Biological Psychiatry: CNNI

Archival Report

Altered Neural Responses to Punishment Learning in Conduct Disorder

Erik M. Elster, Ruth Pauli, Graeme Fairchild, Maria McDonald, Sarah Baumann, Justina Sidlauskaite, Stephane De Brito, Christine M. Freitag, Kerstin Konrad, Veit Roessner, Inti A. Brazil, Patricia L. Lockwood, and Gregor Kohls

ABSTRACT

BACKGROUND: Conduct disorder (CD) is associated with deficits in the use of punishment for reinforcement learning (RL) and subsequent decision making, contributing to reckless, antisocial, and aggressive behaviors. Here, we used functional magnetic resonance imaging (fMRI) to examine whether differences in behavioral learning rates derived from computational modeling, particularly for punishment, are reflected in aberrant neural responses in youths with CD compared with typically developing control participants (TDCs).

METHODS: A total of 75 youths with CD and 99 TDCs (9–18 years, 47% girls) performed a probabilistic RL task with punishment, reward, and neutral contingencies. Using fMRI data in conjunction with computational modeling indices (learning rate α), we investigated group differences for the 3 learning conditions in whole-brain and region of interest (ROI) analyses, including the ventral striatum and insula.

RESULTS: Whole-brain analysis revealed typical neural responses for RL in both groups. However, linear regression models for the ROI analyses revealed that only the response pattern of the (anterior) insula during punishment learning was different in participants with CD compared with TDCs.

CONCLUSIONS: Youths with CD have atypical neural responses to learning from punishment (but not from reward), specifically in the insula. This suggests a selective dysfunction of RL mechanisms in CD that contributes to punishment insensitivity/hyposensitivity as a hallmark of the disorder. Because the (anterior) insula is involved in avoidance behaviors related to negative affect or arousal, insula dysfunction in CD may contribute to inappropriate behavioral decision making, which increases the risk for reckless, antisocial, and aggressive behaviors in affected youth.

https://doi.org/10.1016/j.bpsc.2025.01.003

Conduct disorder (CD) is a common mental disorder in youth characterized by severe and persistent antisocial and aggressive behavior (1). It is hypothesized that the behaviors characteristic of CD may result from deficits in learning mechanisms such as reinforcement learning (RL), which lead to unfavorable behavioral patterns that in turn contribute to impaired social functioning and can have a negative impact on quality of life (2,3). RL describes an individual's ability to adapt their behavior by comparing the expected and actual outcomes of actions (e.g., receiving or not receiving an expected reward or punishment). Behavior is more likely to change when a prediction error (PE) occurs, which signals a discrepancy between the expected and actual outcome. Research highlights that PE signals are essential for updating behavior (i.e., deciding for or against a particular action to obtain a reward or avoid punishment), thereby facilitating RL (4). Growing evidence suggests that impaired RL observed in CD may be due to difficulties in accurately assessing the value of potential behavioral outcomes, particularly the value of punishment. Behavioral deficits in learning to avoid punishment as well as in making decisions under risk (i.e., when potential losses are imminent) have consistently been shown in youth with CD, reflecting punishment insensitivity or hyposensitivity (5–7).

The neuronal underpinnings of RL have been investigated across various experimental tasks and study populations (including CD), highlighting regions such as the (ventral) striatum (VS), prefrontal areas, amygdala, and insular cortex as key brain regions for RL (8). In the context of CD, functional magnetic resonance imaging (fMRI) studies by White *et al.* (9–11) revealed reduced responses in the insula, ventromedial prefrontal cortex (vmPFC), anterior cingulate cortex (ACC), and striatum in youths with pervasive conduct problems (including those with CD and/or oppositional defiant disorder [ODD]), particularly when processing expected values while making choices to avoid punishment in a passive avoidance learning task. The meta-analysis by Alegria *et al.* (12) that focused on a variety of reinforcement-based decision-making paradigms also revealed reduced responses in the vmPFC, ACC, and insula among youths with CD and/or ODD [see also the review by Blair *et al.* (13)]. According to Fairchild *et al.* (14), the abovementioned regions (i.e., VS, vmPFC/orbitofrontal cortex [OFC], ACC, and insula) that showed altered response patterns

in youths with CD are important for outcome valuation. Therefore, they are central for RL and decision making, particularly for learning how to avoid punishment. However, a subsequent review and meta-analysis by Dugré *et al.* (15) that focused on conduct problems and antisocial behaviors more generally, including fMRI studies with youth as well as adults, concluded that the reviewed studies provide rather inconsistent results regarding the directionality of response patterns in core regions relevant to RL in affected individuals compared with healthy controls. More crucially, their meta-analytic results on the few studies of punishment and reward processing (i.e., not necessarily RL paradigms alone) did not show statistically significant response differences in RL-relevant regions of antisocial and typically developing individuals, which underlines the need for further systematic studies on reinforcement-based decision making (including RL from punishment and reward) to clarify the role of RL dysfunction in CD.

More recently, studies have started to utilize computational modeling to better understand RL mechanisms and the involved brain regions in typical and atypical populations. Computational modeling allows researchers to fit theory-informed RL models to behavioral learning data to explain the behavioral and related neural responses that underlie trial-by-trial learning from punishment and/or reward (16). A prevalent model used in the field is the Rescorla-Wagner model, which is an RL model based on PEs (17). Interestingly, Hauser *et al.* (18) used the Rescorla-Wagner model successfully in typical youths to investigate RL from reward and punishment, confirming the striatum, amygdala, vmPFC, and insula as important regions for RL.

