Conduct disorder is associated with heightened action initiation and reduced learning from punishment but not reward

Ruth Pauli[a], Inti Brazil[b], Gregor Kohls[c] Tobias U. Hauser[d,e,f,g], Lisa Gistelinck[h], Dimitris Dikeos[i], Roberta Dochnal[j], Graeme Fairchild[k], Aranzazu Fernández-Rivas[l], Beate Herpertz-Dahlmann[m], Amaia Hervas[n], Kerstin Konrad[o], Arne Popma[p], Christina Stadler[q], Christine M. Freitag[r], Stephane A. De Brito[a,s] , and Patricia L. Lockwood[a,s,t,u]

[a]Centre for Human Brain Health, School of Psychology, University of Birmingham, Birmingham, UK

[b]Radboud University, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands

[c]Department of Child and Adolescent Psychiatry, Faculty of Medicine, TU, Dresden, Germany

[d]Max Planck UCL Centre for Computational Psychiatry and Ageing Research, University College London, London, UK

[e]Wellcome Centre for Human Neuroimaging, University College London, London, UK

[f]Department of Psychiatry and Psychotherapy, Medical School and University Hospital, Eberhard Karls University of Tübingen, Tübingen, Germany

[g]German Center for Mental Health (DZPG), Tübingen, Germany

[h]Katholieke Universiteit Leuven, Leuven, Belgium

[i]Department of Psychiatry, Medical School, National and Kapodistrian University of Athens, Athens, Greece

[j]Faculty of Medicine, Child and Adolescent Psychiatry, Department of the Child Health Center, Szeged University, Szeged, Hungary

[k]Department of Psychology, University of Bath, Bath, United Kingdom

[l]Basurto University Hospital, Bilbao, Spain

[m]Child Neuropsychology Section, Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, RWTH Aachen University, Aachen, Germany

[n]University Hospital Mutua Terrassa, Barcelona, Spain

[o]JARA-Brain Institute, Aachen, Germany

[p]Department of Child and Adolescent Psychiatry, VU University Medical Center, Amsterdam, Netherlands

[q]Department of Child and Adolescent Psychiatry, Psychiatric University Hospital, University of Basel, Basel, Switzerland

[r]Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital Frankfurt, Goethe University, Frankfurt am Main, Germany

[s]Institute for Mental Health, School of Psychology, University of Birmingham, Birmingham, UK

[t]Department of Experimental Psychology, University of Oxford, Oxford, UK

[u]Wellcome Centre for Integrative Neuroimaging, University of Oxford, Oxford, UK

*Correspondence should be addressed to:

Ruth Pauli, Centre for Human Brain Health, University of Birmingham, United Kingdom. E-mail: r.pauli@bham.ac.uk

Patricia L. Lockwood, Centre for Human Brain Health, University of Birmingham, United Kingdom. E-mail: p.l.lockwood@bham.ac.uk

Abstract

Theoretical and empirical accounts of conduct disorder (CD) suggest problems with reinforcement learning as well as heightened impulsivity. These two facets can manifest in similar behaviour, such as risk-taking. Computational models that can dissociate learning from impulsively initiating actions are essential for understanding the cognitive mechanisms underlying CD. Here, a large, international sample of youths from 11 European countries (N = 1418, typically developing (TD) n = 742, CD n = 676) completed a passive avoidance learning task. We used computational modelling to disentangle reward and punishment learning from action initiation. Punishment learning rates were significantly reduced in youths with CD compared to their TD peers, suggesting that they did not update their actions based on punishment outcomes as strongly. Intriguingly, those with CD also had a greater tendency to initiate actions regardless of outcomes, although their ability to learn from reward was comparable to their TD peers. We additionally observed that variability in action initiation correlated with self-reported impulsivity in youths with CD. These findings provide empirical support for a reduced ability to learn from punishment in CD, while reward learning is typical. Our results also suggest that behaviours appearing superficially to reflect reward learning differences could reflect heightened impulsive action initiation instead. Such asymmetric learning from reward and punishment, with increased action initiation, could have important implications for tailoring learning-based interventions to help those with CD.

Introduction

Approximately 2% of children and adolescents in Europe[1] exhibit persistent antisocial, defiant, and aggressive behaviour severe enough to qualify for a diagnosis of conduct disorder (CD). The social and economic costs of CD are high[2,3]. Furthermore, CD is often associated with the core callous-unemotional affective features of psychopathy (e.g., lack of empathy and remorse)[4,5]. Many youths, especially those with the most marked callous-unemotional characteristics, will continue to exhibit antisocial behaviour in adulthood and eventually be diagnosed with antisocial personality disorder[6]. An improved understanding of CD, and the role of callous-unemotional traits in its presentation, is therefore a priority.

Youths with CD might struggle to learn from punishment and reward. Insensitivity to punishment cues was one of the first cognitive features of antisocial behaviour to be verified empirically[7] and is strongly associated with psychopathy in adults[4]. In children, CD has been associated with insensitivity to parental discipline, especially when the child has high levels of callous-unemotional traits[8–13]. Task-based measures of reinforcement learning also highlight difficulties in learning from punishment[14], although these difficulties are not always present[15]. More rarely, behavioural hypersensitivity to reward has been reported[16,17], although in 'risk-taking' rather than standard reinforcement learning contexts, where reward and punishment-related processes cannot easily be distinguished. Studies using various other learning contexts have not reported differences in reward learning[14,15,18] (but see[19]). However, evidence from parenting research highlights that children with conduct problems are responsive to parental praise and encouragement[12,20], suggesting that reward learning, if not heightened, is at least unlikely to be diminished in CD.

Although it is generally accepted that CD is associated with difficulties in learning from punishment and perhaps also with changes in reward sensitivity, the precise nature of these difficulties is still debated[21]. In particular, analytical methods that rely on summary statistics are usually not well suited to isolating learning processes from related but distinct phenomena, such as subjective valuation of outcomes[22] or certain forms of impulsivity[23]. In comparison, computational modelling can distinguish between different latent processes that are associated with similar overt behaviour[24]. For example, reinforcement learning models, which assume that actions and outcomes become associated through experience[25], can distinguish between learning rates (the extent to which recent outcomes influence future choices) and non-learning parameters such as action initiation biases[23,26]. These parameter values are influenced by the task demands as well as participant characteristics[27], but can nonetheless provide useful insights into underlying cognitive processes within specific learning contexts.

To date, very few studies have taken a computational approach to studying reward and punishment learning in CD. In an fMRI study of passive avoidance learning (i.e., learning to respond to rewarding objects and withhold responses to punishing objects), White and colleagues[28] demonstrated a weaker association between ventromedial prefrontal cortex activity and expected outcome value in youths with CD than in typically developing (TD) controls. However, learning rates were fixed at a set rate, which did not allow for modelling of group differences (see also[29]). A more recent study used a probabilistic reinforcement learning task[18]. In this study, youths with CD had significantly lower learning rates than TD controls for punishment stimuli. However, there was no difference in either direction when learning from reward, or for neutral stimuli, suggesting a specific difficulty with punishment learning. Thus, the small amount of computational modelling work to date supports an association with punishment learning difficulties, but no clear evidence for reduced (or heightened) learning from reward.

