Title: Feeling left out or just surprised? Neural correlates of social exclusion and expectancy violations in adolescence

Abbreviated title: Neural correlates of exclusion and expectancy violation

Authors: Theresa W. Cheng*[1], Nandita Vijayakumar*[1,2], John C. Flournoy[1,3], Zdena Op de Macks[1,4], Shannon J. Peake[1], Jessica E. Flannery[1], Arian Mobasser[1], Sarah L. Alberti[1,5], Philip A. Fisher[1], Jennifer H. Pfeifer[1] ; * Co-first authors

Affiliations: [1]Department of Psychology, University of Oregon, Eugene, Oregon, USA

[2]School of Psychology, Deakin University, Melbourne, Australia

[3]Department of Psychology, Harvard University, Cambridge, Massachusetts, USA

[4]Department of Psychology, Leiden University, the Netherlands

[5]Department of Behavioral Neuroscience, Oregon Health & Science University, Portland, Oregon, USA

Corresponding author: Jennifer H. Pfeifer

Address: Department of Psychology // 1227 University of Oregon // Eugene, OR, 97405

Email: jpfeifer@uoregon.edu

Conflict of Interest: The authors declare no competing financial interests.

Abstract

Social belonging and affiliation are important human drives that impact decision-making and health outcomes. While neural responses to social exclusion are increasingly well-characterized, studies rarely identify whether these responses reflect the affective distress or expectancy violation associated with exclusion. The present study compares neural responses to exclusion and over-inclusion, a similarly unexpected social occurrence that violates fair play expectations, but does not involve rejection, with a focus on implications for models of dorsal anterior cingulate cortex (dACC) function. In an fMRI adaptation of the Cyberball paradigm with human adolescents 11.1-17.7 years of age (N=69), we employed parametric modulators to examine scaling of neural signal with cumulative exclusion and inclusion events. The use of parametric modulators overcomes arbitrary definitions of condition onsets and offsets imposed on fluid, continuous gameplay. In support of the notion that the dACC plays a role in processing affective distress, this region exhibited greater signal with cumulative exclusion than cumulative inclusion events. However, this difference was partly driven by diminished signal accompanying cumulative inclusion, a finding that does not conform to either affective distress or expectancy violation model predictions. Additionally, we found that signal in the insula and dorsomedial prefrontal cortex scaled uniquely with social exclusion and over-inclusion, respectively, while conjunction analyses revealed that the rostromedial prefrontal cortex and left intraparietal sulcus responded similarly to both conditions. These findings shed light on which neural regions may exhibit patterns of differential sensitivity to exclusion or over-inclusion, as well as those that are more broadly engaged by both types of social interaction.

Keywords: Cyberball, social exclusion, functional MRI, anterior cingulate cortex

1. Introduction

From the earliest fMRI studies on social rejection, it has been proposed that the dorsal anterior cingulate cortex (dACC) and anterior insula process the affective distress underlying both physical and social pain (Eisenberger, Lieberman, & Williams, 2003; Eisenberger, Jarcho, Lieberman, & Naliboff, 2006). This hypothesis suggests that neural resources dedicated to pain processing were co-opted over the course of evolution to make the experience of social rejection particularly salient, motivating the maintenance of close social ties and ultimately promoting evolutionary fitness (Baumeister and Leary, 1995; Eisenberger, 2012). While recent studies suggest that pain and social rejection have distinct representations in the dACC (Woo et al., 2014; Kragel et al., 2018), many perspectives maintain a role for this region in processing rejection-related affective distress.

In commonly employed social rejection paradigms, the affective distress of rejection co-occurs with unanticipated violations of fairness norms, i.e., expectancy violations. This is true in typical implementations of Cyberball (Williams et al., 2000), a virtual ball-tossing game that is commonly used to simulate social exclusion in neuroimaging studies (Vijayakumar et al., 2017). Studies that have sought to disentangle this confound provide some support for an alternative model of ACC function that posits a dorsal-cognitive/ventral-affective regional division (Bush et al., 2000). For example, some studies with young adults have identified that the dACC is sensitive to expectancy violations (Somerville et al., 2006; Kawamoto et al., 2012), while the vACC is sensitive to social feedback (Somerville et al., 2006). However, other findings do not fully support this dissociation (Bolling et al., 2011b; Kawamoto et al., 2012). As such, there is inconsistent evidence regarding the contributions of ACC subregions to processing aspects of social exclusion.

Inconsistencies are also seen in results from three quantitative meta-analyses examining neural responses to social exclusion across paradigms. The following ACC subregions were associated with social exclusion across meta-analyses: (1) rostral perigenual ACC bordering on mPFC (Cacioppo et al., 2013), (2) perigenual and subgenual ACC (Vijayakumar et al., 2017); and (3) vACC and anterior dACC (Rotge et al., 2015). In comparison, the ventromedial PFC, ventrolateral PFC/lateral OFC, and anterior

insula were more reliably recruited than ACC subregions across whole-brain meta-analyses (Cacioppo et al., 2013; Vijayakumar et al., 2017). Whether the involvement of neural regions identified across studies reflects processes unique to the affective distress of exclusion rather than social expectancy violation remains unknown due to widespread design confounds.

The present study investigates the specificity of neural responses to processing the affective distress or expectancy violation associated with social exclusion. We employ an fMRI adaptation of Cyberball containing periods of over-inclusion and exclusion interspersed with periods of fair play. This enables comparisons of social exclusion and over-inclusion, a condition in which social expectations are violated without eliciting feelings of social rejection. Kawamoto and colleagues (2012) have also employed this strategy by comparing social exclusion and over-inclusion. However, this approach was limited by arbitrary operationalization of inclusion and exclusion on continuous gameplay (e.g., the Nth throw between computer players marks the onset of exclusion). We reduce such constraints via the use of parametric modulators to analyze changes in BOLD signal associated with cumulative exclusion and inclusion events.

We interpret our results with respect to affective distress and expectancy violation models of ACC function. Because testing these models hinges on interrogating the ACC, we examine signal in eight ACC parcels that comprise our regions-of-interest (ROIs). We also employ whole-brain analyses to allow for the possibility that smaller clusters within the ACC meet evidentiary standards for either model, even if their signal is diluted within a parcel due to averaging over larger regions or non-optimal parcel boundaries. We interpret affective distress models to imply that dACC signal will scale with cumulative exclusion but *not* cumulative inclusion events. In other words, we look for regions within the ACC that show increased signal with cumulative exclusion relative to cumulative inclusion, which are specifically driven by sensitivity to exclusion rather than inclusion. In contrast, we interpret expectancy violation models to imply that dACC signal will scale similarly to both types of cumulative events. Therefore, we look for regions within the ACC that do not show differences in cumulative exclusion relative to cumulative inclusion *and* show significant responsiveness to both cumulative exclusion and cumulative

inclusion. Finally, additional analyses explore the specificity of neural responses in social rejection paradigms by identifying responses to short-term expectancy violations, i.e., when a ball toss including or excluding the participant is incongruent with the broader inclusionary or exclusionary context established by the run.

We also examine age-related differences in our sample of 11-17 year-olds, as the neural correlates of processing social exclusion are known to differ between adolescents and adults (Vijayakumar et al., 2017). Developing the capacity to navigate complex social relationships is a critical aspect of adolescence (Blakemore and Mills, 2014). Social evaluation becomes a dominant concern (Somerville, 2013) during this period, with adolescents spending a significant amount of time with peers both on and offline (Csikszentmihalyi and Larson, 1984; Anderson and Jiang, 2018) Accordingly, adolescents experience more severe affective responses to negative social feedback than adults (Sebastian et al., 2010). Such shifts in peer interest and rejection sensitivity may be driven by protracted maturation of neural circuitry during this phase of life (Nelson et al., 2005), and are particularly worth investigating because experiences of peer rejection contribute to the emergence of developmental psychopathology (Deater-Deckard, 2001; Nelson et al., 2005).

2. Materials & Methods

2.1 Participants

Ninety-seven adolescents between the ages of 11 to 17 were recruited from Lane County, Oregon. These participants form the community control group in a study that included additional samples from the child welfare and juvenile justice systems. Anticipating that mental health diagnoses might be prevalent in future high-risk samples, we did not exclude for reported current or prior diagnoses of common psychiatric disorders, including anxiety, depression, attention deficit disorder/attention deficit hyperactivity disorder, oppositional defiant disorder, or conduct disorder. However, no participants from the high risk group are included in any of the analyses presented in this paper. Of the 97 participants recruited from the community sample only, nine dropped out of the study, seven elected not to participate

in the MRI portion, eight completed an alternate (pilot) version of the Cyberball task, one failed to complete the Cyberball task, and one participant's data was not collected due to technical errors. Additionally, two participants were excluded when visual quality inspection of the imaging data revealed extreme motion and/or orbitofrontal signal dropout.