Building on this previous work, in the current study, we aimed to further explore the differences in RL at the neural level in youths with a confirmed CD diagnosis compared with typically developing control participants (TDCs). Participants performed a probabilistic RL task with punishment, reward, and neural contingencies during fMRI. Specifically, we examined the VS, insula, ACC, amygdala, and vmPFC/OFC as regions of interest (ROIs) using model-based fMRI with PEs as parametric modulators derived from computational modeling. We expected that using PEs as parametric modulator would help to best capture the blood oxygen level–dependent signal to reveal group differences in the neural response patterns (19), particularly in relation to RL from punishment, thus reflecting punishment insensitivity/hyposensitivity as a hallmark of youth with CD (5,20). However, in view of the mixed findings in the relevant literature (15), we refrained from making any directional predictions. Because most of the previous research studied predominantly male samples, less is known about neural responses that underlie RL in girls with CD. Accordingly, we explored the influence of participants' sex on differential brain responses in participants with CD versus TDCs in the a priori ROIs.

METHODS AND MATERIALS

Participants

For this study, 248 (f)MRI-eligible participants between 9 and 18 years of age were recruited at Aachen, Germany and Southampton, United Kingdom through community outreach, youth offending services, mental health clinics, and welfare institutions. This was a subsample of the European multisite FemNAT-CD (Neurobiology and Treatment of Female Conduct Disorder) cohort (21) and was investigated in a previous study by Elster *et al.* (20). We excluded 11 participants with CD and 15 TDCs due to missing responses in the experimental task (i.e., more than 5 missing trials in each run and condition) and 17 participants with CD and 31 TDCs due to various artifacts in the imaging data (i.e., lack of quality in anatomical scans and/or head movement during the task). This left a final sample of 174 participants (Aachen: $n = 97$, Southampton: $n = 77$) including 75 participants with CD (29 girls) and 99 TDCs (53 girls). Overall exclusion criteria were autism spectrum disorder, psychosis or schizophrenia, mania or bipolar disorder, genetic syndromes, neurological disorders, and an IQ < 70. The study protocol was approved by the local ethics committees, and participants and their caregivers gave written informed consent. Participants were compensated for their participation, including the money that they gained during the RL task.

The 2 groups were specified as 1) CD (confirmed current diagnosis of CD) and 2) TDC (no current psychiatric diagnoses and no lifetime diagnoses of CD, ODD, or attention-deficit/hyperactivity disorder). Diagnoses were based on DSM-IV-TR criteria (22) assessed with the Kiddie-Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL) (23). The K-SADS-PL is a semistructured clinical interview and was administered separately to participants and their caregivers by trained staff members. Total IQ was estimated with the Wechsler Intelligence Scales (24). Callous/unemotional (CU) traits were assessed by using the total score of the subscales remorselessness, callousness, and unemotionality of the self-report version of the Youth Psychopathic Traits Inventory (25). Socioeconomic status (SES) was estimated based on parental income, education level, and occupation, normalized for the participant's country (26). There were no group differences in sex distribution, but age, estimated IQ, SES, and CU traits differed between the 2 groups, with the CD group being slightly older, having a lower IQ and lower SES, and higher CU traits (Table 1).

Probabilistic RL Task

We used a probabilistic RL task during fMRI, which was originally described by Kim *et al.* (27) and modified by Baumann *et al.* (28), to measure neural RL responses based on monetary punishment (PUN), monetary reward (REW), and neutral (NEUT; i.e., nonreinforcing) contingencies (Figure S1). Behavioral data collection and stimulus presentation were controlled by MATLAB (version R2014a; The MathWorks, Inc.) (29). The purpose of the task was for the participants to learn from feedback which cues were associated with a higher likelihood of reward or lower likelihood of punishment. By learning these contingencies, participants could maximize their monetary gains by choosing the more advantageous cue in each condition.

Prior to performing the task, participants were told that in each trial they would see one of the 3 different pairs of novel cue stimuli (fractals), i.e., they were unknown to the participants at the beginning of the task. Each pair of cues marked the onset of one of the 3 RL conditions (i.e., PUN, REW, and

Table 1. Demographic and Clinical Characteristics of the Sample

	CD, n = 75	TDC, n = 99	Group Comparison
Sex			
Female	29	53	$\chi^2_1 = 3.21$, $p = .07$
Male	46	46	
Age, Years	14.9 (2.4)	13.9 (2.6)	$t_{172} = 2.54$, $p < .05$
Estimated IQ	92.1 (10.8)	104.0 (10.8)	$t_{172} = -7.07$, $p < .001$
Socioeconomic Status	−0.25 (0.87)	0.31 (0.98)	$t_{172} = -3.87$, $p < .001$
CU Traits, YPI Subscale	33.1 (7.4)	28.9 (6.6)	$t_{172} = 3.98$, $p < .001$
CD Symptoms, K-SADS-PL	5.49 (2.28)	0.05 (0.22)	$t_{75} = 19.8$, $p < .001$
Comorbidities			
ODD	52 (69%)	N/A	–
ADHD	37 (49%)	N/A	–
MDD	24 (32%)	N/A	–
PTSD	11 (15%)	N/A	–
SUD	7 (9%)	N/A	–
GAD	5 (7%)	N/A	–
Psychotropic Medication			
Methylphenidate	11 (14.7%)	N/A	–
Antidepressants	5 (6.7%)	N/A	–
Antipsychotics	3 (4%)	N/A	–
Atomoxetine	2 (2.7%)	N/A	–
Lisdexamfetamine	1 (1.3%)	N/A	–

Values are presented as n, mean (SD), or n (%). All diagnoses are based on K-SADS-PL.

ADHD, attention-deficit/hyperactivity disorder; CD, conduct disorder; CU, callous/unemotional; GAD, generalized anxiety disorder; K-SADS-PL, Kiddie Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version; MDD, major depressive disorder; N/A, not applicable; ODD, oppositional defiant disorder; PTSD, posttraumatic stress disorder; SUD, substance use disorder; TDC, typically developing control participant; YPI, Youth Psychopathic Traits Inventory.