One possible reason for inconsistent findings regarding reward learning is that existing studies often have not distinguished reward learning from the tendency to initiate actions without regard to outcomes[15,18], although recent evidence suggests distinct developmental trajectories for these processes[23]. This means that differences in reward responding in some situations could reflect differences in the tendency to initiate actions rather than reward learning per se. Several previous studies have also relied on relatively small sample sizes (e.g.[28]).

In the current study, we sought to overcome these limitations by testing a large, well-characterised, mixed-sex sample from multiple countries, using an experimental paradigm and computational modelling procedure that enabled us to separate reward learning from action initiation. We used a passive avoidance learning task, a classic paradigm in psychopathy and CD research[28,30,31], from which much of the evidence for punishment learning difficulties or insensitivity has been derived[14,15,28,32]. Unlike probabilistic reinforcement learning tasks, in which participants must make a response on each trial, passive avoidance tasks involve learning to inhibit as well as to make responses[26]. This difference is particularly relevant to the question of reward hypersensitivity in CD, since impulsivity – frequently concomitant with CD[33] – can be separated from reward learning in a passive avoidance learning context[23]. Despite the centrality of passive avoidance learning in the literature, no studies to date have applied computational modelling in this learning context (except[28], which did not model learning rates for individual participants). Here, we used computational models of reinforcement learning to quantify differences in reward and punishment learning and impulsive action initiation biases. The large sample also enabled us to examine whether variability in computational parameters was related to self-reported impulsive and callous-unemotional traits, which are highly co-morbid with CD[5], or to CD severity.

We predicted that youths with CD would exhibit lower levels of learning from punishment (i.e., lower punishment learning rates) relative to TD peers. We did not make a directional hypothesis about reward learning rates due to the mixed findings in the literature[18]. Likewise, action initiation biases have not previously been studied in CD.

Methods and Materials

Participants

Participants were selected from the FemNAT-CD dataset[34]. This dataset consists of youths aged 9-18 years old who were either TD or met the criteria for CD. Full inclusion and exclusion criteria are detailed in the Supplementary materials, Inclusion and exclusion criteria. Participants were selected for the current analyses if they had data from a passive avoidance learning task. The total number of eligible participants was 1574. Of these, 156 were excluded due to problems with data quality (see Supplementary materials, Exclusions due to data quality). This left a final sample of 1418 participants (676 with CD (411 females), 742 TD (491 females)). As a control analysis, we repeated the main analyses with all 1574 participants, which did not change our conclusions (see Supplementary materials, Table S2). Demographic characteristics of the sample and subgroups are shown in Supplementary materials, Table S1.

Questionnaire measures

Participants were assessed for current and past psychiatric disorders, including CD, using the K-SADS-PL clinical interview[35]. Impulsive and callous-unemotional traits were assessed using the self-report Youth Psychopathic traits Inventory[36] (YPI). We used the YPI impulsiveness subscale as a measure of impulsivity and the callousness, unemotionality, and remorselessness subscales as the measure of callous-unemotional traits. IQ was assessed with the vocabulary and matrix reasoning subscales of the Wechsler Abbreviated Scale of Intelligence[37], the vocabulary, block design, and matrix reasoning subscales of the Wechsler Scale for Children, or the Wechsler Adult Intelligence Scale[38] according to site and age-specific protocols. Pubertal stage was assessed using the self-report Pubertal Developmental Scale (PDS[39]). Socioeconomic status (SES) was assessed based on parental income, education, and occupation. Full details of the questionnaire measures are available in Supplementary materials, Questionnaire measures. Imputation of missing data for all of the above measures is described in Supplementary materials, Imputation of missing data.

Passive Avoidance Learning task

Participants completed a passive avoidance learning task on a computer in a quiet testing room. The task was adapted from Blair and colleagues[31] using stimuli from Wong and colleagues[40] and presented in E-Prime[41]. The aim of the task was to gain points by pressing a button when presented with 'good' objects (to earn points) and withholding responses when presented with 'bad' objects (to avoid losing points). To maximise their point score, participants had to learn through trial-and-error which objects were associated with reward and which with punishment. There were eight different objects in total, four associated with rewards and four with punishment, with values of +/– 1, 700, 1400, or 2000 points. The point value associated with each object was fixed and did not change throughout the task. The eight objects were each presented 10 times in a random order (thus 80 trials in total). Each response was followed by feedback on the number of points gained or lost plus the running total; when participants did not respond, the value of the object was not revealed (see Figure 1). Stimuli were displayed for 3000ms or until the participant responded, and feedback (or the running total alone) was then displayed for 1000ms. Participants started the task with 10,000 points and could obtain final scores between 51,010 and –31,010.

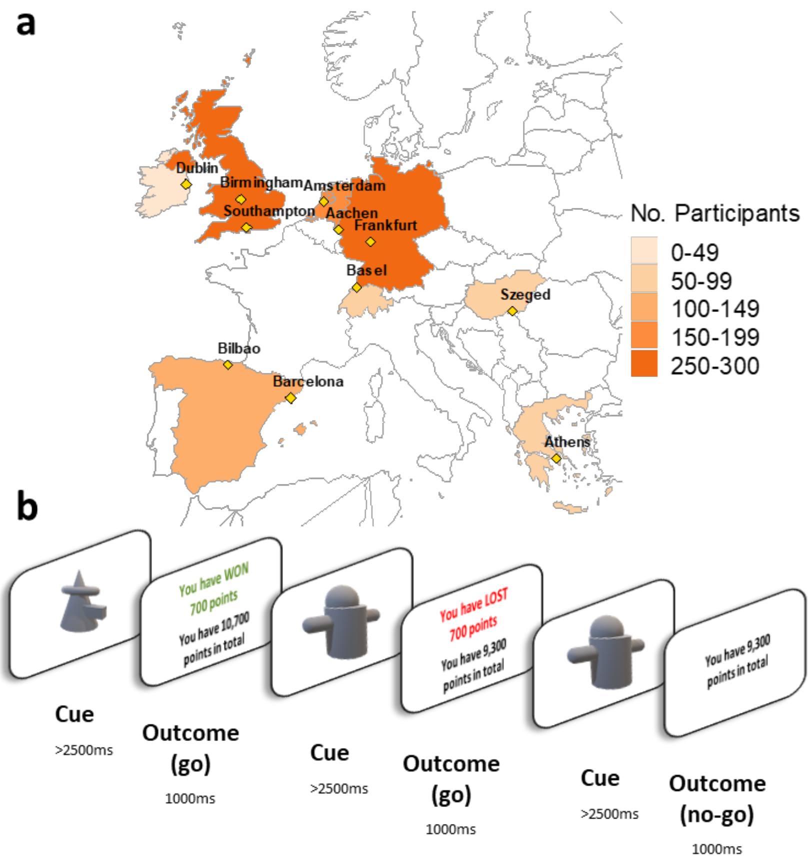

Figure 1. Recruitment sites and learning task. (a) Number of participants recruited from each country. Countries are coloured according to the total number of participants, with individual recruitment sites marked in yellow. **(b)** Details of the