The analyses presented here were conducted with the remaining 69 participants from the community sample aged 11 to 17 years (range= 11.1 to 17.7, M=14.2, SD=1.5; 36 female). Our final sample included four subjects that disclosed psychiatric diagnoses, including two with anxiety disorders and two with attention-related disorders. Six participants reported taking medication for behavioral issues (four subjects reported taking anti-depressant medications, one reported taking a mood stabilizer, and one reported taking stimulants typically prescribed for attention disorders). In total, eight participants reported either receiving psychiatric diagnoses or taking psychotropic medications. We conducted sensitivity analyses by re-running our ROI and whole-brain analyses with the parametric modulators in a sample of 61 participants, omitting those with such diagnoses and medications (see section 1 of the Supplementary Materials).

2.2 Procedures

2.2.1 Study Procedure. Written informed consent was obtained from the parent/guardian, while written assent was obtained from the adolescent. At the first visit, the adolescent was screened for MRI eligibility as per procedures determined by the University of Oregon's Lewis Center for Neuroimaging. The second visit was typically scheduled within one month of the first visit (M=14.8 days, SD=17.1, range=1 to 106). Eligible participants then completed the MRI portion of the study as well as additional questionnaires and behavioral tasks, including the vocabulary and matrix reasoning subtests of the Weschler Abbreviated Scale of Intelligence, second edition (Wechsler, 1999). Participants were compensated for their time, and all materials and procedures were approved by the Institutional Review Board at the University of Oregon.

2.2.2 The Cyberball task. Developed by Williams, Yeager, Cheung, and Choi (2012), Cyberball is an animated interactive ball-tossing computer game used to virtually simulate the experience of social

exclusion. In our study, participants were informed that they were playing Cyberball with two adolescent peers at partner sites in real time via the Internet. However, throws during the game were computer-automated. Similar adaptations of Cyberball have successfully simulated peer rejection in adolescents by leading them to believe that computer players were real peers (Bolling et al., 2011a; Masten et al., 2009). Participants made short video profiles to introduce themselves to the other players and viewed the other players' profiles, further lending credibility to the cover story; to our knowledge, this has not been done in prior studies. Participants were instructed on how to play Cyberball in a mock scan session immediately prior to their MRI session.

Participants played an over-inclusion run of Cyberball followed by an exclusion run, with decision-making tasks between Cyberball runs. Each run contained a total of 44 ball throws. They were instructed to use a button box to indicate which of the two players they wanted to throw to, and if they did not make a decision within 5 s, the ball would be automatically thrown to another player at random. Traditional Cyberball paradigms seek to induce strong feelings of rejection and thus consist of extended periods of exclusion, sometimes taking up close to the entirety of the run. We instead employed periods of exclusion or over-inclusion interspersed with periods of fair play to reflect less stark interactions. In the inclusion run, participants experienced periods of over-inclusion in which computer players repeatedly threw the ball to the participant rather than to one another. Periods of over-inclusion were interspersed with periods of fair play such that, overall, 17 of the 27 throws by the other players were to the participant (63%). In the exclusion run, participants experienced periods in which the computer players only threw the ball to one another and not to the participant. These were interspersed with periods of fair play such that, overall, six of the other players' total combined 38 throws were to the participant (16%). In social judgment and evaluation tasks, adolescents' (ages 11-17) mean expectations of being liked are between 40-55%, and on average, such ratings are lower in mid-adolescence, i.e., between the ages of 12-14 (Gunther Moor et al., 2010; Rodman et al., 2017 p.201). Although our choice of paradigm differs, we suggest that the rates of social exclusion (16% of throws to the participant) and over-inclusion (63% of throws to the participant) in our study are extreme enough to produce a violation of expectations in both

conditions across participants. (Note that these percentages are true of the overall run, and within each run participants experienced even more extreme periods of over-inclusion or exclusion interspersed with periods of fair play). As other computer players still both received the ball some of the time and at approximately even rates, it is generally the case that neither computer player was explicitly excluded by greater participant inclusion. The time elapsed between each computer player catching and throwing the ball was a random number between 0 and 3 s (M=1.5, SD=0.9), and the ball took approximately 1.4 s to travel through the air. Therefore, the timing of events, including participant button presses, varied from trial to trial and did not systematically align with the TRs. See Figure 1 for a schedule of throws in both runs.

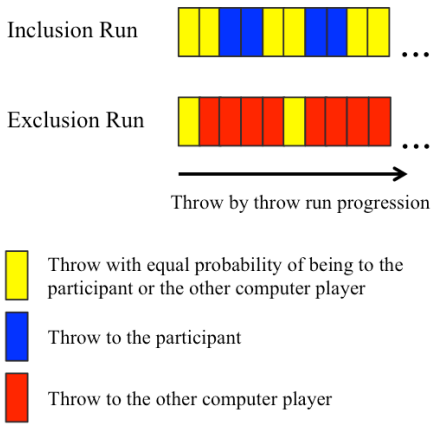

Figure 1. Partial schedules of throws for each of the computer players in both runs are displayed. Each rectangle represents one throw by the computer player. The color of the rectangle indicates the recipient of the ball. The complete schedules extend the pattern displayed to the total number of computer programmed throws per run, and both of the computer players follow the same schedule within a run.

In the protocol for this study, participants first played two rounds of the Yellow Light Game (YLG; more details about this task and the overall protocol can be found in (Op de Macks et al., 2018), a driving simulation adapted from the Stoplight Task (Gardner and Steinberg, 2005) outside of the scanner.

Afterward, within the scanner, participants (a) played two rounds of the YLG, (b) played the Cyberball inclusion run, (c) were introduced to virtual peers and observed them playing the YLG, (d) played two peer-observed runs of YLG, (e) played the Cyberball exclusion run, and (f) played two final peer-observed runs of YLG. The total length of the scan protocol was 1 hr 15 min. The inclusion run always preceded the exclusion run; Cyberball conditions were not counterbalanced due to concern about carryover of negative affect following exclusion. Participants were led to believe that peers from the YLG observations were the same as those in the Cyberball game.

2.2.3 Behavioral measures. Following the MRI scan (and typically within 30 minutes of the end of the scan), adolescents completed the Need-Threat Scale (NTS) to assess their subjective experience of distress during Cyberball. We used the 12-item version (Zadro et al., 2004), which has been validated as an overall measure of need-threat (Gerber et al., 2017). The 12 items on the NTS included ratings of belongingness, control, meaningfulness, and self-esteem on a scale from 1 ("not at all") to 5 ("very much so"), with higher NTS scores reflecting more threat. The overall scale exhibited good reliability in our sample (standardized Cronbach's α=0.81). We assessed the believability of the deception by asking participants "Did you think the peers could actually see you playing [the driving game]?" This question was part of a task experience survey administered at the very end of the session, just prior to participant debriefing. From this self-report, we inferred whether or not participants believed they were interacting with real peers during Cyberball.

We were interested in evaluating adolescents' baseline expectations of social inclusion and exclusion. Adolescents completed a modified version of the Rejection Sensitivity Questionnaire (RSQ; (Downey et al., 1998) with 10 total scenarios rather than the 12 from the original survey. Our analysis did not use the full RSQ, as our interest was exclusively in expectations of acceptance or rejection, rather than anxious or angry perceptions and reactions to rejection. We conducted our analyses on the questions (one per vignette) in which adolescents self-reported their expectations of social acceptance on a six-point Likert scale ("Extremely unlikely" to "Extremely likely"), where higher scores reflected greater expectations of acceptance. (Example item: "I would expect that the person would want to hang out with

me.") The subset of questions pertaining to expectations of acceptance exhibited fairly good internal consistency in our sample (standardized Cronbach's α=0.80). Two participants did not have responses on the RSQ, and their scores were mean-imputed in the brain-behavior analyses. We implement the abbreviation "RSQ-expectations" to refer to averaged scores among the subset of RSQ items pertaining to expectations of acceptance.

2.2.4 fMRI data acquisition. Data were acquired on a 3T Siemens Skyra MRI scanner at the Lewis Center for Neuroimaging at the University of Oregon. High-resolution T1-weighted structural images were collected with the MP-RAGE sequence (TE=3.41 ms, TR=2500 ms, flip angle=7°, 1.0 mm slice thickness, matrix size=256 x 256, FOV=256 mm, 176 slices, bandwidth=190 Hz/pixel). Two functional runs of T2*-weighted BOLD-EPI images were acquired with a gradient echo sequence (TE=27 ms, TR=2000 ms, flip angle = 90°, 2.0 mm slice thickness, matrix size=100 x 100, FOV=200mm, 72 slices, bandwidth=1786 Hz/pixel). There were 60 to 87 images per run, as run length varied with participants' response times during Cyberball. To correct for local magnetic field inhomogeneities, a field map was also collected (TE=4.37 ms, TR=639.0 ms, flip angle=60°, 2.0 mm slice thickness, matrix size=100 x100, FOV=200 mm, 72 slices, bandwidth=1515 Hz/pixel).