NEUT). The assignment of the cue pairs to the different contingency conditions was counterbalanced across participants. In each trial, participants were instructed to always choose by button press (i.e., right or left) one of the 2 simultaneously presented cues. Depending on their choices, participants then received probabilistic feedback in the form of losing money (PUN condition), winning money (REW condition), or obtaining a neutral outcome (NEUT condition). It was stressed that, based on the feedback, participants should learn to always select the cue associated with the better outcome (i.e., 70% [high] vs. 30% [low] probability of avoiding punishment in the PUN condition or receiving a reward in the REW condition) to win as much money as possible during the experiment. Each participant started the experiment with €10 in Germany or £10 in the United Kingdom and was told that any wins or losses would be added or subtracted, respectively, from this total. As per instructions, participants were paid according to their performance at the end of the experiment. The task lasted ∼25 minutes (including short breaks of ∼1 minute between runs) and had 45 trials per contingency across 3 runs (i.e., 15 trials × 3 contingencies × 3 runs = 135 trials in total). The order of trials was pseudorandomized to ensure that the same condition was never presented twice or more consecutively and that there were a similar number of trials in each condition. A more detailed description of the task and task procedures can be found in the Supplement and in Baumann *et al.* (28) and Elster *et al.* (20).

Computational RL Modeling and Statistical Analyses

We used the computational RL modeling parameters from our previous study (20) to assess each participant's individual learning behavior in conjunction with fMRI data applying the Rescorla-Wagner learning rule (17). The RL model that we used belongs to the PE-based models and is able to estimate individual learning rates α and exploration parameters β. α represents the speed at which individuals update their expectation of a certain outcome with newer information from trial to trial (higher α = quicker and more efficient value updating), and β represents an individual's randomness in choice behavior while learning (lower β = more consistent choices of the best option) (16). Model fitting, comparison, and identifiability procedures revealed that 3 separate learning rates α and 3 separate exploration parameters β, 1 for each of the 3 contingencies, most accurately and parsimoniously captured the learning behavior underlying the choices made by each participant over the course of the experiment (i.e., this model had best exceedance probability >.99, highest log model evidence = −16,681, and lowest integrated Bayesian information criterion = 33,164) (Figure S2). Information on parameter recovery can be found in Elster *et al.* (20).

First, we tested whether there were group differences regarding α, using a repeated-measures analysis of variance (ANOVA) with group (CD vs. TDC) and sex (male vs. female) as between-subjects factors and condition (REW vs. PUN vs. NEUT) as a within-subjects factor, followed by post hoc comparisons in case of significant main or interaction effects. The alpha level was set at .05. Effect sizes were calculated using η_p^2, where 0.01, 0.06, and 0.14 represent small, medium, and large effects, respectively (30). Analyses were conducted in R with RStudio (version 4.3.2) using the rstatix, emmeans, and BayesFactor packages. Additionally, we calculated Bayes factors (BFs) for post hoc comparisons using a BF that corresponds to how many times more likely the data are under the alternative hypothesis than under the null hypothesis (i.e., BF_{10}). A $BF_{10} > 3$ is considered substantial evidence in favor of the alternative hypothesis. A BF_{10} between 1/3 and 3 indicates that the data cannot clearly differentiate between the 2 hypotheses, and a $BF_{10} < 1/3$ favors the null hypothesis (31).

fMRI Acquisition

T2*-weighted images were obtained with echo-planar imaging using a 3T Siemens scanner (Aachen: Prisma, Southampton: Trio) and a 20-channel head coil. Whole-brain volumes of 41, 3-mm thick transversal slices (TR/TE = 2500/30 ms; flip angle = 83°; FOV = 192 × 192 mm^2; matrix size = 64 × 64, and voxel size = 3 × 3 × 3 mm^3) were collected throughout 3 functional runs. A total of 465 functional volumes (plus 5 dummy scans per run allowing for T1 magnetic saturation) were acquired for each participant. Prior to the functional runs, high-resolution

T1-weighted structural images of the entire brain were acquired using a magnetization prepared rapid acquisition gradient echo (MPRAGE) sequence (TR = 1900 ms, TE [Aachen/Southampton] = 3.4/4.1 ms; flip angle = 9°; FOV = 192 × 192 mm^2; matrix size = 256 × 256, and voxel size = 1 × 1 × 1 mm^3). Each site underwent a site qualification procedure before starting data collection to ensure comparability of MRI data acquisition and quality [see Raschle *et al.* (32) for details].

Whole-Brain and ROI Data Analysis

fMRI preprocessing and analysis were conducted using SPM12 (http://www.fil.ion.ucl.ac.uk/spm). Images were realigned to the first volume in the time series, anatomical scans were coregistered to the mean image and spatially normalized into a standard anatomical reference space, and spatial smoothing was applied using a Gaussian kernel with a full width at half maximum of 6 mm. To account for interscanner variability and minimize site-related effects, ComBat harmonization was applied to preprocessed data features (i.e., averaged beta weights from each ROI) prior to statistical analysis. ComBat, which has been shown to be effective in multisite imaging harmonization, adjusts for scanner differences while preserving the biological variability relevant to the study (33). Regression analysis was carried out on the preprocessed functional time series of each participant using a general linear model for an event-related design as implemented in SPM12. For the whole-brain analysis, simple contrasts with model-derived PEs (see Figure 1 and the Supplement for calculation) [see (16)] as parametric modulators at the time of outcome were created for the REW, PUN, and NEUT conditions. Cue onsets, choices, missing responses, and motion parameters were entered as regressors of no interest.

