equal treatment in those conditions for includers (Equal treatment excluders > Equal treatment includers; median amount of trials equal treatment excluders = 11; range = 1–39; median amount of trials equal treatment includers = 33; range = 1–40). For completeness, we contrasted unequal treatment of excluders in the beneficial inequality and non-costly inequality condition with equal treatment in those conditions for includers (Unequal treatment excluders > Equal treatment includers). For group comparisons, contrast images were entered into separate second-level analyses.
for each contrast of interest, where peer status history (chronically rejected vs. stably highly accepted) was the between-subjects variable in independent samples t-tests. To explore individual differences, we ran two whole-brain regression analyses on the forgiveness contrast (Equal treatment excluders > Equal treatment includers) with self-reported trait perspective-taking or parent-reported behavioral regulation problems as predictors. Both for behavioral and neuroimaging analyses, we ran analyses in the following order: i) main effects across the sample, ii) peer status history group differences, and iii) individual differences associated with perspective-taking and executive control.

3. Results

3.1. Behavioral results: punishment and forgiveness

First, we examined allocation behavior toward includers and excluders in all conditions. We performed a repeated measures ANOVA with recipient (2 levels: includers vs. excluders) and inequality condition (3 levels: beneficial inequality, non-costly inequality and prosocial inequality) as within-subjects factors and peer status history (2 levels: chronically rejected vs. stably highly accepted) as a between-subjects factor for the percentage of unequal distributions chosen in the Dictator Game. This analysis yielded main effects of recipient, $F(1, 41) = 30.37, p < 0.001, \eta^2_p = 0.43$, inequality condition, $F(2, 82) = 35.90, p < 0.001, \eta^2_p = 0.47$, and a recipient x inequality condition interaction, $F(2, 82) = 34.62, p < 0.001, \eta^2_p = 0.46$. Participants chose the unequal distribution more often for the excluders than for the includers in the beneficial and non-costly inequality conditions (both $p < 0.001$; see Fig. 2). In contrast, they chose the unequal distribution more often for the includers than for excluders in the prosocial inequality condition ($p = 0.031$). Neither interaction effects with peer status history or a main effect of peer status history were found. Thus, both stably highly accepted and chronically rejected adolescents punished the excluders by choosing unequal distributions of money for the excluders and they did this to a similar extent.

Beneficial punishment (beneficial inequality excluders - beneficial inequality includers) and non-costly punishment (non-costly inequality excluders - non-costly inequality includers) were highly correlated ($r = 0.81, p < 0.001$) and did not differ from another ($t = -1.18, p = 0.245$). Behavior in the prosocial inequality condition was unrelated to behavior in the other inequality conditions, both for includers (both $r < 0.12, both p > 0.447$) and for excluders (both $r < 0.11, both p > 0.513$). Together these analyses confirm: (i) the validity of our operationalization of punishment/forgiveness, (ii) the decision to collapse choices in the beneficial inequality and non-costly inequality conditions for our neuroimaging analyses of punishment and forgiveness, and (iii) the decision to exclude the prosocial inequality condition from our neuroimaging analyses of punishment and forgiveness.

Chronically rejected adolescents ($M = 44.8; SEM = 2.8$) had more parent-reported behavioral regulation difficulties than stably highly accepted adolescents ($M = 34.8; SEM = 1.1$), $t = -3.32, p = 0.003$, also after controlling for gender and ADHD diagnosis. Self-reported perspective-taking skills did not differ between groups ($p = 0.108$). Punishment behavior (beneficial and non-costly unequal treatment of excluders – includers) was not correlated with behavioral regulation difficulties ($p = 0.393$) or perspective taking ($p = 0.190$).

To investigate reaction time differences, we performed a repeated-measures ANOVA with recipient (2 levels: includers vs. excluders) and inequality condition (3 levels: beneficial inequality, non-costly inequality and prosocial inequality) as within-subjects factors and peer status history (2 levels: chronically rejected vs. stably highly accepted) as a between-subjects factor and reaction times as the dependent variable. This analysis yielded a main effect of inequality condition, $F(2, 82) = 12.47, p < 0.001, \eta^2_p = 0.23$, but no further significant main or interaction effects, all $F < 2.46, ps > 0.10, \eta^2_p < 0.058$. Follow-up pairwise comparisons showed that reaction times were faster in the Beneficial Inequality condition ($M = 1591$ ms; $SEM = 69$ ms) than in Non-Costly Inequality ($M = 1744$ ms; $SEM = 76$) and Prosocial Inequality ($M = 1798$ ms; $SEM = 74$) conditions (both $ps < 0.018$), and the latter two conditions did not differ from each other, $p = 0.206$.