2.2.5 fMRI processing. Raw DICOM image files were converted to the NifTI format with MRIConvert (Smith, 2011). Data were preprocessed using Statistical Parametric Mapping software (SPM12, Wellcome Department of Cognitive Neurology, London, UK). Participants' anatomical images were coregistered to the 152 Montreal Neurological Institute stereotaxic template, segmented into six tissue types, and used to create a group anatomical template using DARTEL. Next, functional images were unwarped using field maps and corrected for head motion via image realignment. A group averaged field map was created and used as a substitute for two participants: one that did not have a field map and another whose fieldmap was not well-aligned with their functional volumes. Unwarped and realigned mean functional images were coregistered to the anatomical image (that had in turn been coregistered to the MNI template). Transformations were applied to warp the functional data to the DARTEL group template, and these normalized images were smoothed using a 6-mm FWHM Gaussian kernel.

Preprocessing scripts used for this analysis are available on GitHub at https://github.com/dsnlab/TDS_scripts/tree/cheng_cyb_main/fMRI/ppc/spm/tds2 (SPM scripts) and https://github.com/dsnlab/TDS_scripts/tree/cheng_cyb_main/fMRI/ppc/shell/schedule_spm_jobs/tds2 (shell scripts).

Motion artifacts were identified using an in-house automated script that evaluates changes in image intensity relative to the mean across all subjects, as well as volume-to-volume changes in Euclidean distance. This regressor of no interest was constructed by marking volumes of the following types: (a) volumes with greater than 0.3 mm of motion in Euclidian distance relative to the previous volume, (b) volumes for which the *mean intensity across voxels* was extreme (3 SDs above or 1.5 SDs below) relative to the mean intensity across subjects and runs, and (c) volumes for which the *standard deviation across voxels* was extreme (3 SDs above or below) relative to the mean standard deviation across subjects and runs. The mean intensity and standard deviation scores were calculated across all runs for all subjects, including volumes collected during the YLG and while participants observed others playing the YLG. Volumes immediately preceding and following marked volumes were also flagged. This script identified head motion in 36 out of 138 total Cyberball runs across 69 participants. Of the volumes marked for motion, the script flagged an average of 3.94 (5.7%) volumes per run (SD=3.58, range=1 to 17, or up to 23.3%). Additionally, our models included four motion parameters (absolute distance, absolute rotation, and the first derivatives of each of these values) as regressors of no interest. As mentioned previously, two participants were excluded on the basis of head motion/signal dropout. The first was identified for exclusion based on visual inspection of pre-processed images that revealed extreme signal dropout in the orbitofrontal gyrus. The second participant was identified for exclusion based on initial visual inspection of contrasts from their single-subject models, which revealed a clear pattern of motion-related striping that indicated that their data was contaminated. No participants were excluded for exceeding an *a priori* threshold of >25% of marked motion volumes within a single run.

2.3 Experimental Design and Statistical Analysis

Cyberball was modeled as an event-related design. For each participant's fixed-effects analysis, a general linear model was created with two regressors of interest modeled as zero-duration events: the computer-generated throws were each modeled as either an Inclusion Event (IncEvent, i.e. throws to the participant) or an Exclusion Event (ExcEvent, i.e. throws to the computer players). These events occurred within an Inclusion Context (IncContext, i.e. inclusion run) or an Exclusion Context (ExcContext, i.e. exclusion run). An additional zero-duration event regressor of no interest marked when participants' computer avatar "caught" the ball, signaling the participant's turn to throw the ball. This regressor was added to control for neural responses related to participants' decision-making and subsequent button-press, as has been used in previous studies with event related designs (Bolling et al., 2011b). There was also a "wait" period at the start of each run, consisting of 6 s where the screen displayed the words "Connecting to other players…" along with a progress bar, plus additional time until the first throw of the game (on average 8 s in total).

Parametric modulators were created to track consecutive IncEvents within the IncContext (referred to as Increasing Inclusion) and consecutive ExcEvents within the ExcContext (referred to as Increasing Exclusion). Parametric modulators were not created for ExcEvents in the IncContext or vice versa due to an insufficient number of such events in respective runs. Each parametric modulator was mean-centered relative to the average number of continuous throws of that type (for the inclusion parametric modulator: M=2.73, SD=1.82; for exclusion: M=5.97, SD=3.89). The model was convolved with the canonical hemodynamic response function, and parameter estimates from the GLM were used to create six linear contrast images: one for each of the four types of events (*i.* IncEvent in IncContext, *ii.* ExcEvent in IncContext*, iii.* IncEvent in ExcContext, and *iv.* ExcEvent in ExcContext) relative to wait periods (collapsed across both runs), and one for each of two parametric modulators (Increasing Inclusion and Increasing Exclusion).

2.3.1 ROI analyses. Given the focus on the ACC in the prior literature, we selected ROIs that were independently defined using Craddock and colleagues' (2012) parcellation scheme, which is derived from cluster analyses of resting state functional neuroimaging scans. We selected eight parcels (six

regions, two were represented with separate parcels in each hemisphere) of the ACC from the 250-parcel brain map (see Figure 2). We used MarsBar (version 0.21) to extract parameter estimates of average signal associated with each of the parametric modulators within these parcels. To test the affective distress model, we look for parcels that show increased signal with Increasing Exclusion relative to Increasing Inclusion, where the difference is driven by sensitivity to Increasing Exclusion rather than Increasing Inclusion. To test the expectancy violation model, we look for parcels that do not show differences with Increasing Exclusion and Increasing Inclusion, but rather are sensitive to both. While we are primarily interested in whether the dACC is identified as meeting these logical requirements, we test parcels along the full extent of the ACC to account for models that include a dorsal-cognitive/ventral-affective split in explaining ACC functioning overall (e.g., Bush et al., 2000).

To consider all of these possibilities, we first tested whether Increasing Exclusion was significantly different from Increasing Inclusion via paired, two-tailed student's t-tests. Furthermore, we separately tested whether Increasing Inclusion and Increasing Exclusion significantly differed from zero. We control the false discovery rate at .05 using the Benjamini-Hochberg (Benjamini and Hochberg, 1995) procedure. We do not apply the Bonferroni-correction because the full set of tests are non-independent; adjacent parcels are spatially correlated both because parcels do not necessarily reflect strict signal boundaries and because of smoothing, and the results of different tests are correlated, i.e., if the difference between cumulative exclusion and cumulative inclusion is significantly different, it follows that at least one them is significantly different from zero.

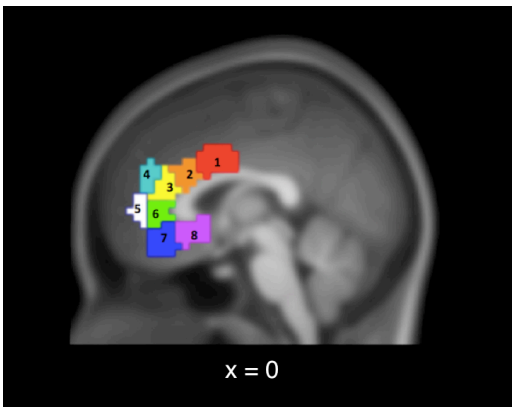

Figure 2. Regions of interest within the anterior cingulate cortex (ACC). A total of nine regions of interest were examined. Eight of these run along the extent of the ACC and were identified using the Craddock (2011) parcellation scheme with 250 parcels.

2.3.2 Whole-brain analyses. It is possible that smaller clusters within the ACC meet the evidentiary standards for the affective distress and/or expectancy violation models, but that this is not visible from the ROI analyses due to signal diluation when averaging over larger regions and/or non-optimal parcel boundaries. Whole-brain analyses allow us to examine whether smaller clusters within the ACC reflect our hypotheses. We conducted group-level analyses based on these fixed-effects (single subject) contrast images, modeling the subject variable as a random effect. We ran whole-brain conjunction and subtraction analyses with the parametric modulators using paired samples t-tests. We also used a repeated-measures flexible factorial ANOVA with a 2x2 design to examine the interaction between the event (IncEvent and ExcEvent) and context (IncContext and ExcContext) factors. This enabled comparisons between Context Congruent (i.e. IncEvent in IncContext and ExcEvent in ExcContext) and Context Incongruent (i.e. IncEvent in ExcContext and ExcEvent in IncContext) events. We also report on the main effects of event and context, although these were not main interests of the study.

Unless otherwise specified, reported results exceed the minimum cluster size threshold needed for a .05 family-wise error (FWE) rate given a voxel-wise threshold of p=.001 (bi-sided, NN=3, k=68). The core questions posed by our analyses involves comparisons of Increasing Exclusion > Increasing

Inclusion and vice versa. Thus, only these two contrasts were additionally explored at an alternative threshold, which still achieved an FWE rate of .05, albeit with a lower vowel-wise threshold of $p<.005$ and a higher cluster-extend threshold of $k=153$. Cluster extent thresholds were identified using AFNI 3dClustSim, version AFNI_17.1.01 (Apr 12 2017; version update accounts for software bugs identified in (Eklund et al., 2016). Smoothness estimates entered into 3dClustSim were spatial autocorrelation function (acf) parameters averaged from each individual's first level model residuals as calculated by 3dFWHMx (acf parameters: 0.70986 4.667 8.5925).