At the second level, simple contrasts were entered into a full-flexible ANOVA with group (CD vs. TDC) as the between-subjects factor and condition (REW vs. PUN vs. NEUT) as the within-subjects factor. We used a cluster extent threshold of $k = 10$ voxels in combination with a voxel-level threshold of $p < .001$, which is a common approach for whole-brain analyses in the neuroimaging literature (34). This combination is designed to balance sensitivity and specificity, ensuring robust detection of significant clusters while maintaining a familywise error–corrected significance level of $p < .05$, thereby effectively controlling for multiple comparisons across the whole brain (35).

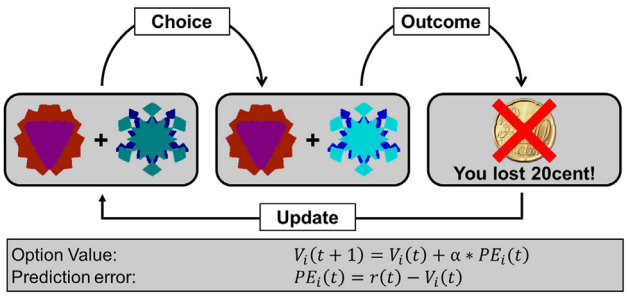

Figure 1. Punishment trial, option value, and prediction error (PE) calculation. An example trial of the reinforcement learning task for the punishment condition is shown here. At trial onset, a pair of abstract cue stimuli is presented. The participant then chooses one of them (the chosen cue is subsequently shown brighter) and receives a probabilistically determined outcome. Both of the cue stimuli have a specific option value *V*. For our task, we have V_i with $i \in$ [A, B] representing 2 possible options to choose from. The value for a chosen option $V_i = t + 1$ is updated with each new choice and depends on the expected value $V_i(t)$ that it had before and the $PE_i(t)$ combined with the learning rate α, where t represents the time point of choice. $PE_i(t)$ is derived from the comparison between the expected value $V_i(t)$ and the actual result of that choice $r(t)$.

Our a priori structurally defined ROIs comprised the right and left VS, (anterior) insula, ACC, vmPFC/OFC, and amygdala; we eventually used the bilateral masks because there were no lateralization effects. They were analyzed using the MarsBaR toolbox for SPM (36). MarsBaR extracts and averages time-series data within a predefined ROI and uses the statistical tests implemented in SPM12 to evaluate the significance of neural responses. By concentrating on specific brain regions, ROI analyses enhance the sensitivity to detect subtle effects. This is particularly useful when investigating small or specific brain areas that often do not survive correction for multiple comparisons at a whole-brain level. Responses within ROIs were assessed using the standard effects of interest *F* contrast in MarsBaR. This *F* contrast tests for significant response patterns across groups and conditions, as determined by our full-flexible ANOVA approach at the second level in SPM12. Anatomical masks for the ROI analyses were created in standard Montreal Neurological Institute (MNI) space using the FSL Harvard-Oxford atlas (37).

To better understand the influence of our model-derived learning rates α on brain responses and to specifically test our hypothesis of neural punishment insensitivity/hyposensitivity in CD, we then fitted linear regression models for each ROI that showed a significant effect in the ROI ANOVA models. While ANOVA models are commonly used for comparing group means, regression models allow for a more flexible approach when dealing with a mix of categorical and metric variables, especially when controlling for the effects of other covariates. We analyzed the influence of the learning rate α on parameter estimates derived from the SPM ROI analysis for each condition separately. The first model of each regression included the parameter estimates for the corresponding RL condition (i.e., REW, PUN, or NEUT) as the dependent variable and the corresponding α and group (i.e., CD, TDC) as independent variables. To assess whether including sex as an independent variable would improve model fit, we also tested another regression model:

$$\textit{Regression model } 1 : \textit{Parameter estimates}_i = \alpha_i * \textit{group} \quad (1)$$

$$\textit{Regression model } 2 : \textit{Parameter estimates}_i = \alpha_i * \textit{group} + \textit{sex} * \textit{group} \quad (2)$$

$$\textit{with } i \in [\textit{REW}, \textit{PUN}, \textit{NEUT}]$$

We compared the regression models to determine which model fit the data best using an ANOVA approach. We also tested whether specific covariates, such as age, IQ, SES, and site of data collection, would further improve the model fit of the regression models compared before. Below, we only report the winning regression models, but all results can be found in the Supplement.

RESULTS

Learning Rates α

The repeated-measures ANOVA for the learning rates α revealed a main effect of condition ($F_{2,340}$ = 34.67, $p < .001$, η_p^2 = 0.17) but also significant interactions of group × sex ($F_{1,170}$ = 4.65, $p < .05$, η_p^2 = 0.03), group × condition ($F_{2,340}$ = 5.30, $p < .01$, η_p^2 = 0.03), and sex × condition ($F_{2,340}$ = 3.15, $p < .05$, η_p^2 = 0.02). For the condition effect, the Bonferroni-adjusted post hoc comparisons revealed that overall participants had a higher learning rate for PUN than REW and NEUT ($mean_{Difference}$ [$\alpha_{REW} - \alpha_{PUN}$] = −0.15; 95% CI, −0.20 to −0.10; $p < .001$; $BF_{10} > 100$; $mean_{Difference}$ [$\alpha_{PUN} - \alpha_{NEUT}$] = 0.15; 95% CI, 0.10 to 0.20; $p < .001$; $BF_{10} > 100$), while the learning rates for reward and neutral did not differ significantly ($mean_{Difference}$ [$\alpha_{REW} - \alpha_{NEUT}$] < 0.01; 95% CI, −0.05 to 0.05; $p = .91$; BF_{10} = 0.11). The post hoc comparisons for all interaction effects were nonsignificant (*p*s > .15).

Whole-Brain Imaging Analyses

The whole-brain results are listed in the Supplement. We found only a main effect of condition (Table S2), indicating different brain response patterns in the insula, striatum, and ACC, as well as in other regions that were not relevant to our hypotheses (Tables S2–S5).