3.2. Neuroimaging results

3.2.1. fMRI results: punishment and forgiveness across the sample

Before we examined differences in brain responses between chronically rejected and stably highly accepted adolescents, we first investigated the neural correlates of punishment and forgiveness across the sample. The punishment contrast (Unequal treatment excluders > Unequal treatment includers) resulted in activation in bilateral ventral striatum (peaks at −12, 20, 7 and 9, 20, 4), right dIPFC (36, 29, 37) and bilateral parietal cortex (peaks at 30, −58, 61 and −45, −45, 52) (Fig. 3A). The forgiveness contrast (Equal treatment excluders > Equal treatment includers) contrast resulted in activation in the dmPFC (peak at 6, 47, 22; Fig. 3B). Alternative punishment (Unequal treatment excluders > Equal treatment excluders) and forgiveness (Equal treatment excluders > Unequal treatment excluders) contrasts did not result in significant clusters of activation at our chosen threshold. All significant clusters are reported in Table 2.

3.2.2. Individual differences associated with chronic peer group rejection, perspective-taking skills and behavioral regulation problems

To examine how neural processes involved in punishment and forgiveness of exclusion varied as a function of chronic peer rejection, we compared the two groups using two-sample t-tests on both contrasts outlined above. A two-sample t-tests on the punishment contrast (Unequal treatment excluders > Unequal treatment includers) showed that chronically rejected and stably highly accepted adolescents showed no differential brain activity during punishment of excluders. During forgiveness of excluders, chronically rejected adolescents showed enhanced activity in lIPFC (peak at 36, 44, 4) and dorsal striatum (peak at 9, 11, 10) when directly compared to stably rejected adolescents in a two-sample t-tests on the forgiveness contrast (Equal treatment excluders > Equal treatment includers). As can be seen in Fig. 4, the group differences are driven by heightened lIPFC and dorsal striatum responses during
equal treatment of excluders relative to includers in chronically rejected adolescents, and a reverse pattern of activity in LPFC and no differentiation in dorsal striatum in stably highly accepted adolescents.

To explore how individual differences in perspective taking and executive control were associated with neural activity during forgiveness, we ran two whole-brain regression analyses on the forgiveness contrast (Equal treatment excluders > Equal treatment includers) with self-reported trait perspective-taking skills or parent-reported behavioral regulation problems as predictors. Participants with higher levels of perspective taking showed higher levels of activity in a region of the dmPFC (peak at −3, 50, 37) during forgiveness, which overlapped with the cluster of activation obtained in the main forgiveness contrast (Fig. 5A). Participants with more behavioral regulation problems showed more activity in the right dorsal AI (peak at 39, 17, 10) and the pre-SMA/ACC.

![Brain regions revealed by whole-brain contrasts of treatment of the excluders vs. includers in the Dictator Game across the sample (all thresholded p < 0.001 uncorrected, >10 voxels).](image)

**Table 2**

| Anatomical region                                      | L/R | Voxels | z     | MNI coordinates |
|--------------------------------------------------------|-----|--------|-------|-----------------|
| Punishment (Inequality excluders > inequality includers) |     |        |       |                 |
| Caudate                                                | L   | 50     | 4.09  | −12             |
| Dorsolateral Prefrontal Cortex                         | R   | 71     | 3.94  | 36              |
| Caudate                                                | R   | 57     | 3.83  | 12              |
|                                                     |     |        | 3.67  | 20              |
|                                                     |     |        | 3.40  | 20              |
| Superior parietal lobule                               | L   | 38     | 3.80  | −39             |
|                                                     |     |        | −39   | −49             |
| Inferior parietal lobule                               | R   | 40     | 3.30  | −58             |
|                                                     |     |        | 3.07  | −58             |
| Forgiveness (Equality excluders > equality includers)  |     |        |       |                 |
| Dorsomedial Prefrontal Cortex                          | R   | 11     | 3.54  | 6               |

Note: L/R=Left/Right; k=cluster size in 3 x 3 x 3 mm voxels; z=z-score; MNI coordinates=x y z voxel coordinates in MNI space of the peak voxel.