2.3.2.1 Age-related changes in the neural correlates of social exclusion/expectancy violation. Using whole-brain regression, we examined associations with age in both the 2x2 ANOVA and in the model with parametric modulators. We were interested in interactions between the linear age term and the context and throw factors in the 2x2 ANOVA, as well as between the linear age term and the parametric modulators. We also examined whether controlling for age in these models altered any of the results.

2.3.3 Examination of a functionally defined dACC ROI. To explore the pattern of signal change responsible for dACC signal differences in exclusion and over-inclusion, we extracted mean parameter estimates from a functionally-defined dACC ROI. This dACC ROI is a sphere with a radius of 4 mm constructed around the most ventral sub-peak of a dACC-SMA cluster (MNI coordinates: -8, 6, 38) identified from the contrast of Increasing Exclusion > Increasing Inclusion (see Figure 3A). Among cluster sub-peaks, the most ventral peak was visually identified as most clearly within the dACC and thus least likely to include SMA signal. As this ROI was selected based on significance, we visualize the pattern of signal in this ROI (Figure 3B) without conducting further null hypothesis significance testing in order to avoid the problem of circular analysis.

2.3.4 Exploratory analysis: Baseline expectations of social acceptance. We sought to evaluate whether accounting for state-level expectations of social acceptance (RSQ-expectations) altered conclusions with respect to our hypotheses. For each of the eight ACC ROIs, we used linear regression to predict parameter estimates for the inclusion and exclusion parametric modulators from RSQ-expectations (e.g., regressing $BOLD_{inclusion}$ on RSQ-expectations). Significance was evaluated with a

procedure controlling the false discovery rate (Benjamini and Hochberg, 1995) , considering adjusted p-values from the set of 16 tests. We also assessed the impact of RSQ-expectations as a covariate in the whole-brain model containing parametric modulators. Since the aim of this analysis was to examine the RSQ as a potential confound in interpreting our primary hypotheses, we did not examine the effect of RSQ on the repeated measures 2x2 ANOVA (which provides interesting additional information on the specificity of neural regions in response to social exclusion but does not test our central hypotheses).

2.3.5 Code accessibility. High motion volumes were identified using an in-house automated script that is publically available (Cosme et al., 2018). We refer interested readers to the most recent version (https://github.com/dsnlab/auto-motion), and the branch used in this analysis is available at https://github.com/dsnlab/TDS_scripts/tree/cheng_cyb_main/fMRI/fx/motion/auto-motion.

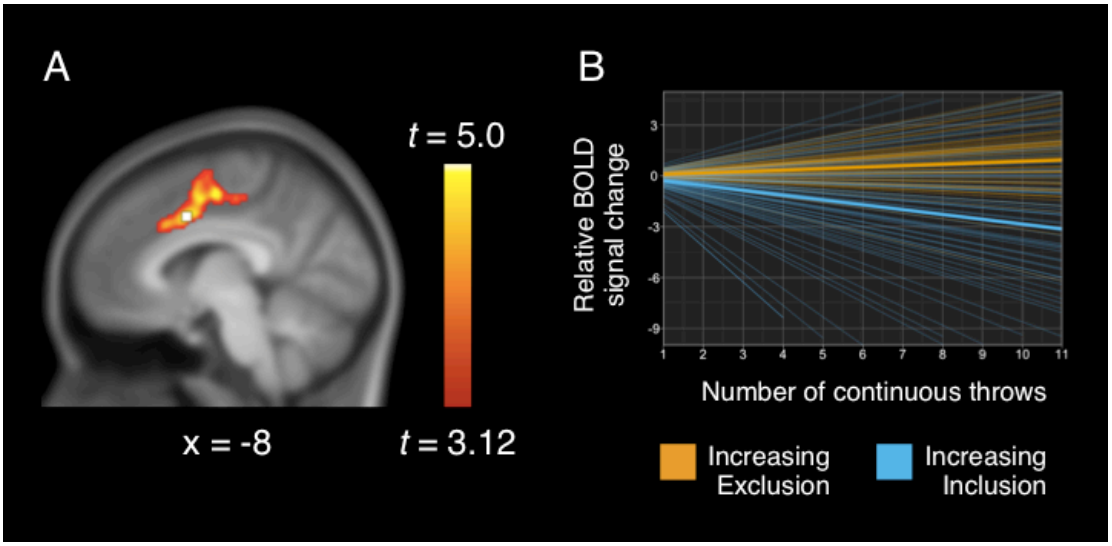

Figure 3. Examination of a functionally-defined dACC ROI. (A) The center of the 4-mm radius sphere was the most ventral sub-peak of a cluster encompassing the dACC and the supplementary motor area identified from the contrast of Increasing Exclusion > Increasing Inclusion. (B) Plot of parameters extracted from this sphere for Increasing Exclusion (yellow) and Increasing Inclusion (blue). Faint lines represent individual estimates, while bold lines reflect the group average. These parameters were

extracted for visualizaiton only, and no further statistical tests were conducted on this functionally-defined ROI.

3. Results

3.1 Behavioral results

To infer whether or not participants believed they were interacting with real peers throughout the paradigm, participants were asked in a post-task survey "Did you think the peers could actually see you playing [the driving game]?". Three participants (4%) did not respond, eight (12%) expressed disbelief, and the remaining majority (58 participants, or 84%) believed that they were interacting with real peers. We did not exclude participants on the basis of their response on this survey item because feelings of exclusion can be induced even when participants know they are playing with computer players (Zadro et al., 2004). We conducted sensitivity analyses by re-running our ROI and whole-brain analyses with the parametric modulators in a sample of 61 participants, omitting those who indicated that they did not believe the manipulation (see section 1 of the Supplementary Materials).

On average, participants reported a moderate level of subjective need threat during the Cyberball game (mean NTS score=3.23, SD=.73; higher scores reflect more threat) as measured on the NTS, comparable to levels reported in other studies using Cyberball with adolescents (Masten et al., 2009; Bolling et al., 2011b). Subjects completed the Need Threat Scale in reference to the Cyberball game in general (run unspecified), preventing analyses with self-reported distess in response to exclusion specifically. However, participants report average levels of need threat comparable to those in previous studies of exclusion only. This may be because the exclusion run took place at least ten minutes after the inclusion run, and was thus the most recent run to participants completing the scale.

Participants' mean response across ten RSQ-expectations scenarios indicated that they felt, on average, that social acceptance was "Likely" (but not very/extremely likely; M = 3.92, SD= .62 on a 6-point Likert scale). Age was not correlated with NTS (Pearson's r = .18) or expectations of acceptance (as

identified from RSQ-expectations, Pearson's r = 0.06). Furthermore, visual inspection of the scatterplots do not suggest that these variables display other developmental effects (i.e., quadratic trends).

3.2 fMRI results

3.2.1 ROI analyses. See Table 1 for the results of statistical comparisons across regions-of-interest. We test whether signal scales differently for cumulative exclusion and inclusion events in eight parcels across the ventral and dorsal extent of the ACC from Craddock and colleagues' (2012) 250-parcel map. We take affective distress models to be consistent with the presence of a dACC parcel for which signal is greater in cumulative exclusion than cumulative inclusion, such that the region is responsive to cumulative exclusion events, but not to cumulative inclusion events. On the other hand, expectancy violation is a plausible functional explanation for parcels where signal does not differ between the cumulative exclusion and inclusion, but rather is responsive to cumulative exclusion events only. We found that, of the eight parcels, Increasing Exclusion was associated with significantly greater signal than Increasing Inclusion in the most superior and caudal dACC parcel alone (ROI 1 in Figure 2; $t(68)=2.97$, $p=.004$). Signal from this dACC parcel appeared to exhibit both significant increases with Increasing Exclusion ($M=0.06$, $SE= 0.02$, $t(68)=3.02$, $p=.003$) *and* non-significant decreases with Increasing Inclusion ($M=-0.11$, $SE= 0.05$, $t(68)=-2.06$, $p=.043$). Of the remaining seven ROIs, none exhibited significant increases in *both* the Increasing Exclusion and Increasion Inclusion conditions. We note that significance is established here by examining FDR-adjusted p-values; when applying Bonferroni-correction across a total of 24 tests ($p<.0021$), there are no ROIs for which Increasing Exclusion and Increasing Inclusion are significantly different. However, such stringent correction may not be inappropriate due to non-independence of the test statistics (see section 2.3.1).

Table 1.