ROI Analyses

Subsequent ROI analyses were carried out on parameters estimated from the 5 ROIs to specifically test our hypothesis of neural punishment insensitivity/hyposensitivity in participants with CD compared with TDCs. Therefore, here we report only the effect that was consistent with our hypothesis. The full results of the ROI analyses, including the results for the reward condition, which showed no relevant group effects (which would indicate reward hyper- or hyposensitivity during reward learning), can be found in the Supplement.

The response pattern of the insula was exclusively consistent with the punishment insensitivity/hyposensitivity hypothesis of CD: The linear model that best predicted the parameter estimates in the punishment condition for the insula (which showed a significant response in the *F* contrast: $F = 2.87$, $p < .01$) included the variables group, learning rate (α_{PUN}), sex, and site (but did not include the covariates age, IQ, and SES):

$$PUN_{Insula} = group * \alpha_{PUN} + group * sex + site \quad (3)$$

This model explained a moderate proportion of the variance (R^2 = 14%, $F_{6,167}$ = 4.55, $p < .001$, $R^2_{adjusted}$ = 11%). As shown in Table 2, the individual predictors indicated that group, α_{PUN}, the interaction group × α_{PUN} (Figure 2), and the interaction group × sex were significant predictors, while sex was not. It is worth mentioning that this model showed a particularly large effect size (doubled R^2) when only the anterior insula was considered (Table S6).

DISCUSSION

In the current study, we aimed to test the punishment insensitivity/hyposensitivity hypothesis of CD by examining differences in neural responses during RL between youths with CD and TDCs through a combination of computational modeling of RL behavior (i.e., the learning rate α) and fMRI responses in core RL (i.e., PE-related) brain areas. Consistent with the punishment insensitivity/hyposensitivity hypothesis, we found that compared with TDCs, individuals with CD exhibited altered neural responses to RL from punishment, particularly in the (anterior) insula, while reward learning was unaffected. Importantly, although some of the ROI analyses also revealed interaction effects between sex and group, these effects were unrelated to RL behavior, suggesting that girls and boys with

Table 2. Regression Coefficients From the Linear Regression Model for Punishment Learning in the Insula

Explanatory Variables	Estimate	95% CI	*p* Value
Group	−3.82	−5.73 to −1.91	<.001
α_{PUN}	−5.20	−8.83 to −1.58	<.01
Sex	−1.14	−2.76 to 0.48	.17
Site	−1.45	−2.51 to −0.40	<.01
Group × α_{PUN}	7.02	2.27 to 11.78	<.01
Group × Sex	2.90	0.77 to 5.03	<.01

α_{PUN}, punishment learning rate.

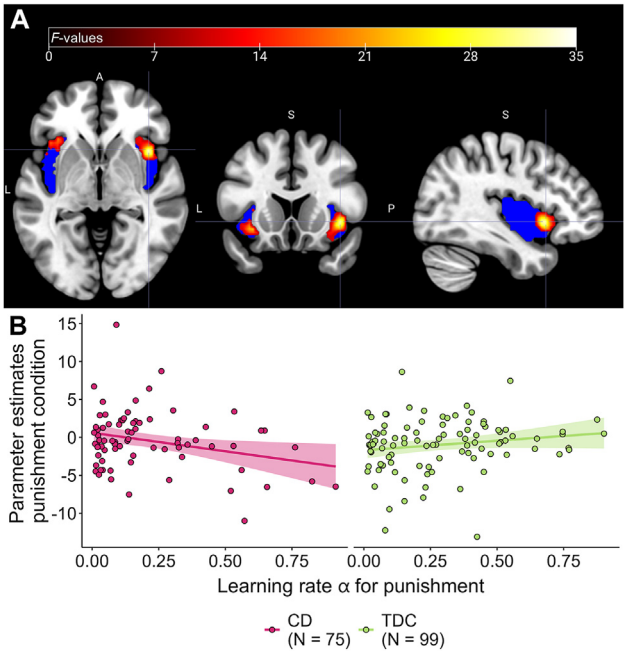

Figure 2. The group × learning rate (α) interaction effect for the insula response in the punishment condition. **(A)** The insula as region of interest defined by the Harvard-Oxford atlas in blue with the neural response (in hot colors) from the punishment > baseline contrast as indicated through the color bar with peak response at [x = 39, y = 15, z = −3] Montreal Neurological Institute (MNI) coordinates. Response is thresholded at $p < .001$ uncorrected for display purposes, and the color bar indicates *F* values. **(B)** The linear relationship between the response in the insula and the learning rate for punishment (α_{PUN}) is reversed in participants with conduct disorder (CD) vs. typically developing control participants (TDCs), suggesting that the response pattern of the insula to learning from punishment is different in the 2 groups. The shaded areas around the regression line indicate the standard error of the estimated response for the modeled linear relationship. A, anterior; L, left; P, posterior; S, superior.

CD show common alterations in the neural mechanisms of punishment learning.