learning task (shown here in the English language version). The aim of the task was to learn whether to respond or withhold responses to stimuli in order to earn points. Participants learnt by trial and error whether to make or withhold a button press to obtain a reward (points) or avoid punishment (losing points). Eight unfamiliar stimuli were presented individually for 3000ms or until a button press response was made. Responses were followed by feedback on the outcome (1000ms) or a running total alone if the participant did not respond. Each stimulus had a fixed value of +/– 1, 700, 1400, or 2000 points and was shown once per 'block' for 10 blocks, with a randomised order within blocks.

Statistical analysis of behavioural responses

All statistical analyses were conducted in R (v. 4.1.1 and v. 4.1.2) through RStudio[42]. Analysis of group differences in participants' behavioural responses was conducted using a nested linear mixed effects model. These analyses were conducted using R's lme4 package glmer function[43]. Participants' responses were coded as 1 (active response) or 0 (no response) and were predicted from group (TD = 0, CD = 1), age, sex (0 = male, 1 = female), IQ, SES, object repetition number (1-10), and object valence (0 = reward, 1 = punishment) as fixed effects, with varying intercepts allowed for responses grouped by participant nested within site as random effects. All continuous variables were z-scored, and discrete variables (participant response, group, sex) were recoded so that the two levels summed to zero (e.g., 0 and 1 becomes –0.5 and 0.5). To analyse whether group differences in responding across trials were driven by reward or punishment, post-hoc analyses were conducted by repeating the above analysis separately for reward and punishment stimuli. The strength of any null effects was interpreted using Bayes factors calculated with the BIC method[44] and R's lme4 BIC function[43], using standard priors and the language suggested by Jeffreys[45].

Model fitting and comparison procedure

Five different reinforcement learning models were constructed. For each model, rewards were coded as 1, neutral outcomes (when no response was made) as 0, and punishments as −1. Expected values for 'go' and 'no-go' actions were initialised at 0 at the task outset, which is the midpoint between the possible outcomes 1 and −1 and reflects participants' initial lack of knowledge about stimuli values. (An exception to this was for models with an initial action initiation bias; see below). First, we constructed a basic reinforcement learning model, in which learning was captured by a single learning rate (α) parameter and a temperature parameter β, which captures noisiness in responding. In this model, the expected value V of a response on trial t is updated with a reward prediction error PE scaled by the learning rate α, where the prediction error is the discrepancy between the outcome *r* (1, 0, or −1) and the expected value:

$$\text{If go: } V_{(t+1)} = V_{(t)} + (\alpha * PE_{(t)})$$

$$\text{If no-go: } V_{(t+1)} = V_{(t)}$$

where

$$PE_{(t)} = r_{(t)} - V_{(t)}$$

(1): basic model

The expected values are then converted to response probabilities using the Softmax equation, where the temperature parameter *β* adds noise:

$$\text{Probability of observed response} = e^{Vgo(t)/\beta} / (e^{Vgo(t)/\beta} + e^{Vnogo(t)/\beta})$$

(2): softmax

Using the model comparison procedure illustrated in Figure 2, we constructed four further models with combinations of additional parameters. These parameters were separate learning rates for reward versus punishment outcomes (3) and two versions of an action initiation bias towards responding regardless of anticipated outcome (4-5). We did not include outcome magnitude sensitivity parameters because in our previous work these parameters did not demonstrate good parameter recovery[23].

$$\text{For reward outcomes: } V_{(t+1)} = V_{(t)} + (\alpha_r * PE_{(t)})$$

$$\text{For punishment outcomes: } V_{(t+1)} = V_{(t)} + (\alpha_p * PE_{(t)})$$

(3): two learning rates

For models that included the initial 'go' bias, the starting value of responding to each object was increased (or decreased) by an amount *b_i* on the first presentation of the object only:

$$V_{(1)} = b\_i$$

(4): initial 'go' bias

For models that included the constant 'go' bias, the value of responding to each object was increased (or decreased) by an amount *b_c* on each presentation of the object:

$$V_{biased(t)} = V_{(t)} + b\_c$$

(5): constant 'go' bias

V_{biased} was used only to calculate the response probability for the current trial, so that the bias did not accumulate over repeated presentations of the object.

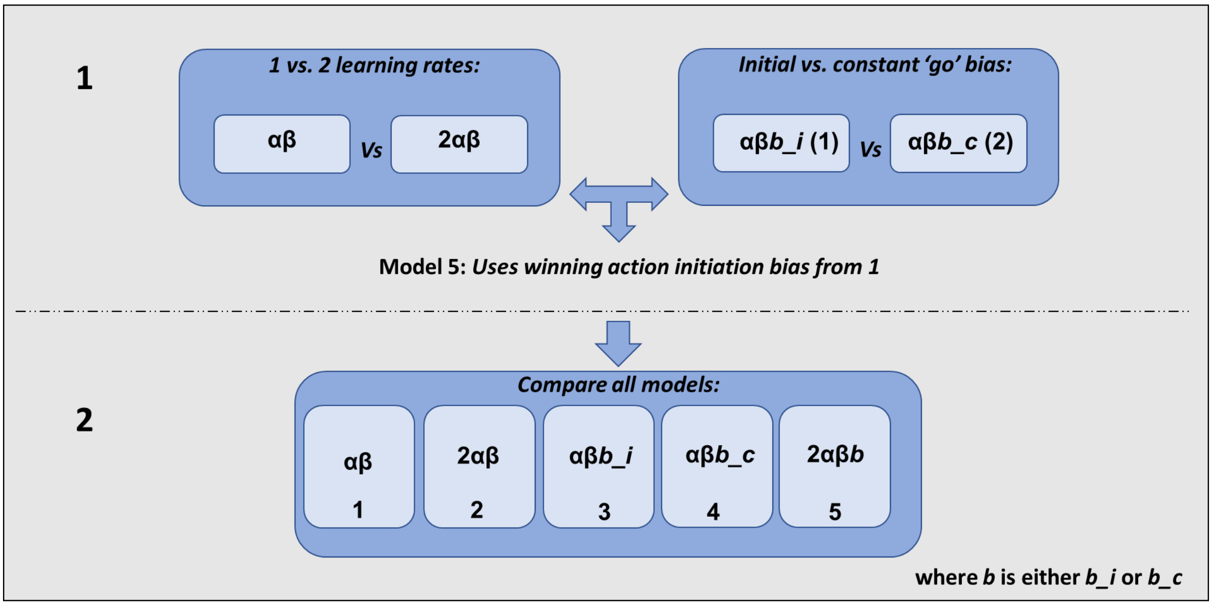

Figure 1. Steps in model construction procedure. In the first step (**1**), models with one versus two learning rates (for reward and punishment) were compared, and separately, models with an initial versus constant action initiation bias were compared. A fifth model was then constructed by combining all parameters from the winning models in step **1** (i.e., one versus two learning rates and the winning action initiation bias). In step **2**, to confirm that the winning model from step **1** was the best overall model, we compared models 1-5 directly.