Evaluating parametric modulators within regions-of-interest

ROI	Increasing Exclusion M (SE)	H_0: Increasing Exclusion = 0 *t* (*p*)	Increasing Inclusion M (SE)	H_0: Increasing Inclusion = 0 *t* (*p*)	H_0: Increasing Exclusion = Increasing Inclusion *t* (*p*)
1	.06 (.02)	3.02 (3.51e-3)^{**s}	-.11 (.05)	-2.06 (.043)	2.97 (4.04e-3)^{**s}
2	.05 (.02)	3.09 (2.94e-3)^{**s}	-.03 (.06)	-.53 (.598)	1.40 (.167)
3	.08 (.02)	3.69 (4.46e-4)^{***s}	.01 (.08)	0.12 (.906)	.85 (.396)
4	.03 (.02)	1.77 (.081)	.07 (.07)	1.05 (.296)	-.56 (.579)
5	.11 (.02)	4.95 (5.17e-6)^{***s}	.13 (.08)	1.66 (.101)	-.26 (.792)
6	.06 (.02)	3.27 (1.70e-3)^{**s}	-.01 (.06)	-.15 (.878)	1.16 (.249)
7	.07 (.03)	2.89 (5.11e-3)^{**s}	.11 (.07)	1.42 (.160)	-.44 (.664)

8	.12 (.03)	4.86 (7.12e-6)^{***s}	.22 (.10)	2.19 (.032)	-.93 (.358)
dACC ROI	.08 (.02)	*NA*	-.28 (.07)	*NA*	*NA*

Caption: M=Mean beta, SE= Standard error; H_0= null hypothesis; **p<.01, ***p<.001, ^{s}: significant at FDR-adjusted *q*=0.05. Note that we do not conduct significance testing on the dACC subpeak (statistics are listed as NA), because the dACC subpeak was identified from our results for displaying greater signal with IE than II.

3.2.1.1 Examination of a functionally defined dACC ROI. We conducted descriptive analyses on a dACC ROI identified as having significantly greater signal in Increasing Exclusion as compared to Increasing Inclusion (see Figure 3A). This allowed for the possibility that a sub-region of the ACC might meet the evidentiary standard for the affective distress model even if its signal was diluted in parcel-based analyses. This ROI was examined to characterize the degree to which cumulative exclusion and inclusion contributed to the observed effect. Mean beta parameter estimates from this ROI suggest that signal differences between the parametric modulators were driven by increases with Increasing Exclusion (M=0.08, SE= 0.02) as well as decreases with Increasing Inclusion (M=-.28, SE=.07). We also sought to determine if signal in this dACC cluster had been amplified by signal from the nearby supplementary motor cortex (SMA) due to smoothing. This concern was minimized as we found the same statistical pattern of ROI results in a separate unsmoothed model.

3.2.2 Modulated BOLD response. See Table 2 for a summary of whole-brain results involving parametric modulators. Whole-brain analyses serve to complement parcel-based analyses by identifying smaller clusters for further exploratory analyses (see above). These analyses identified one functionally-defined dACC ROI for exhibiting greater signal with Increasing Exclusion than Increasing Inclusion (a

candidate for the affective distress model). We identified no ACC clusters from the conjunction of Increasing Exclusion and Increasing Inclusion that were candidates for further exploration in support of the expectancy violation model.

Other findings reported here compare and contrast neural responses to cumulative exclusion and inclusion, broadly characterizing the specificity of neural responses to different types of social interactions within Cyberball. Increasing Inclusion > Increasing Exclusion identified significant clusters in the supplementary motor area, as well as the right cuneus and sub-gyral region ventral to the cingulate gyrus. In addition, a dorsomedial PFC cluster was identified at an exploratory threshold that maintains FWE=.05 (lower voxel-wise threshold of $p<0.005$ and higher cluster extent threshold of $k>183$). The contrast of Increasing Exclusion > Increasing Inclusion identified the left motor cortex (encompassing pre- and postcentral gyri and extending into the insula), as well as the right posterior insula and supplementary motor area extending into the dACC. Examination of this contrast at the same lower voxel-wise threshold yielded no additional clusters. The conjunction of Increasing Inclusion and Increasing Exclusion identified shared signal decreases in the left intraparietal sulcus, as well as shared increases in the rostromedial PFC, as illustrated in Figure 4. The conjunction of signal positively associated with Increasing Exclusion and negatively associated with Increasing Inclusion identified clusters in the left primary motor cortex, supplementary motor area, and left anterior insula. Statistical maps are available on NeuroVault (https://neurovault.org/collections/3794).

Table 2.

Whole-brain results with parametric modulators

Condition	Peak Region	BA	Extent	Peak t-statistic	x	y	z
II>IE	R cuneus		192	4.80	18	-100	16
	R sub-gyral region, inferior to cingulate gyrus		87	4.50	-18	-12	26
	R SMA	6	74	4.04	6	-22	66
IE>II	L postcentral gyrus extending into insula	3	3177	9.06	-42	-18	56
	SMA/dACC		1055	7.30	0	-2	52
	L precentral gyrus	6	112	4.84	-58	6	30
	R posterior insula	13	101	4.48	48	-20	20
II & IE	Rostromedial PFC	9	107	4.04	-4	58	22
II- & IE-	L intraparietal sulcus		94	4.50	-34	-48	38
II- & IE	L postcentral gyrus extending into precentral gyrus	2, 3, 4	975	6.57	-42	-18	54
	Supplementary motor area	6	234	5.07	0	-2	48
	Left anterior insula	13	122	4.22	-42	-22	20

Caption: IE: Increasing Exclusion; II: Increasing Inclusion; -: negative association; PFC: prefrontal cortex; ACC: anterior cingulate cortex; SMA: supplementary motor cortex; BA: Brodmann's area; *Cluster is no longer present when controlling for age. No effects were identified in the conjunction of IE- and II+. No linear effects of age or interactions between age and our parametric modulators were

identified. Results are FWE cluster corrected at p<.05 (voxel-wise p<.001, k=68). Our primary contrasts of interest (IE > II, II > IE, and the conjunction of II and IE) were also examined at a lower voxel-wise threshold (p<0.005) which, in conjunction with a higher spatial extent threshold (k>153) also achieved a p<0.05 family-wise error rate (calculated by 3dClustSim). Comparisons of II>IE at this lower voxel-wise threshold identified an additional cluster in the dorsomedial prefrontal cortex (Broadmann's area 8, 284-voxel extent, located at coordinates 2, 44, 50), while the other two contrasts did not identify any novel clusters. Whole-brain findings associated with the parametric modulators themselves are available in Supplementary Table E.

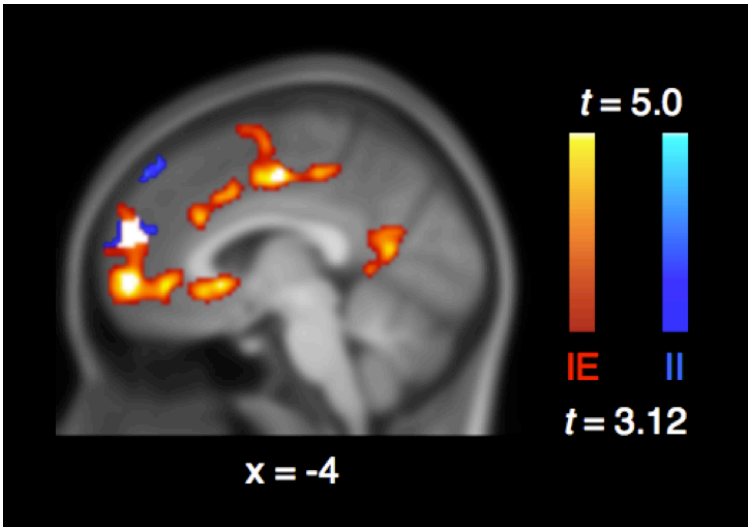

Figure 4. Conjunction of Increasing Exclusion and Increasing Inclusion. The rostromedial prefrontal cortex is identified from the conjunction (white) of parametric modulators for Increasing Exclusion (red) and Increasing Inclusion (blue). Results are FWE cluster-corrected at p<.05 (voxel-wise p<.001, k=68).

3.2.3 Repeated measures 2x2 ANOVA For detailed whole-brain results of this 2x2 ANOVA, including main effects of event and context, see Table 3. To characterize the specificity of neural responses during Cyberball, we examined the interaction between context and event type, comparing when these two factors were congruent (i.e., IncEvent-IncContext and ExcEvent-ExcContext) versus

incongruent (i.e., IncEvent-ExcContext and ExcEvent-ExcContext). The contrast of Context Incongruent > Context Congruent events identified the bilateral putamen, bilateral inferior parietal lobule (extending into right posterior superior temporal sulcus and posterior insula), bilateral inferior frontal gyrus, left anterior insula, left precentral gyrus, left precuneus, left posterior superior temporal sulcus, left middle frontal gyrus, right superior temporal gyrus, supplementary motor area, cerebellum, thalamus, and dACC. See Figure 5 to view the left insula cluster. Comparing Context Congruent > Context Incongruent events yielded no significant clusters.

Table 3.