This larger-scale fMRI study of probabilistic RL in CD considerably extends previous work by White *et al.* (9–11), who also examined RL-based neural response patterns in smaller samples of individuals with conduct problems (including individuals with CD and/or ODD). However, in contrast to our study, White *et al.* applied a preset, generic learning rate for all participants and RL conditions instead of individual learning rates for each condition. This approach did not take 2 factors into account: 1) that learning rates differ across individuals and 2) the so-called learning rate asymmetry, meaning that learning rates in probabilistic RL tasks are typically higher for punishment than reward (38); this was also observed here. Interestingly, though, we could confirm a different PE-related insula response in youths with CD compared with TDCs while learning to avoid punishment [similar to White *et al.* (9)]. More specifically, in the TDC group, we found greater (anterior) insula responses to be associated with higher learning rates for punishment. However, in the CD group this relationship was reversed; in individuals with greater (anterior) insula responses, punishment learning was worse. One could speculate that this effect may reflect different styles of punishment learning that characterize the 2 groups: when faced with punishment, typical youth learn the value of it more quickly and efficiently due to a stronger avoidance signal from the (anterior) insula (39) and then adjust their behavior accordingly. In contrast, youth with CD learn the value of punishment more slowly and less efficiently in the context of enhanced (anterior) insula responses, which may contribute to a less risk-averse but more harmful behavioral repertoire. Given the role of dopaminergic signaling in the (anterior) insula when processing PEs that drive RL, particularly from punishment (40), this insula finding may suggest, at least in part, altered dopaminergic signaling in youth with CD that potentially interferes with effective punishment learning and thus contributes to social norm violations (41). However, correlating heterogeneity factors (such as individual learning rates in CD vs. TDC groups) with task-related fMRI parameter estimates should only be done with caution and requires replication in follow-up studies to substantiate conclusions based on such analyses (42).

Although the factor of group (i.e., CD vs. TDC) was a significant predictor of differential ACC or amygdala responses in the punishment learning condition, we did not find a significant interaction effect with the behavioral learning rate for punishment. The response patterns of the vmPFC/OFC and VS/nucleus accumbens were also not consistent with our hypotheses. These results underscore the complexity (including heterogeneity) of the neural RL mechanisms that underlie CD, as well as the challenge of replicating previous study results using different methodologies [e.g., individual learning rates in the current study vs. generic, nonindividual learning rates as in (9)] and sample characteristics [e.g., confirmed CD diagnosis in the current study vs. a mixed sample of youths with CD and/or ODD in (9)]. Therefore, additional computational neuroimaging research that relates individual learning behavior to PE-based neural response patterns in youth with CD is warranted.

Our whole-brain analyses revealed a main effect of condition, indicating different response patterns in core RL brain areas and in the occipital pole across the 3 conditions in the entire sample. The activation in the occipital pole likely reflects visual processing of feedback stimuli, which differs between the 3 conditions. It is known that stimuli (such as rewards) engage visual regions in reward-related tasks because participants tend to process motivationally relevant outcomes more extensively [e.g., (43,44)]. However, we found no significant group differences in this analysis, suggesting that neural RL dysfunction in CD may be much more subtle (e.g., more heterogeneous) and not as widespread in this population as has been theorized (13). This idea fits well with the meta-analytic results by Dugré *et al.* (15), who were unable to reveal robust between-group response differences in core RL circuits when they compared individuals with antisocial behavior (including youths with CD) and TDCs who were performing punishment and/or reward-processing tasks.

Our study has several strengths, including 1) a clinically well-characterized and relatively large sample of female and male youths with a reliable diagnosis of CD as well as TDCs and 2) a rigorous methodological approach that combines fMRI data and computational modeling of RL behavior to test the punishment insensitivity/hyposensitivity hypothesis of CD. However, the use of more advanced computational RL models (45), modern statistical analysis approaches (46), and alternative experimental task designs (e.g., reversal RL tasks) in larger samples (with sufficiently powered sex distributions) will likely further deepen our understanding of aberrant punishment learning in CD beyond the current status quo, including possible sex-specific effects.

Conclusions

Our study provides new evidence that CD is associated with atypical neural responses to learning from punishment, but not from reward, specifically in the (anterior) insula. This suggests a selective dysfunction of (neural) RL mechanisms in CD thereby contributing to punishment insensitivity/hyposensitivity as a hypothesized hallmark of both girls and boys with the disorder (20). Because the (anterior) insula is critically involved in avoidance behaviors (i.e., avoidance of harm, loss, or risk in general) related to negative affective states or physiological arousal, it can be speculated that (anterior) insula dysfunction in CD leads to poor behavioral decision making, which increases the risk for reckless, antisocial, and aggressive behaviors in affected youth (47). However, to substantiate such a conclusion, additional studies are required that examine, for example, physiological markers (such as heart rate and/or electrodermal activity) in parallel with learning indices and brain responses. More importantly, because CD is a disorder in which interpersonal (i.e., social) functioning is impaired, investigating the neural mechanisms of social RL mechanisms will be crucial. This knowledge can then be incorporated into new intervention approaches that are tailored to the specific needs of individuals affected by CD [e.g., insula response regulation during RL from punishment via neurofeedback (48)].

ACKNOWLEDGMENTS AND DISCLOSURES

This work was supported by the European Commission's Seventh Framework Programme (Grant No. 602407 [FemNAT-CD; coordinator: CMF]). GK and EME were supported by a 2023 NARSAD Young Investigator Grant from

the Brain & Behavior Research Foundation (Grant No. 30849 [principal investigator, GK]). RP was supported by an ESRC postdoctoral fellowship award (Fellowship No. ES/V011324/1). PLL was supported by a Jacobs Foundation Research Fellowship, a Leverhulme Prize (Fellowship No. PLP-2021-196), a Wellcome Trust/Royal Society Sir Henry Dale Fellowship (Fellowship No. 223264/Z/21/Z), and a UKRI EPSRC Frontiers Research Guarantee/ERC Starting Grant (Grant No. EP/X020215/1). IAB was supported by a VIDI grant awarded by the Dutch Research Council (Grant No. VI.Vidi.2021G.017). SDB was supported by an ESRC grant (Grant No. ES/V003526/1). The project was funded by the Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung) as part of the German Center for Child and Adolescent Health (Grant No. 01GL2405B).

We thank the Federal Ministry of Education and Research and the state governments (http://www.nhr-verein.de/unsere-partner) for supporting this study as part of the joint funding of National High-Performance Computing. Furthermore, we thank Professor Stefan Ehrlich and the members of his group, especially Arne Doose, for their feedback on the model-based fMRI analysis. We also thank Dr. Josefine Rothe, who was partially funded by a Brain & Behavior Research Foundation grant (Grant No. 30849 [principal investigator, GK]).