Model fitting and comparison were conducted in MATLAB 2019b[46]. We used an iterative maximum a posteriori (MAP) approach for all model fitting, in line with previous work using reinforcement learning models[47–50]. Models were compared on exceedance probability (where a probability over .95 provides strong evidence for the best fitting model), log model evidence (LME; more positive values indicating better fit[51]) and the integrated BIC score (BIC_{int}), which penalises more complex models so that a lower BIC_{int} indicates better performance. Full details of the model fitting and comparison procedure are provided in Supplementary materials, Model fitting and comparison procedure.

Statistical analysis of model parameters

Next, we investigated associations between group status (CD versus TD) and model parameters from the winning model. Since parameter values were not normally distributed, we used robust linear mixed effects regression models using the rlmer[43] function in R. We tested whether each parameter was predicted by group status, with age, IQ, SES, and sex as covariates (fixed effects) and varying intercepts for different sites of data collection (random effects). Discrete variables were recoded so that contrasts summed to zero, and continuous variables were z-scored. As above, the strength of null effects was interpreted using Bayes factors calculated with the BIC method[44] using the language suggested by Jeffreys[45].

Associations between model parameters, callous-unemotional and impulsive traits, and CD severity

Finally, we investigated whether relevant model parameters were associated with self-reported callous-unemotional and impulsive traits in youths with CD, because these traits were theoretically linked to the model parameters that differed between groups (see Results). For punishment learning rates, we tested for associations with the YPI callousness, unemotionality, and remorselessness subscales using three separate robust linear mixed effects regression models. Parameter values were predicted from the relevant YPI subscale (callousness, unemotionality, or remorselessness) and age, sex, and IQ as covariates. Site of data collection was included as a random effect. We corrected for multiple comparisons across the three models using Bonferroni correction. Next, we repeated this procedure to test for an association between action initiation biases and impulsivity using a fourth regression model with the same covariates as above. Where significant associations emerged, we conducted a simple Pearson correlation to check that this was also significant.

Results

Demographic and clinical characteristics

Demographic and clinical characteristics of the sample are shown in Table 1 and Supplementary materials, Table S1. Youths with CD and TD controls differed on IQ and SES, and there was also a group difference in the proportion of females. (Note that the over-representation of females in both groups represents a deliberate sampling strategy of the FemNAT-CD project[34]). We repeated our main analyses with groups that were matched for age, pubertal stage, number of males and females, IQ, and estimated SES. All results remained virtually the same (see Supplementary materials, Group matching procedure and Tables S1-S2).

Table 1. Demographic and clinical characteristics

	All participants	**Conduct disorder**	**Typically developing**	
N	1418	676	742	
Num. females (%)	902 (64%)	411 (61%)	491 (66%)	$X^2_{(1)} = 4.18$, $p = .04$
Age (years)	14.12 [2.40]	14.26 [2.31]	13.99 [2.48]	$t_{(1415.20)} = 2.15$, $p = .03$
Pubertal stage (median)	4: late pubertal	4: late pubertal	4: late pubertal	$t_{(1404.90)} = 0.99$, $p = .32$

IQ	99.28 [12.97]	94.87 [12.35]	103.43 [12.16]	$t_{(1350.80)} = -12.95$, $p < .001$
SES	0.02 [1.01]	−0.36 [0.97]	0.35 [0.92]	$t_{(1257.40)} = -13.63$, $p < .001$
Num. comorbid ADHD	N/A	290 (43%)	N/A	

Notes: ADHD = attention deficit/hyperactivity disorder, SES = socioeconomic status. Columns 2-4 show means [standard deviation] unless otherwise specified. Column 5 shows statistical tests for differences between the two groups (columns 3-4).

Youths with CD made fewer correct responses for punishment but more for reward

First, we compared overall performance in youths with CD versus TD controls (Figure 3). The two groups made similar numbers of correct responses overall (main effect of group: odds ratio (OR) = 1.02 [0.97, 1.07], $z = 0.61$, $p = .54$, $BF_{01} = 266.00$) but there was a significant group by valence interaction (OR = 0.55 [0.53, 0.59], $z = -20.22$, $p < .001$). Separate follow-up models for reward and punishment stimuli revealed that youths with CD made more correct responses than TD controls for reward stimuli (OR = 1.28 [1.15, 1.42], $z = 4.66$, $p < .001$), but fewer correct responses for punishment stimuli (OR = 0.78 [0.71, 0.88], $z = -4.40$, $p < .001$).

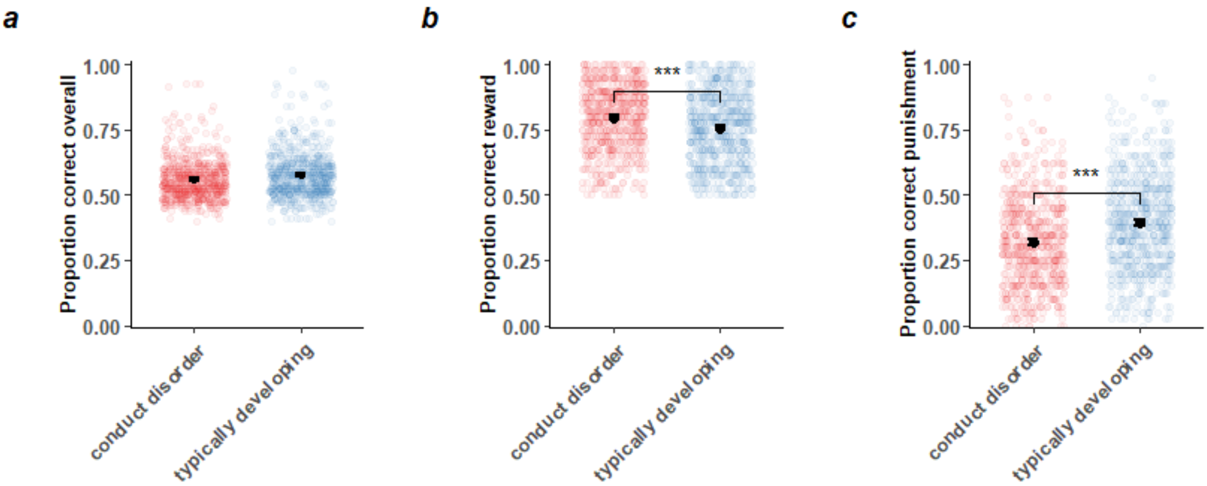

Figure 3. Youths with CD made fewer correct responses for punishment but more for reward. **(a)** Group differences in overall performance. Youths with CD and TD controls exhibited similar performance (OR = 1.02 [0.97, 1.07], $z = 0.61$, $p = .54$, $BF_{10} = 0.004$, $BF_{01} = 274.35$). **(b)** Group differences in performance for reward stimuli. Youths with CD significantly outperformed TD controls (OR = 1.28 [1.15, 1.42], $z = 4.66$, $p < .001$). **(c)** Group differences in performance for punishment stimuli. TD controls significantly outperformed youths with CD (OR = 0.78 [0.71, 0.88], $z = -4.40$, $p < .001$). Coloured circles represent individual participants and black circles represent group means with 95% confidence intervals. *** $p < .001$.