Whole-brain results with 2x2 factorial ANOVA

Condition	Peak Region	BA	Extent	Peak t-statistic	x	y	z
ExcEvent> IncEvent	L middle occipital gyrus extending into bilateral temporal-occipital cortex	19	11332	8.78	-50	-72	10
	Bilateral inferior parietal lobule		2997	6.25	-46	-38	60
			834	4.77	40	-40	44
	Bilateral middle frontal gyrus	6	4629	5.83	30	20	56
		46	772	5.39	-48	48	2
	L orbitofrontal cortex		90	5.60	-24	32	-24
	L posterior insula		163	5.41	-40	-12	16

	L postcentral gyrus/superior temporal gyrus	40	127	5.30	-60	-24	14
	L postcentral gyrus	3	165	5.05	-42	-18	40
	Temporal pole extending into middle temporal gyrus	21	271	5.13	40	12	-46
	L parahippocampal gyrus extending into superior temporal gyrus	20	112	4.72	-26	-4	-44
	R supramarginal gyrus	22/40	186	4.30	64	-44	40
	R uncus*	20	68	4.19	30	-6	-42
	R dorsolateral PFC	9	123	3.84	46	26	32
IncEvent > ExcEvent	L precentral gyrus		1231	8.62	-34	-18	52
	R medial frontal gyrus*	6	395	6.03	16	-12	66
	R thamalus (nucleus?)		72	3.76	10	-12	-4
ExcContext >IncContext	R precentral gyrus		194	4.74	54	-2	16
		4	157	4.73	52	-4	42
	R inferior parietal lobule		236	4.16	58	-22	24
	L posterior insula		103	3.83	-42	-18	20
IncContext> ExcContext	Bilateral inferior parietal lobule	40	545	4.99	-48	-34	46
			269	4.82	42	-38	40

	Bilateral superior and middle frontal gyri	6	317	4.66	-24	0	56
			328	4.47	24	-2	52
	Bilateral precuneus		233	4.43	18	-72	56
		7	182	4.25	-18	68	58
Context Incongruent > Context Congruent	Bilateral putamen		176	5.51	-14	0	4
			89	4.36	22	8	4
	L precentral gyrus	9	188	5.46	-46	2	36
		6	254	4.49	-20	-10	58
	Bilateral inferior parietal lobule, R extends into posterior superior temporal sulcus and posterior insula	40	986	5.34	60	-30	30
		40	115	4.19	-46	-36	24
	Supplementary motor area	6	417	5.15	4	2	54
	L anterior insula		147	4.75	-30	22	-6
	R thalamus (medial dorsal nucleus)		104	4.73	10	-16	6
	Bilateral inferior frontal gyri	9	196	4.73	-56	6	22
			84	3.97	42	8	22

R superior temporal gyrus		123	4.32	50	-10	-16
L precuneus		85	4.13	-18	-76	40
L posterior superior temporal sulcus		274	4.04	-40	-52	16
L middle frontal gyrus	9	72	4.04	-26	34	22
R cerebellum		184	4.02	28	-50	-12
dACC	24	71	3.88	-8	6	38

Caption: IncEvent: inclusion event, i.e., throw to the participant; ExcEvent= exclusion event, i.e., throw to computer player; IncContext: inclusion context, ie., the inclusion run; ExcContext: exclusion context, i.e., the exclusion run; Context Congruent: events which match the context, i.e., IncEvents in the IncContext and ExcEvents in the ExcContext; Context Incongruent: events which do not match the context, i.e., IncEvents in the ExcContext and ExcEvents in the IncContext; *Cluster is not present when controlling for age. There were no main effects of IncContext>ExcContext or Context Congruent > Context Incongruent. Results are FWE cluster corrected at p<.05 (voxel-wise p<.001, k=68).

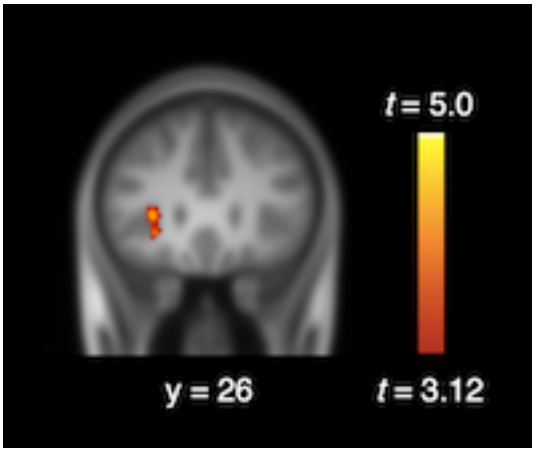

Figure 5. Context Incongruent > Context Congruent results of the event by context interaction. Sagittal and coronal views display the left anterior insula cluster. Results are FWE cluster corrected at p<.05 (voxel-wise p<.001, k=68).

3.2.4 Age. There were no interactions between age and the parametric modulators. Age effects and age interactions in the repeated measures 2x2 ANOVA did not reveal clusters in the ACC or insula, and findings from these analyses are reported in the Supplementary Materials. Controlling for age did not substantively alter findings from either the parametric modulators or the repeated measures 2x2 ANOVA (minor changes noted in the Supplementary Materials).

3.2.5 Expectations of acceptance. In the eight ACC regions-of-interest, expectations of acceptance (RSQ-expectations) were not significantly associated with signal related to increasing exclusion or inclusion events, even with a lenient threshold uncorrected for multiple comparisons ($p<.05$). Furthermore, whole-brain analyses controlling for expectations of acceptance did not substantively alter the main findings for the parametric modulators. However, one novel cluster was identified in the left middle temporal gyrus (MNI coordinates: -58, -12, -12; k = 73) in the conjunction of increasing inclusion and exclusion after controlling for self-reported state expectations of acceptance.

4. Discussion

This study probes the specificity of neural responses to social exclusion and evaluates models of dACC function that differentially emphasize its involvement in processing affective distress versus expectancy violations. We disentangle these processes by comparing social exclusion to over-inclusion, a condition that elicits social expectancy violations without the experience of rejection.

4.1 Modulated BOLD responses

4.1.1 ROI analyses of ACC involvement in affective distress versus social expectancy violation. We interpret affective distress models to imply that dACC signal increases with Increasing Exclusion, but not Increasing Inclusion; in contrast, expectancy violation accounts predict similar dACC responses to both conditions. Of the eight ACC parcels examined, only the most caudal dACC parcel exhibited greater average signal with Increasing Exclusion as compared to Increasing Inclusion. This parcel spatially overlapped with the dACC ROI functionally-defined from whole brain analyses. An examination of the parametric modulators shows that greater signal in the dACC for Increasing Exclusion

than Increasing Inclusion was driven not only by significant increases in signal with Increasing Exclusion, but also by non-significant *decreases* in signal with Increasing Inclusion. As this latter result was non-significant, the statistical pattern of findings is consistent with the affective distress model. However, these signal decreases with Increasing Inclusion are notable in magnitude, and reflect a significant caveat suggesting further examination. Intriguingly, both social acceptance and rejection have been demonstrated to engage a similar region of the dACC in a study using a social feedback task (Dalgleish et al., 2017); however, in this study, the region responded positively to both conditions. Future studies with detailed measures of participant alertness and affective state might address the possibility that there is a non-zero baseline for affective distress and/or cognitive monitoring in the scanner environment, and that signal decreases with Increasing Inclusion are due to attenuation of these processes over time.

4.1.2 Whole-brain analyses. Neural responses to the parametric modulators also differed in regions aside from the dACC. The dorsal aspect of the mPFC (identified at a joint threshold of lower magnitude and higher spatial extent threshold) was relatively more responsive during Increasing Inclusion as compared to Increasing Exclusion. This potentially reflects the role of dorsal mPFC regions in complex social cognitive processes (Masten et al., 2011; Eickhoff et al., 2014; Martin et al., 2017) particularly for more effortful social cognitive tasks (Hartwright et al., 2013). We therefore speculate that for typically-developing adolescents, social cognition may be engaged to a greater extent during cumulative inclusion relative to cumulative exclusion. Meanwhile, comparing Increasing Exclusion to Increasing Inclusion identified the posterior insula. The posterior insula is thought to reflect visceral and autonomic states (Craig, 2003), contributing to the sensory rather than affective aspects of pain (Eisenberger, 2012), and has also been proposed to play a role in particularly intense affective experiences (Kross et al., 2011). Our findings suggest that, controlling for expectancy violation, signal in this region scales with the intensity of social exclusion.

Prior meta-analyses have consistently implicated the anterior PFC in social exclusion paradigms (Cacioppo et al., 2013; Vijayakumar et al., 2017), with recent work indicating that this region processes negative affect (Kragel et al., 2018). Our findings suggest that this region processes unexpected and/or

affectively charged social interaction more generally, as we identified greater rostromedial PFC signal with *both* Increasing Inclusion and Increasing Exclusion. Its involvement may also reflect engagement of self-oriented processing, as the anterior PFC is implicated in perspective-taking (Ames et al., 2008), prospection (Addis et al., 2007) and self-reflection (Modinos et al., 2009), suggesting cumulative interactions may be increasingly self-relevant. Meanwhile, we identified reduced signal in the left intraparietal sulcus in both Increasing Inclusion and Increasing Exclusion. The intraparietal sulcus is part of a task positive dorsal attention network (Corbetta and Shulman, 2002), and decreases in signal with cumulative events is consistent with repetition suppression BOLD effects that have implicated this region in representing the goals of others' actions (Hamilton and Grafton, 2006), suggesting that increasing exclusion and inclusion both lead to enhanced attention and/or consideration of other players' goals. The left anterior insula was identified as exhibiting significantly lower signal with Increasing Inclusion *and* significantly greater signal with Increasing Exclusion. Visceral information from the posterior insula is re-represented by the anterior insula, which may act as an integrative hub for internal and external information (Craig, 2003; Seeley et al., 2007; Singer et al., 2009; Menon and Uddin, 2010). Pain is informed by bodily experience, and insula involvement is frequently interpreted to reflect somatosensory and affective pain in social rejection paradigms (Kross et al., 2011; Eisenberger, 2012). One interpretation of this region's involvement with both cumulative inclusion and exclusion (but in different directions) is that it tracks with negative affect. However, other interpretations are plausible (e.g., participant response demands also scale in opposite directions for cumulative exclusion and inclusion) and our paradigm is unable to disentangle these competing explanations.