CMF receives royalties for books on attention-deficit/hyperactivity disorder, autism spectrum disorder, and major depressive disorder. All other authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Department of Child and Adolescent Psychiatry, Faculty of Medicine, Dresden University of Technology, German Center for Child and Adolescent Health, partner site Leipzig/Dresden, Dresden, Germany (EME, MM, VR, GK); Centre for Human Brain Health, School of Psychology, University of Birmingham, Birmingham, United Kingdom (RP, SDB, PLL); Department of Psychology, University of Bath, Bath, United Kingdom (GF); Child Neuropsychology Section, Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital RWTH Aachen, Aachen, Germany (SB, KK); Department of Experimental-Clinical and Health Psychology, Ghent University, Ghent, Belgium (JS); Institute for Mental Health, School of Psychology, University of Birmingham, Birmingham, United Kingdom (SDB, PLL); School of Psychology, University of Birmingham, Birmingham, United Kingdom (SDB); Centre for Developmental Science, School of Psychology, University of Birmingham, Birmingham, United Kingdom (SDB); Birmingham Centre for Neurogenetics, University of Birmingham, Birmingham, United Kingdom (SDB); Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital Frankfurt, Goethe University, Frankfurt am Main, Germany (CMF); JARA-Brain Institute II, Molecular Neuroscience and Neuroimaging, RWTH Aachen and Research Centre Jülich, Jülich, Germany (KK); and Radboud University, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, the Netherlands (IAB).

Address correspondence to Erik M. Elster, M.Sc., at erik.elster@ukdd.de.

Received Oct 29, 2024; revised and accepted Jan 3, 2025.

Supplementary material cited in this article is available online at https://doi.org/10.1016/j.bpsc.2025.01.003.

REFERENCES

1. American Psychiatric Association (2013): Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5). Washington, DC: American Psychiatric Association.
2. Blair RJR (2013): The neurobiology of psychopathic traits in youths. Nat Rev Neurosci 14:786–799.
3. Pauli R, Lockwood PL (2023): The computational psychiatry of antisocial behaviour and psychopathy. Neurosci Biobehav Rev 145:104995.
4. Sutton RS, Barto AG (2018): Reinforcement Learning: An Introduction, 2nd ed. Cambridge, MA: The MIT Press. 526 p.
5. Byrd AL, Loeber R, Pardini DA (2014): Antisocial behavior, psychopathic features and abnormalities in reward and punishment processing in youth. Clin Child Fam Psychol Rev 17:125–156.
6. Fairchild G, van Goozen SHM, Stollery SJ, Aitken MRF, Savage J, Moore SC, Goodyer IM (2009): Decision making and executive function in male adolescents with early-onset or adolescence-onset conduct disorder and control subjects. Biol Psychiatry 66:162–168.
7. Kohls G, Baumann S, Gundlach M, Scharke W, Bernhard A, Martinelli A, *et al.* (2020): Investigating sex differences in emotion recognition, learning, and regulation among youths with conduct disorder. J Am Acad Child Adolesc Psychiatry 59:263–273.
8. Garrison J, Erdeniz B, Done J (2013): Prediction error in reinforcement learning: A meta-analysis of neuroimaging studies. Neurosci Biobehav Rev 37:1297–1310.
9. White SF, Pope K, Sinclair S, Fowler KA, Brislin SJ, Williams WC, *et al.* (2013): Disrupted expected value and prediction error signaling in youths with disruptive behavior disorders during a passive avoidance task. AJP. Am J Psychiatry 170:315–323.
10. White SF, Fowler KA, Sinclair S, Schechter JC, Majestic CM, Pine DS, Blair RJ (2014): Disrupted expected value signaling in youth with disruptive behavior disorders to environmental reinforcers. J Am Acad Child Adolesc Psychiatry 53:579–588.e9.
11. White SF, Tyler PM, Erway AK, Botkin ML, Kolli V, Meffert H, *et al.* (2016): Dysfunctional representation of expected value is associated with reinforcement-based decision-making deficits in adolescents with conduct problems. J Child Psychol Psychiatry 57:938–946.
12. Alegria AA, Radua J, Rubia K (2016): Meta-analysis of fMRI studies of disruptive behavior disorders. Am J Psychiatry 173:1119–1130.
13. Blair RJR, Veroude K, Buitelaar JK (2018): Neuro-cognitive system dysfunction and symptom sets: A review of fMRI studies in youth with conduct problems. Neurosci Biobehav Rev 91:69–90.
14. Fairchild G, Hawes DJ, Frick PJ, Copeland WE, Odgers CL, Franke B, *et al.* (2019): Conduct disorder. Nat Rev Dis Primers 5:43.
15. Dugré JR, Radua J, Carignan-Allard M, Dumais A, Rubia K, Potvin S (2020): Neurofunctional abnormalities in antisocial spectrum: A meta-analysis of fMRI studies on Five distinct neurocognitive research domains. Neurosci Biobehav Rev 119:168–183.
16. Lockwood PL, Klein-Flügge MC (2021): Computational modelling of social cognition and behaviour-A reinforcement learning primer. Soc Cogn Affect Neurosci 16:761–771.
17. Rescorla RA, Wagner AR (1972): A theory of Pavlovian conditioning: Variations on the effectiveness of reinforcement and non-reinforcement. In: Black AH, Prokasy WF, editors. Classical Conditioning II: Current Research and Theory. New York, NY: Appleton-Century-Crofts, 64–99.
18. Hauser TU, Iannaccone R, Walitza S, Brandeis D, Brem S (2015): Cognitive flexibility in adolescence: Neural and behavioral mechanisms of reward prediction error processing in adaptive decision making during development. Neuroimage 104:347–354.
19. Cohen MS (1997): Parametric analysis of fMRI data using linear systems methods. Neuroimage 6:93–103.
20. Elster EM, Pauli R, Baumann S, De Brito SA, Fairchild G, Freitag CM, *et al.* (2024): Impaired punishment learning in conduct disorder. J Am Acad Child Adolesc Psychiatry 63:454–463.
21. Freitag CM, Konrad K, Stadler C, De Brito SAD, Popma A, Herpertz SC, *et al.* (2018): Conduct disorder in adolescent females: Current state of research and study design of the FemNAT-CD consortium. Eur Child Adolesc Psychiatry 27:1077–1093.
22. American Psychiatric Association (2000): Diagnostic and Statistical Manual of Mental Disorders: DSM-IV-TR. Washington, DC: American Psychiatric Association.
23. Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, *et al.* (1997): Schedule for affective disorders and schizophrenia for school-age children-Present and lifetime version (K-SADS-PL): Initial reliability and validity data. J Am Acad Child Adolesc Psychiatry 36:980–988.
24. Wechsler D (1999): Wechsler Abbreviated Scale of Intelligence. San Antonio, TX: The Psychological Corporation.
25. Andershed HA, Kerr M, Stattin H, Levander S (2002): Psychopathic Traits in Non-referred Youths: A New Assessment Tool. Psychopaths: Current International Perspectives. Den Haag, Netherlands: Elsevier.
26. Pauli R, Tino P, Rogers JC, Baker R, Clanton R, Birch P, *et al.* (2021): Positive and negative parenting in conduct disorder with high versus low levels of callous-unemotional traits. Dev Psychopathol 33: 980–991.