* significant using FWE cluster-correction, p<0.05 (cluster forming threshold p<0.001).
Table 3

Brain regions revealed by whole-brain individual differences analyses of sharing equally with the excluders relative to the includers in the Dictator Game (all thresholded $p < 0.001$ uncorrected, $>10$ voxels).

| Anatomical region                          | L/R | Voxels | z     | MNI coordinates |
|--------------------------------------------|-----|--------|-------|-----------------|
|                                            |     |        |       |                 |
| Chronically rejected adolescents $>$ Stably highly accepted adolescents $[$Equality excluders $>$ Equality includers$]$ |     |        |       |                 |
| Lateral Prefrontal Cortex                  | R   | 35     | 3.95  | 36 44 4        |
| Caudate                                    | R   | 15     | 3.79  | 9 11 10        |
| Positive correlation with perspective-taking $[$Equality excluders $>$ Equality includers$]$ |     |        |       |                 |
| Dorosmedial Prefrontal Cortex              | L   | 191    | 4.13  | −3 50 37"      |
| Positive correlation with behavioral regulation problems $[$Equality excluders $>$ Equality includers$]$ |     |        |       |                 |
| Anterior insula                            | L   | 81     | 4.07  | 42 5 13        |
| Pre-supplementary motor area/dACC          | R   | 14     | 3.46  | 3 11 55        |

Note: L/R = Left/Right; k = cluster size in $3 \times 3 \times 3$ mm voxels; $Z$ = $z$-score; MNI coordinates = xyz voxel coordinates in MNI space of the peak voxel. * significant using FWE cluster-correction, $p < 0.05$ (cluster forming threshold $p < 0.001$). dACC = dorsal Anterior Cingulate cortex.

4. Discussion

This study examined the neural processes involved in punishment and forgiveness of excluders and how these processes vary as a function of chronic peer rejection. The first main finding was that both chronically rejected and stably highly accepted adolescents punished the excluders by selectively decreasing their monetary outcomes; both when punishment resulted in gains and when punishment had no monetary benefits. Social exclusion thus elicits a tendency to retaliate against the sources of exclusion even when this does not result in material gain, which is not modulated by a history of chronic peer rejection.

The second main finding was that punishment was associated with heightened activity in the striatum, the dIPFC and parietal cortex in both groups of adolescents. The striatum has previously been shown to be involved in processing both primary (e.g. food) and social rewards, including punishment of unfair interaction partners (Singer et al., 2006). Furthermore, striatum activation during punishment has been shown to correlate with the self-reported desire for revenge and has therefore been suggested to code for the rewarding aspect of retaliation. Activation in the dIPFC has been linked to punishing criminal intent (Buckholtz et al., 2008) and economic unfairness (Güroğlu et al., 2011). The dIPFC – in concert with the parietal cortex – has been proposed to integrate different value signals and execute a punishment response among competing response options (Buckholtz and Marois, 2012).

Contrary to our hypotheses, chronically rejected adolescents did not show lower levels of forgiveness than stably highly accepted adolescents. Nonetheless, despite both groups displaying similar levels of forgiveness behavior, meaningful differences emerged on a neural level. To be specific, the third main finding was that chronically rejected adolescents, compared to stably highly accepted adolescents, displayed higher levels of activity in the dorsal striatum and the IPFC when they forgave the excluders. The dorsal striatum is strongly connected with a dorsal frontal-parietal network that plays a vital role executive control (Haber and Knutson, 2010; van den Bos et al., 2014). Together with the chronically rejected adolescents’ higher levels of behavioral regulation difficulties (indexing problems in inhibition, shifting and emotional control), heightened activity in this dorsal frontal-striatal network suggests that chronically rejected adolescents have to exert greater levels of control in order to act prosocial toward those who previously excluded them. This dovetails with findings showing that when children are asked to react to a peer hassle under time pressure, those with a rejected status generate more aggressive reactions than children with an accepted status (Rabiner et al.,
consistent with their higher levels of aggressive behavior in daily life (Dodge et al., 2003a; Ladd, 2006). However, when they were given enough time to think about their reactions, the reactions of most children with a rejected status became indistinguishable from their peers with an accepted status. Taken together, a greater demand on neural circuitry implicated in cognitive control during prosocial reactions to exclusion could be a correlate of these reactions being more effortful for adolescents who were chronically rejected.

The fourth main finding was that analyses of individual differences yielded insights into processes underlying refraining from retaliation. Two mechanisms have been shown to be involved in succeeding or failing to refrain from retaliation after being excluded: i) the ability to take other people’s perspectives (Will et al., 2015) and ii) the ability to cognitively control behavior (Chester et al., 2013). Our findings show that these two mechanisms are associated with activity in functionally separable neural networks. That is, during forgiveness, perspective-taking skills scaled with activity in the dmPFC, which is a hub in the ‘theory of mind’ network (Koster-Hale and Saxe, 2013) and behavioral regulation problems correlated positively with activity in the pre-SMA/ACC and dorsal AI. A meta-analysis has shown that the dorsal AI (overlapping with the cluster in our study) is strongly connected to the pre-SMA/ACC and dIPFC and is implicated in executive control functions, including inhibition and switching (Chang et al., 2013). Taken together, these findings show that individual differences in cognitive functions can be reliably linked to activity in functionally separable neural networks supporting social cognition (e.g. dmPFC) or executive control (e.g. ACC and dorsal AI) and they underscore the importance of examining individual differences in neural processes underlying prosocial reactions to social exclusion.