4.2 Repeated measures 2x2 ANOVA

We also examined neural responses associated with violations of short-term expectancies developed through repeated events of one type establishing a context within a run. This was examined by comparing Context Incongruent (i.e., exclusion events in the inclusion context and inclusion events in the exclusion context) and Context Congruent (i.e., exclusion events in the exclusion context and inclusion events in the inclusion context) events. This contrast identified regions including the dACC, right

posterior insula, left anterior insula (dorsal to the cluster identified in the conjunction of signal negatively associated with Increasing Inclusion and positively associated with Increasing Exclusion), and the precuneus. As it seems unlikely that unexpected (Context Incongruent) events elicit distress across both contexts, involvement of these regions may not necessarily be associated with pain-related processing within these paradigms. This finding is notable as it runs contrary to widespread interpretations of these regions' involvement in social exclusion neuroimaging tasks. This further cautions against reverse inference of dACC and insula involvement as indicators of distress (unless otherwise corroborated).

4.3 Examining age and expectations of acceptance as covariates

Behaviorally, there were no effects of age on need threat or expectations of acceptance. Neural responses to cumulative interactions (i.e., models containing parametric modulators) largely exhibited neither age effects nor associations with expectations of acceptance (RSQ-expectations). These findings are consistent with evidence that Cyberball is a "strong situation" fairly robust to many individual differences in healthy samples (McDonald & Donnellan, 2012). However, they are not consistent with studies that have identified effects with continuous linear age associations (e.g., Boilling et al., 2011a) or pubertal development (Masten et al., 2013), although these studies examine different age ranges. A number of age-related findings from the repeated measures 2x2 ANOVA are discussed in section 3 of the Supplementary Materials.

4.4 Strengths and Limitations

The current study includes noteworthy methodological advancements. Parametric modulators model exclusion and inclusion-related neural changes in gameplay without arbitrary definitions of when these conditions begin and end. In contrast, the literature widely employs block and event-related designs that model social experience assuming binary, static, and independent periods or occurrences. Our modeling approach may identify neural processes that are missed by traditional block and event-related designs. This approach may better capture changes in neural signal when gameplay is more fluid, reflecting naturalistic social interactions that often involve greater ambiguity than typical laboratory tasks. Additionally, our central findings pertaining to the ACC are robust to two sensitivity analyses—one

excluding subjects who have psychiatric diagnoses and/or are taking psychotropic medications, as well as a second that excludes subjects who indicated that they did not believe that the manipulation involved real peers (see section 1 of the Supplementary Materials).

However, the results from the present study must be considered in light of several limitations. Our study did not account for possible order effects, as over-inclusion always preceded exclusion. We acknowledge that earlier runs of inclusion may influence the extent of expectancy violation in the later exclusion run, as it is possible that initially encountering expectancy violations during over-inclusion may attenuate later expectancy violations in the exclusion run. The inverse is also possible, as over-inclusion may induce higher expectations of inclusion that are then more profoundly violated by exclusion, such that the fixed order sensitizes rather than attenuates later expectancy violations. Ultimately we did not counterbalance because exclusion paradigms have been shown to induce distress among adolescents (Masten et al., 2009; Peake, Dishion, Stormshak, Moore, & Pfeifer, 2013; Sebastian, Viding, Williams, & Blakemore, 2010), and we sought to minimize the potential carryover of negative affect into the subsequent round of Cyberball. Although counterbalancing would ameliorate order effects, shuffling conditions may also significantly alter the social context of each run, subsequently altering the participant experience. Indeed, White and colleagues (2013) identified within-subjects neural differences in inclusion runs occurring before and after exclusion runs in the Cyberball paradigm. By not counterbalancing, we prioritized maintaining a consistent affective experience overall. Additionally, given the fixed order of our overall protocol, it is possible that some of the differences between conditions might be attributed to having viewed participant introduction videos and game play between the social inclusion and exclusion runs. Watching one anothers' videos and gameplay provided a plausible basis for (computer) players to negatively evaluate the participant, and was found to be a critical component to the believability of the manipulation during pilot testing.

Another limitation is that the intensity of over-inclusion and exclusion was imbalanced, with a greater average number of cumulative events in Increasing Exclusion (5.97 cumulative exclusion throws in the exclusion context) than in Increasing Inclusion (2.73 cumulative inclusion throws in the inclusion

context). Relatedly, there were few incongruent throws in each context, and a particularly low number of inclusion throws in the exclusion context. Thus, we thus have relatively less precision in our estimates of some conditions, weakening certain inferences due to our varying ability to detect effects. Additionally, it is unknown whether the true modulation of neural signal driven by social interactions is linear, particularly as interactions become prolonged (i.e., as the number of events becomes large). Our results provide a valuable estimate of linear signal scaling with a low yet affectively meaningful (based on NTS responses) number of events, but do not elucidate the true shape of the trajectory—with an average of less than three consecutive inclusion throws, we were not able to examine such effects.

Another concern is that our model may not have fully accounted for preparatory motor responses, even after modeling the period during participants' button press as a regressor (as in Bolling et al., 2011a; Bolling et al., 2011b; Bolling, Pelphrey, & Wyk, 2016). As the dACC integrates information to guide motor behavior (Paus, 2001), signal changes associated with Increasing Inclusion may have been partly driven by motor habituation. To minimize the influence of asymmetrical motor demands across conditions, future studies might require participants' button press response even when the ball is not thrown to them.

Finally, for the analyses with RSQ-expectations, we acknowledge that we only evaluate trait-level expectations, and the extent to which fluctuating state-level expectations during the Cyberball task itself impact our findings remains unknown. However, asking participants about their expectations of acceptance or rejection prior to or during any of the task runs would conspicuously draw attention to this aspect of the paradigm, impacting the psychological experience and believability of the task.

4.5 Conclusions

We used parametric modulators to examine the specificity of neural responses to social exclusion and over-inclusion during Cyberball. There was greater signal in the caudal dACC during Increasing Exclusion (i.e. consecutive exclusion events in the exclusion context) than Increasing Inclusion (i.e., consecutive inclusion events in the inclusion context), which is statistically consistent with the affective distress model. However, this difference was partly driven by diminished signal with Increasing

Inclusion, a finding that is not does not clearly conform to either affective distress or expectancy violation dACC models. Compared to Increasing Exclusion, Increasing Inclusion was associated with relatively greater BOLD signal in the dmPFC, potentially reflecting greater social cognitive engagement with cumulative inclusion. Signal increases in the rostromedial PFC were observed in both Increasing Exclusion and Increasing Inclusion, suggesting that this region plays a role in processing the unexpected and/or affectively charged nature of social interaction across both conditions. Finally, we note that responses in the dACC, inusla, and precuneus were elicited by oddball events violating short-term expectancies established by the context of the run (i.e. Context Incongruent versus Context Congruent events). These findings caution against interpreting involvement of these regions as necessarily reflecting aspects of pain processing in social rejection paradigms.

Acknowledgements: The authors wish to express gratitude to Rebecca Calcott for consultation on the analysis and to Danielle Cosme for scripts to process motion and to access brain parcellation maps.

Funding: This work was supported by the grants P50 DA035763 (PIs: Chamberlain and Fisher) and R01 MH107418 (PI: Pfeifer). Author TWC was supported by National Center for Advancing Translational Sciences of the National Institutes of Health under award number TL1TR002371. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

Addis DR, Wong AT, Schacter DL (2007) Remembering the past and imagining the future: common and distinct neural substrates during event construction and elaboration. Neuropsychologia 45:1363–1377.

Ames DL, Jenkins AC, Banaji MR, Mitchell JP (2008) Taking Another Person's Perspective Increases Self-Referential Neural Processing. Psychol Sci 19:642–644.

Anderson M, Jiang J (2018) Teens, Social Media & Technology 2018. Pew Research Center.

Baumeister RF, Leary MR (1995) The need to belong: Desire for interpersonal attachments as a fundamental human motivation. Psychol Bull 117:497–529.

Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B Methodol 57:289–300.

Blakemore S-J, Mills KL (2014) Is adolescence a sensitive period for sociocultural processing? Annu Rev Psychol 65:187–207.

Bolling DZ, Pelphrey KA, Wyk BCV (2016) Unlike adults, children and adolescents show predominantly increased neural activation to social exclusion by members of the opposite gender. Soc Neurosci 11:475–486.

Bolling DZ, Pitskel NB, Deen B, Crowley MJ, Mayes LC, Pelphrey K a. (2011a) Development of neural systems for processing social exclusion from childhood to adolescence. Dev Sci 14:1431–1444.