Behavioural evidence for learning from reward and punishment in CD and TD

There was a significant main effect of stimulus repetition on the number of correct responses made, with performance improving throughout the task (i.e., behavioural evidence for learning; OR = 1.19 [1.16, 1.20], $z = 22.81$, $p < .001$). The level of overall learning was similar for the two groups (group by stimulus by repetition interaction: OR = 0.99 [0.96, 1.02], $z = -0.67$, $p = .50$, $BF_{01} = 256.00$). However, a significant group by valence by repetition interaction suggested that the two groups learned differently from reward versus punishment (OR = 0.78 [0.74, 0.83], $z = -8.56$, $p < .001$, Figure 4). Specifically, separate follow-up models for reward and punishment stimuli revealed that, on a behavioural level, youths with CD seemed to learn more from reward than TD controls (OR = 1.13 [1.08, 1.19], $z = 5.45$, $p < .001$) while for punishment TD controls learned more than youths with CD (OR = 0.86 [0.83, 0.90], $z = -7.37$, $p < .001$).

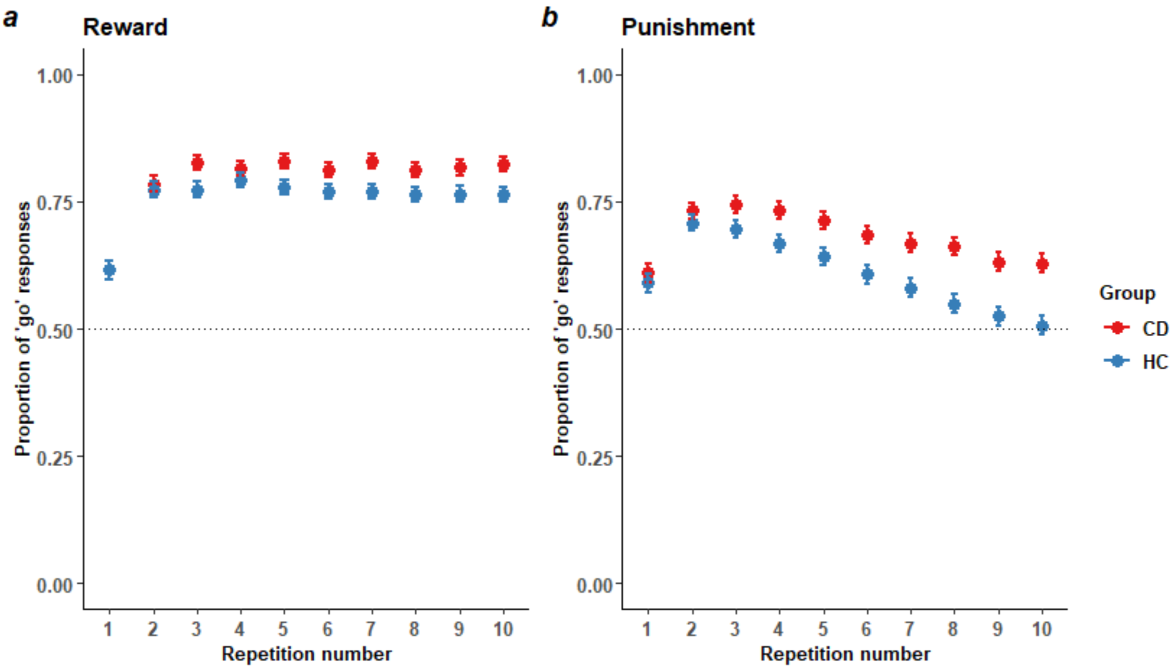

Figure 4. Reward and punishment responding across stimulus repetitions for youths with CD and TD **controls. (a)** Proportion of 'go' responses to reward stimuli across repeated stimulus presentations. Youths with CD learned better from reward than TD controls (OR = 1.13 [1.08, 1.19], $z = 5.45$, $p < .001$). **(b)** Proportion of 'go' responses to punishment stimuli across repeated stimulus presentations. TD controls learned better from punishment than youths with CD (OR = 0.86 [0.83, 0.90], $z = -7.37$, $p < .001$). Error bars represent 95% confidence intervals.

Finally, we noted that there was a sharp increase in responses on repetition 2 followed by more gradual changes thereafter. To ensure that our results were not driven by this initial pattern alone, we replicated our analyses with stimulus presentations 2-10 only. This did not change our findings (see Supplementary materials, Table S2).

Computational modelling of reward and punishment learning and action initiation

We next compared the five computational reinforcement learning models. Models were compared on exceedance probability, Log Model Evidence (LME), and the integrated Bayesian Information Criterion (BIC_{int}). Model 5 (2αβ*b*_c) won on all three measures (see Figure 5). Correlations between parameter values and task performance are shown in Supplementary materials, Tables S3-S4.