27. Kim H, Shimojo S, O'Doherty JP (2006): Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. PLoS Biol 4:e233.
28. Baumann S, Hartz A, Scharke W, De Brito SA, Fairchild G, Herpertz-Dahlmann B, *et al.* (2022): Differentiating brain function of punishment versus reward processing in conduct disorder with and without attention deficit hyperactivity disorder. World J Biol Psychiatry 23:349–360.
29. MATLAB, version 8.3.0, vol. R2014a. Natick, MA: The MathWorks Incorporated (2014).
30. Cohen J (2013): Statistical power analysis for the Behavioral Science, Internet, 0th ed. Abingdon, UK: Routledge. Available at: https://www.taylorfrancis.com/books/9781134742707. Accessed July 2, 2024.
31. Marsman M, Wagenmakers EJ (2017): Bayesian benefits with JASP. Eur J Dev Psychol 14:545–555.
32. Raschle NM, Fehlbaum LV, Menks WM, Martinelli A, Prätzlich M, Bernhard A, *et al.* (2019): Atypical dorsolateral Prefrontal Activity in Female Adolescents with Conduct Disorder during Effortful Emotion Regulation. Biol Psychiatry Cogn Neurosci Neuroimaging 4:984–994.
33. Hu F, Chen AA, Horng H, Bashyam V, Davatzikos C, Alexander-Bloch A, *et al.* (2023): Image harmonization: A review of statistical and deep learning methods for removing batch effects and evaluation metrics for effective harmonization. Neuroimage 274:120125.
34. Yeung AWK (2024): A citation analysis of (f)MRI papers that cited Lieberman and Cunningham (2009) to justify their statistical threshold. PLoS One 19:e0309813.
35. Woo CW, Krishnan A, Wager TD (2014): Cluster-extent based thresholding in fMRI analyses: Pitfalls and recommendations. Neuroimage 91:412–419.
36. Brett M, Anton JL, Valabregue R, Poline JB (2002): Region of interest analysis using the MarsBar toolbox for SPM 99. Neuroimage 16:S497.
37. Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, *et al.* (2006): An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 31:968–980.
38. Gershman SJ (2015): Do learning rates adapt to the distribution of rewards? Psychon Bull Rev 22:1320–1327.
39. Kuhnen CM, Knutson B (2005): The neural basis of financial risk taking. Neuron 47:763–770.
40. Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD (2006): Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature 442:1042–1045.
41. Xiang T, Lohrenz T, Montague PR (2013): Computational substrates of norms and their violations during social exchange. J Neurosci 33:1099–1108a.
42. Lebreton M, Bavard S, Daunizeau J, Palminteri S (2019): Assessing inter-individual differences with task-related functional neuroimaging. Nat Hum Behav 3:897–905.
43. Anderson BA (2017): Reward processing in the value-driven attention network: Reward signals tracking cue identity and location. Soc Cogn Affect Neurosci 12:461–467.
44. Pessoa L, Engelmann JB (2010): Embedding reward signals into perception and cognition. Front Neurosci 4:17.
45. LaFollette K, Yuval J, Schurr R, Melnikoff D, Goldenberg A (2023): Data driven equation discovery reveals non-linear reinforcement learning in humans. Available at: https://osf.io/65jqh. Accessed June 4, 2024.
46. Haines N, Sullivan-Toole H, Olino T (2023): From classical methods to generative models: Tackling the unreliability of neuroscientific measures in mental health research. Biol Psychiatry Cogn Neurosci Neuroimaging 8:822–831.
47. Sonuga-Barke EJS, Cortese S, Fairchild G, Stringaris A (2016): Annual Research Review: Transdiagnostic neuroscience of child and adolescent mental disorders—Differentiating decision making in attention-deficit/hyperactivity disorder, conduct disorder, depression, and anxiety. J Child Psychol Psychiatry 57:321–349.
48. Lawrence EJ, Su L, Barker GJ, Medford N, Dalton J, Williams SCR, *et al.* (2014): Self-regulation of the anterior insula: Reinforcement learning using real-time fMRI neurofeedback. Neuroimage 88:113–124.