Bolling DZ, Pitskel NB, Deen B, Crowley MJ, McPartland JC, Mayes LC, Pelphrey K a. (2011b) Dissociable brain mechanisms for processing social exclusion and rule violation. NeuroImage 54:2462–2471.

Bush G, Luu P, Posner M (2000) Cognitive and emotional influences in anterior cingulate cortex. Trends Cogn Sci 4:215–222.

Cacioppo S, Frum C, Asp E, Weiss RM, Lewis JW, Cacioppo JT (2013) A Quantitative Meta-Analysis of Rejection. Nat Sci Rep:10–12.

Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci 3:201–215.

Cosme D, Flournoy J, DeStasio K (2018) dsnlab/auto-motion: In house (Version v0.2-alpha). Available at: http://doi.org/10.5281/zenodo.1240528.

Craddock RC, James GA, Holtzheimer PE, Hu XP, Mayberg HS (2012) A whole brain fMRI atlas generated via spatially constrained spectral clustering. Hum Brain Mapp 33:1914–1928.

Craig A (2003) Interoception: the sense of the physiological condition of the body. Curr Opin Neurobiol 13:500–505.

Csikszentmihalyi M, Larson R (1984) Being Adolescent: Conflict and Growth in the Teenage Years. New York: Basic Books.

Dalgleish T, Walsh ND, Mobbs D, Schweizer S, Harmelen A-L van, Dunn B, Dunn V, Goodyer I, Stretton J (2017) Social pain and social gain in the adolescent brain: A common neural circuitry underlying both positive and negative social evaluation. Sci Rep 7:42010.

Deater-Deckard K (2001) Annotation: Recent Research Examining the Role of Peer Relationships in the Development of Psychopathology. J Child Psychol Psychiatry 42:565.

Downey G, Lebolt A, Rincón C, Freitas AL (1998) Rejection Sensitivity and Children's Interpersonal Difficulties. Child Dev 69:1074–1091.

Eickhoff SB, Laird AR, Fox PT, Bzdok D, Hensel L (2014) Functional Segregation of the Human Dorsomedial Prefrontal Cortex. Cereb Cortex:bhu250.

Eisenberger NI (2012) The pain of social disconnection: examining the shared neural underpinnings of physical and social pain. Nat Rev Neurosci 13:421–434.

Eisenberger NI, Jarcho JM, Lieberman MD, Naliboff BD (2006) An experimental study of shared sensitivity to physical pain and social rejection. Pain 126:132–138.

Eisenberger NI, Lieberman MD, Williams KD (2003) Does rejection hurt? An FMRI study of social exclusion. Science 302:290–292.

Eklund A, Nichols TE, Knutsson H (2016) Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. Proc Natl Acad Sci 113:7900–7905.

Gardner M, Steinberg L (2005) Peer Influence on Risk Taking, Risk Preference, and Risky Decision Making in Adolescence and Adulthood: An Experimental Study. Dev Psychol 41:625–635.

Gerber JP, Chang S-H, Reimel H (2017) Construct validity of Williams' ostracism needs threat scale. Personal Individ Differ 115:50–53.

Gunther Moor B, Leijenhorst L van, Rombouts SARB, Crone EA, Molen MWV der (2010) Do you like me? Neural correlates of social evaluation and developmental trajectories. Soc Neurosci 5:461–482.

Hamilton AF de C, Grafton ST (2006) Goal representation in human anterior intraparietal sulcus. J Neurosci Off J Soc Neurosci 26:1133–1137.

Hartwright CE, Apperly IA, Hansen PC (2013) Representation, Control, or Reasoning? Distinct Functions for Theory of Mind within the Medial Prefrontal Cortex. J Cogn Neurosci 26:683–698.

Kawamoto T, Onoda K, Nakashima K, Nittono H, Yamaguchi S, Ura M (2012) Is dorsal anterior cingulate cortex activation in response to social exclusion due to expectancy violation ? An fMRI study. Front Evol Neurosci 4:1–10.

Kragel PA, Kano M, Oudenhove LV, Ly HG, Dupont P, Rubio A, Delon-Martin C, Bonaz BL, Manuck SB, Gianaros PJ, Ceko M, Losin EAR, Woo C-W, Nichols TE, Wager TD (2018) Generalizable representations of pain, cognitive control, and negative emotion in medial frontal cortex. Nat Neurosci 21:283–289.

Kross E, Berman MG, Mischel W, Smith EE, Wager TD (2011) Social rejection shares somatosensory representations with physical pain. Proc Natl Acad Sci 108:6270–6275.

Martin AK, Dzafic I, Ramdave S, Meinzer M (2017) Causal evidence for task-specific involvement of the dorsomedial prefrontal cortex in human social cognition. Soc Cogn Affect Neurosci 12:1209–1218.

Masten CL, Eisenberger NI, Borofsky LA, Pfeifer JH, McNealy K, Mazziotta JC, Dapretto M (2009) Neural correlates of social exclusion during adolescence: understanding the distress of peer rejection. Soc Cogn Affect Neurosci 4:143–157.

Masten CL, Eisenberger NI, Pfeifer JH, Colich NL, Dapretto M (2013) Associations Among Pubertal Development, Empathic Ability, and Neural Responses While Witnessing Peer Rejection in Adolescence. Child Dev 84:1338–1354.

Masten CL, Morelli SA, Eisenberger NI (2011) An fMRI investigation of empathy for "social pain" and subsequent prosocial behavior. NeuroImage 55:381–8.

McDonald MM, Donnellan MB (2012) Is ostracism a strong situation? The influence of personality in reactions to rejection. J Res Personal 46:614–618.

Menon V, Uddin LQ (2010) Saliency, switching, attention and control: a network model of insula function. Brain Struct Funct 214:655–667.

Modinos G, Ormel J, Aleman A (2009) Activation of Anterior Insula during Self-Reflection. PLoS ONE 4:1–8.

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.

Op de Macks ZA, Flannery JE, Peake SJ, Flournoy JC, Mobasser A, Alberti SL, Fisher PA, Pfeifer JH (2018) Novel insights from the Yellow Light Game: Safe and risky decisions differentially impact adolescent outcome-related brain function. NeuroImage 181:568–581.

Paus T (2001) Primate anterior cingulate cortex: where motor control, drive and cognition interface. Nat Rev Neurosci 2:417–424.

Peake SJ, Dishion TJ, Stormshak EA, Moore WE, Pfeifer JH (2013) Risk-taking and social exclusion in adolescence: neural mechanisms underlying peer influences on decision-making. NeuroImage 82:23–34.

Rodman AM, Powers KE, Somerville LH (2017) Development of self-protective biases in response to social evaluative feedback. Proc Natl Acad Sci 114:13158–13163.

Rotge J-Y, Lemogne C, Hinfray S, Huguet P, Grynszpan O, Tartour E, George N, Fossati P (2015) A meta-analysis of the anterior cingulate contribution to social pain. Soc Cogn Affect Neurosci 10:19–27.

Sebastian C, Viding E, Williams KD, Blakemore S-J (2010) Social brain development and the affective consequences of ostracism in adolescence. Brain Cogn 72:134–145.

Seeley WW, Menon V, Schatzberg AF, Keller J, Glover GH, Kenna H, Reiss AL, Greicius MD (2007) Dissociable Intrinsic Connectivity Networks for Salience Processing and Executive Control. J Neurosci Off J Soc Neurosci 27:2349–2356.

Singer T, Critchley HD, Preuschoff K (2009) A common role of insula in feelings, empathy and uncertainty. Trends Cogn Sci 13:334–340.

Smith J (2011) MRIConvert. Eugene, OR: The Lewis Center for Neuroimaging, University of Oregon.

Somerville LH (2013) The Teenage Brain: Sensitivity to Social Evaluation. Curr Dir Psychol Sci 22:121–127.

Somerville LH, Heatherton TF, Kelley WM (2006) Anterior cingulate cortex responds differentially to expectancy violation and social rejection. Nat Neurosci 9:1007–1008.

Vijayakumar N, Cheng TW, Pfeifer JH (2017) Neural correlates of social exclusion across ages: A coordinate-based meta-analysis of functional MRI studies. NeuroImage 153:359–368.

Wechsler D (1999) Wechsler Abbreviated Scale of Intelligence. San Antonio, TX: The Psychological Corporation.

White LO, Wu J, Borelli JL, Mayes LC, Crowley MJ (2013) Play it again: neural responses to reunion with excluders predicted by attachment patterns. Dev Sci:n/a-n/a.

Williams KD, Cheung CK, Choi W (2000) Cyberostracism: effects of being ignored over the Internet. J Pers Soc Psychol 79:748–762.

Woo C-W, Koban L, Kross E, Lindquist MA, Banich MT, Ruzic L, Andrews-Hanna JR, Wager TD (2014) Separate neural representations for physical pain and social rejection. Nat Commun 5:5380.

Zadro L, Williams KD, Richardson R (2004) How low can you go? Ostracism by a computer is sufficient to lower self-reported levels of belonging, control, self-esteem, and meaningful existence. J Exp Soc Psychol 40:560–567.