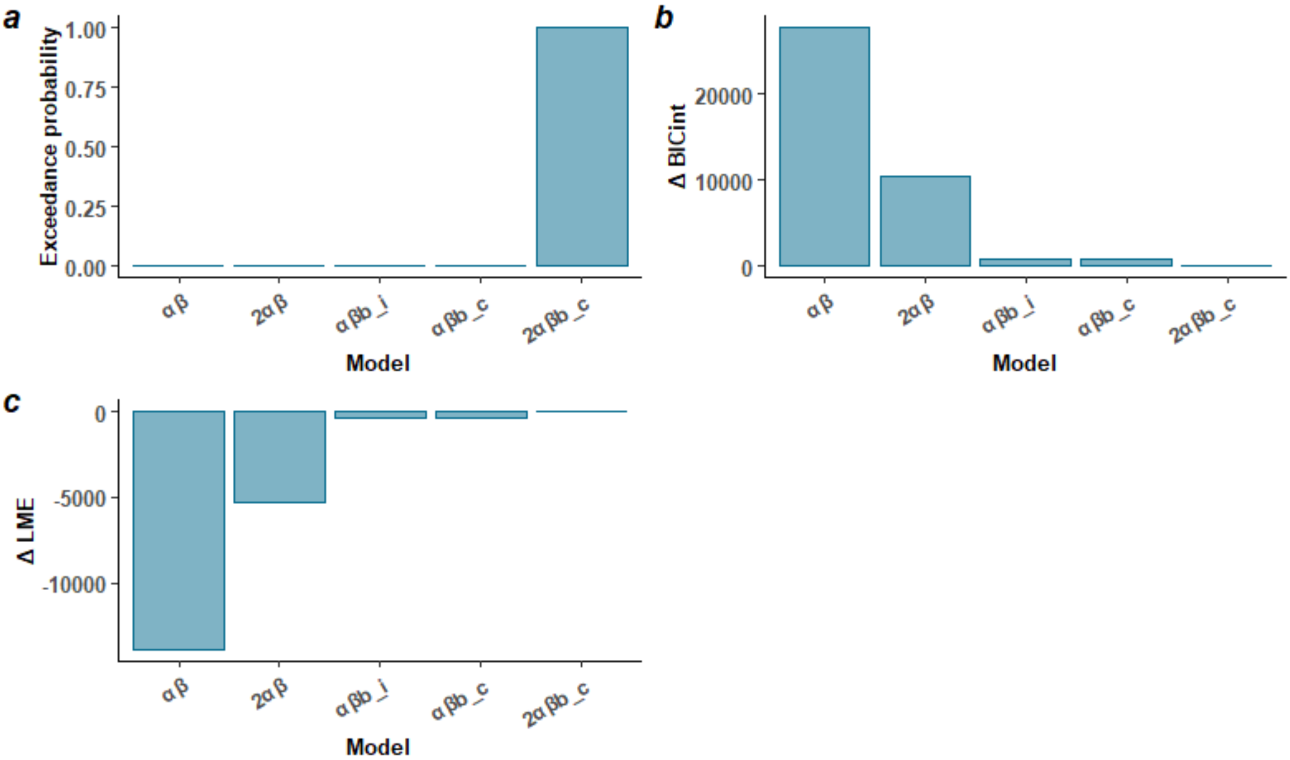

Figure 5. Model performance and validation. (a) Exceedance probability for the five computational models. The winning model was the 2αβ*b_c* model, with separate reward and punishment learning rates and a constant action initiation bias. **(b)** ΔBIC_{int}, relative to the winning model (2αβ*b_c*). **(c)** ΔLME, relative to the winning model (2αβ*b_c*).

Youths with CD had lower punishment learning rates

We next assessed whether parameter values from the winning model varied in youths with CD versus TD controls. Youths with CD exhibited significantly lower punishment learning rates than TD youths ($\beta = -0.10$ [-0.16, -0.03], $z = -2.81$, $p = .01$) but, despite making more responses to reward stimuli, did not differ from TD controls on reward learning rates, with Bayesian evidence strongly supporting the null effect (β = 0.09

[–0.01, 0.19], $z = 1.67$, $p = .09$, $BF_{10} = 0.002$, $BF_{01} = 433.72$). Additionally, youths with CD exhibited significantly higher action initiation biases than TD controls ($\beta = 0.35$ [0.22, 0.47], $z = 5.31$, $p < .001$). There was also a significant difference in temperature parameter between the groups ($\beta = -0.14$ [–0.02, 0.06], $z = -2.43$, $p = .02$), but we did not attempt to interpret this difference because it was less robust across control analyses (Supplementary materials, Table S2) and did not relate to our hypotheses.

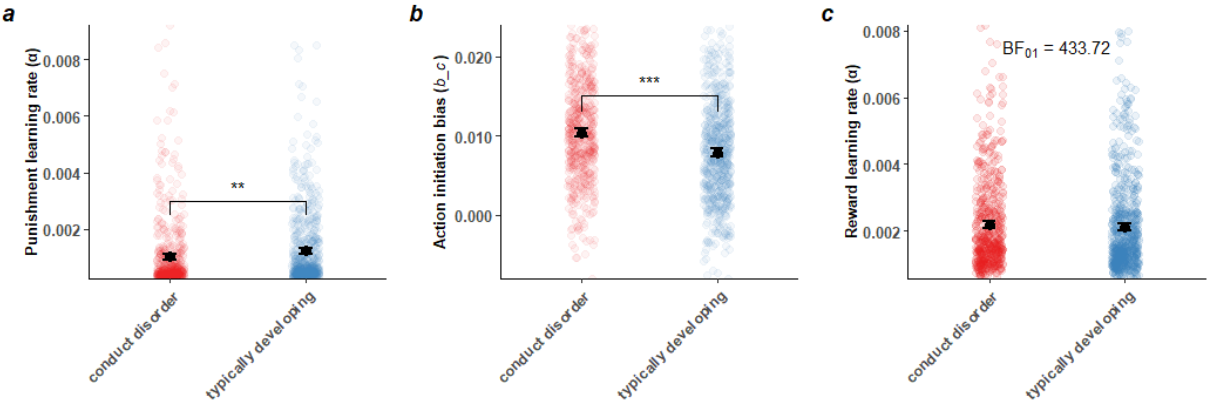

Figure 6. Youths with CD had lower punishment learning rates but similar reward learning rates compared to TD youths. (a) Group differences in punishment learning rates. Youths with CD exhibited significantly lower punishment learning rates ($\beta = -0.10$ [–0.16, –0.03], $z = -2.81$, $p = .01$). **(b)** Group differences in action initiation biases. Youths with CD exhibited significantly higher action initiation biases ($\beta = 0.35$ [0.22, 0.47], $z = 5.31$, $p < .001$). **(c)** Group differences in reward learning rates. Youths with CD did not differ from TD youths on reward learning rates ($\beta = 0.09$ [–0.01, 0.19], $z = 1.67$, $p = .09$, $BF_{10} = 0.002$, $BF_{01} = 433.72$). Coloured circles represent individual participants and black circles represent group means with 95% confidence intervals. *** $p < .001$.

Associations between model parameters, impulsive and callous-unemotional traits, and CD severity

First, we tested for associations between punishment learning rates and the three callous-unemotional trait subscales of the YPI. There were no significant associations with callousness ($\beta = 0.004$ [–0.01, 0.02], $z = 0.62$, $p = .54$, $BF_{10} = 0.04$, $BF_{01} = 23.47$), unemotionality ($\beta = -0.002$ [–0.01, 0.01], $z = -0.42$, $p = .68$, $BF_{10} = 1.00$, $BF_{01} = 1.00$), or remorselessness ($\beta = -0.001$ [–0.01, 0.01], $z = -0.21$, $p = .83$, $BF_{10} = 0.04$, $BF_{01} = 23.19$). Next, we tested for an association between action initiation bias and impulsivity scores. Higher impulsivity scores predicted significantly higher action initiation biases ($\beta = 0.04$ [0.02, 0.07], $z = 3.54$, $p < .001$, Figure 7). A simple Pearson correlation between impulsivity and action initiation biases was also significant ($r_{(550)} = .10$, $p = .02$). There was no association between the number of CD symptoms and punishment learning rates ($\beta = 0.01$ [–0.01, 0.02], $z = 1.03$, $p = .30$, $BF_{10} = 0.06$, $BF_{01} = 16.09$) or action initiation biases (β = 0.03 [−0.01, 0.01], z = 1.64, p = .10, BF_{10} = 0.20, BF_{01} = 4.94).