Several limitations of this study deserve to be mentioned. First, the data do not speak to the causal question whether observed differences in forgiveness-related neural activity were caused by chronic peer rejection, or whether they reflect a propensity that was already present before the emergence of a rejected peer status. High levels of behavioral problems are known to predict peer rejection and peer rejection in turn is known to predict increases in behavioral problems above and beyond prior levels of problem behavior (Sturaro et al., 2011). The observed differences in neural processing between the two groups may thus reflect: 1) a predisposition that was already present before the emergence of a rejected peer status, 2) a result of chronic peer rejection or 3) a combination of both. Longitudinal studies with multiple MRI scans and classroom assessments of peer rejection at different time-points across development are needed to better understand the nature of the relationship between brain development and chronic peer rejection. Second, differences between the two groups could both be the result of greater exposure to negative peer interactions in the rejected group or greater exposure to positive peer interactions in the accepted group. Future studies should compare rejected adolescents to adolescents with a so-called ‘average’ social status to disentangle the influence of positive and negative experiences in the peer group on neural processing involved in social behavior. Third, our sample of chronically rejected adolescents was relatively small. Rejected adolescents form a heterogeneous group consisting of aggressive and non-aggressive subtypes (Ladd, 2006). Future research with larger samples could test whether the absence of behavioral differences between the two groups in our study could be due to heterogeneity in the rejected group. Fourth, we did not collect self-reports of trait forgiveness to validate our behavioral measure of forgiveness. Nonetheless, other studies have validated this behavioral measure of forgiveness by showing that sharing valuable resources with an offender in the Dictator Game correlates with self-reports of trait forgiveness (van der Wal et al., 2014). Moreover, our index of forgiveness comprises (at least) two components that are central to forgiveness: i) refraining from punishment and ii) a prosocial concern for the offender’s well being. Nonetheless, cautiousness is warranted in drawing firm conclusions about the extent to which our behavioral measure reflects the multi-faceted construct that is forgiveness.

5. Conclusions

Our results show that punishment and forgiveness of excluders rely on distinct neural networks implicated in emotional and cognitive processes involved in social decision-making. Although chronically rejected adolescents were equally likely to forgive excluders as stably highly accepted adolescents, they displayed enhanced recruitment of the dorsal striatum and IPFC when forgiving excluders, suggesting that they might have to exert greater levels of executive control in order act prosocial toward peers who excluded them. Consequently, our results have implications for understanding the processes through which peer rejection’s adverse effects are transmitted across development.

A greater demand on executive control functions might entail greater difficulties controlling retaliatory responses to negative treatment in the heat of the moment of everyday interactions with peers. In turn, retaliatory responses are likely to provoke excluders, which could elicit new episodes of exclusion and thereby further consolidate a rejected status in the peer group (Sandstrom, 2004). A priority for future research is therefore to further examine longitudinal associations between behavioral reactions to social exclusion, peer status, and different subcomponents of cognitive control (e.g. inhibition, shifting, emotion regulation tasks) to get a better understanding of which executive functions might underlie prosocial reactions to exclusion. Such endeavors can inform research that could test whether training of executive functions might facilitate adaptive social responses to exclusion, which ultimately could lead to greater acceptance among peers.

Conflict of interest

None.

Author contributions

All authors contributed to the study design. Data collection was performed by G.-J.W. with the help of research assistants. G.-J.W. analyzed the data under the supervision of E.A.C. and B.G. All authors were involved in writing of the manuscript and they all approved the final version of the manuscript for submission.

Funding

This work was supported by the Netherlands Organization for Scientific Research (NWO; grant number 056-34-010 to E.A.C. and grant number 451-10-021 to B.G.). The funding source had no involvement in: preparation of the article, design of the study, collection, analysis and interpretation of data, writing of the report, or in the decision to submit the article for publication.

Acknowledgements

The authors would like to thank all adolescents and their parents for participating in the study, Marthe de Jong, Ursula Steenwijk and Jorien van Hoorn for their assistance during data collection and Marieke Buij for her help in contacting participants and their parents.