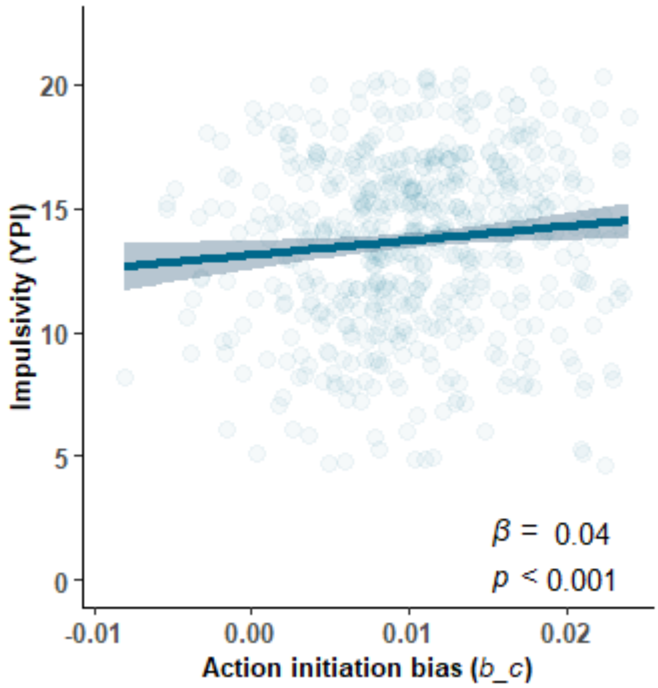

Figure 7. Association between action initiation bias and impulsive traits in youths with CD. There was a significant positive association between the action initiation bias and YPI impulsivity scores (β = 0.04 [0.02, 0.07], z = 3.54, p < .001).

Discussion

CD and related disorders have often been associated with poor reinforcement learning from punishment[7,14,15,18] and, less consistently, with hypersensitivity to reward[16,17]. However, none of these studies could distinguish reward learning from impulsively initiating actions. Here, we show using computational modelling that punishment learning rates are reduced in youths with CD. Moreover, while reward learning rates are comparable to TD youths, action initiation biases are heightened in youths with CD. This tendency to impulsively initiate actions was also positively associated with self-reported trait impulsivity. Together, these findings suggest that an apparent hypersensitivity to reward might sometimes reflect an impulsive action initiation bias, rather than enhanced reward learning.

The finding that punishment has a reduced deterrent effect in CD has been reported previously in smaller samples[8–15,18]. Our study adds important evidence in a large international cohort that this is a true learning effect, in line with a recent study that used a probabilistic reinforcement learning task[18]. Interestingly, however, we did not observe an association between punishment learning rates and callous-unemotional traits. On one level, this is surprising, given that callous-unemotional traits are core features of psychopathy and punishment insensitivity has often been considered central to psychopathic traits[7,52]. However, empirical evidence has not always supported this association[53]. Our study provides further evidence that the severity of callous-unemotional traits might not be intrinsically linked to heightened punishment insensitivity, at least within the range of callous-unemotional traits seen in this sample.

Previous investigations of reward learning in CD have yielded mixed findings[14–18] with a minority reporting reward hypersensitivity[16,17]. These were studies that used a 'risk-taking' task in which reward sensitivity was captured by the tendency to pursue rewards despite negative outcomes. In this context, it is unclear whether behaviour is driven by reward sensitivity or punishment insensitivity. Our findings suggest that previous reports of reward hypersensitivity might reflect decreased punishment learning, combined with an impulsive action initiation bias. This interpretation is consistent with previous studies showing no difference in reinforcement learning from reward[14,15,18], and helps to explain why the literature on reward learning appears inconsistent. Additional research into learned versus impulsive responding will be valuable for further elucidating reward-related behaviour in CD.

Our findings have implications for educational interventions. In particular, while the deterrent effect of punishment is likely to be reduced in CD, the motivating effect of reward is unlikely to be enhanced. While rewards (including praise) are an important element of behaviour management in most schools, evidence for their effectiveness in the general pupil population is quite limited[54]. The current findings do not suggest that rewards will be more effective for youths with CD than for other youths. Thus, it might be more effective for teachers to focus on reducing impulsivity, for example by helping youths with CD to pre-plan responses to aggravating situations, rather than rely on consequence-based strategies.

This study has several strengths, including the large, international, clinically well-characterised sample and the use of a classic paradigm in CD research[30,31]. However, we note some limitations. First, due to high levels of comorbidity, we were not able to control for the presence of ADHD. This level of comorbidity does, nonetheless, reflect clinical reality[21]. While this is therefore a shortcoming for theoretical understanding, the practical implications of our findings would likely stand regardless of the role of ADHD. Second, our task design did not allow us to distinguish between inhibiting actions to avoid punishment and inhibiting actions to gain reward, nor between making responses to gain reward and making responses to avoid punishment[26]. This would have allowed a more nuanced understanding of how action biases interact with outcome valence. Finally, it should be noted that modelling parameters are always a reflection of the task as well as participant characteristics[27,55], and modelling parameters should be subject to the same scrutiny and independent replication as any other psychological measurement.

In summary, we demonstrated that youths with CD exhibited lower punishment learning rates, higher action initiation biases, and similar reward learning rates to TD youths. This is one of the first studies to apply computational modelling techniques to studying CD. Our findings suggest that impulsive behaviour is an important element of apparently reward-oriented behaviour in CD, and that practitioners should be aware of this alongside more widely recognised difficulties in punishment learning.

Acknowledgments

R.P was supported by an ESRC post-doctoral fellowship award. This work was supported by a Medical Research Council Fellowship (MR/P014097/2), a Jacobs Foundation Research Fellowship, a Leverhulme Prize (PLP-2021-196) a Wellcome Trust/Royal society Sir Henry Dale Fellowship (223264/Z/21/Z) and a UKRI EPSRC Frontiers Research Guarantee/ERC Starting Grant (EP/X020215/1) to P. L. Lockwood. G.K was supported by a 2023 NARSAD Young Investigator Grant from the Brain & Behavior Research Foundation (BBRF; Grant No. 30849). The FemNat-CD consortium was funded by the European Commission under the 7th Framework Health Program, Grant Agreement no. 602407. We are grateful to all our participants and their families, and to Jo Cutler, Anthony Gabay, Miriam Klein-Flugge, Marco Wittmann, and Stefano Palminteri for helpful discussions and advice.

Author contributions

R.P previously collected data for this study as part of the FemNat-CD consortium. G.K adapted the learning task for use in this study. G.K and I.B collated the learning task data and conducted preliminary data pre-processing and quality checks. R.P conducted all analyses and wrote the manuscript. P.L assisted with coding for the computational modelling analyses, contributed to manuscript preparation, and provided guidance and oversight on all aspects of the analyses. All other authors contributed substantially to study design, acquisition of funding, project supervision and management, and/or data collection as part of the FemNAT-CD consortium. All authors read and approved the final manuscript.

Declaration of interests

C.M.F receives royalties for books on attention-deficit/hyperactivity disorder and autism spectrum disorder. She has served as consultant to Desitin and Roche. No other authors report any conflicts of interest.

Ethics Declarations

The FemNAT-CD project received ethical approval from the relevant local ethics committees, as follows: Aachen: Ethik Kommission Medizinische Fakultät der Rheinisch Westfälischen Technischen Hochschule Aachen (EK027/14). Amsterdam: Medisch Etische Toetsingscommissie (2014.188). Athens: Election Committee of the First Department of Psychiatry, Eginition University Hospital (641/9.11.2015). Barcelona: Child and Adolescent Mental Health—University Hospital Mutua Terrassa (acta 12/13). Basel: Ethik Kommission Nordwest- und Zentralschweiz (EKNZ 336/13). Bilbao: Hospital del Basurto. Birmingham and Southampton: University Ethics Committee and National Health Service Research Ethics Committee (NRES Committee West Midlands, Edgbaston; REC reference 13/WM/0483). Dublin: SJH/AMNCH Research Ethics Committee (2014/04/Chairman (3)). Frankfurt: Ethik Kommission Medizinische Fakultät Goethe Universität Frankfurt am Main (445/13). Szeged (Hungary): Egészségügyi Tudományos Tanács Humán Reprodukciós Bizottság (CSR/039/00392–3/2014). This study was conducted in accordance with the ethical standards of the 1964 Declaration of Helsinki and its later amendments.

References

1. Sacco, R., Camilleri, N. & Umla-Runge, K. The prevalence of conduct disorders among young people in europe: A systematic review and meta-analysis. *Eur. Psychiatry* **64**, S637–S638 (2021).
2. Romeo, R., Knapp, M. & Scott, S. Economic cost of severe antisocial behaviour in children - and who pays it. *Br. J. Psychiatry* **188**, 547–553 (2006).
3. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet Psychiatry* **9**, 137–150 (2022).
4. Cleckley, H. *The Mask of Sanity*. (Mosby, Saint Louis, 1976).
5. Rowe, R. *et al.* The role of callous and unemotional traits in the diagnosis of conduct disorder. *J. Child Psychol. Psychiatry* **51**, 688–695 (2010).
6. Gelhorn, H. L., Sakai, J. T., Price, R. K. & Crowley, T. J. DSM-IV conduct disorder criteria as predictors of antisocial personality disorder. *Compr. Psychiatry* **48**, 529–538 (2007).
7. Lykken, D. T. A study of anxiety in the sociopathic personality. *J. Abnorm. Soc. Psychol.* **55**, 6–10 (1957).
8. Wootton, J. M., Frick, P. J., Shelton, K. K. & Silverthorn, P. Ineffective parenting and childhood conduct problems: The moderating role of callous-unemotional traits. *J. Consult. Clin. Psychol.* **65**, 301–308 (1997).
9. Dadds, M. R. & Salmon, K. Punishment Insensitivity and Parenting: Temperament and Learning as Interacting Risks for Antisocial Behavior. *Clin. Child Fam. Psychol. Rev.* **6**, 69–86 (2003).

10. Oxford, M., Cavell, T. A. & Hughes, J. N. Callous/Unemotional Traits Moderate the Relation Between Ineffective Parenting and Child Externalizing Problems: A Partial Replication and Extension. *J. Clin. Child Adolesc. Psychol.* **32**, 577–585 (2003).
11. Kochanska, G. Toward a Synthesis of Parental Socialization and Child Temperament in Early Development of Conscience. *Child Dev.* **64**, 325–347 (1993).
12. Waller, R., Gardner, F. & Hyde, L. W. What are the associations between parenting, callous–unemotional traits, and antisocial behavior in youth? A systematic review of evidence. *Clin. Psychol. Rev.* **33**, 593–608 (2013).
13. Briggs-Gowan, M. J. *et al.* Punishment insensitivity and impaired reinforcement learning in preschoolers. *J. Child Psychol. Psychiatry* **55**, 154–161 (2014).
14. Kohls, G. *et al.* Investigating Sex Differences in Emotion Recognition, Learning, and Regulation Among Youths With Conduct Disorder. *J. Am. Acad. Child Adolesc. Psychiatry* **59**, 263–273 (2020).
15. Kohls, G. *et al.* Neuropsychological Subgroups of Emotion Processing in Youths With Conduct Disorder. *Front. Psychiatry* **11**, (2020).
16. O'Brien, B. S. & Frick, P. J. Reward dominance: Associations with anxiety, conduct problems, and psychopathy in children. *J. Abnorm. Child Psychol.* **24**, 223–240 (1996).
17. Tenenbaum, R. B. *et al.* Specificity of Reward Sensitivity and Parasympathetic-Based Regulation among Children with Attention-Deficit/Hyperactivity and Disruptive Behavior Disorders. *J. Abnorm. Child Psychol.* **46**, 965–977 (2018).

18. Elster, E. M. *et al.* Impaired Punishment Learning in Conduct Disorder. *J. Am. Acad. Child Adolesc. Psychiatry* (2023) doi:10.1016/j.jaac.2023.05.032.

19. Zhang, R. *et al.* Dysfunction in differential reward-punishment responsiveness in conduct disorder relates to severity of callous-unemotional traits but not irritability. *Psychol. Med.* 1–11 (2021) doi:10.1017/S0033291721003500.

20. Waller, R. *et al.* Callous-Unemotional Behavior and Early-Childhood Onset of Behavior Problems: The Role of Parental Harshness and Warmth. *J. Clin. Child Adolesc. Psychol.* **44**, 655–667 (2015).

21. Pauli, R. & Lockwood, P. L. The computational psychiatry of antisocial behaviour and psychopathy. *Neurosci. Biobehav. Rev.* **145**, 104995 (2023).

22. Jean-Richard-dit-Bressel, P. *et al.* Punishment insensitivity in humans is due to failures in instrumental contingency learning. *eLife* **10**, e69594 (2021).

23. Pauli, R. *et al.* Action initiation and punishment learning differ from childhood to adolescence while reward learning remains stable. *Nat. Commun.* **14**, 5689 (2023).

24. Lockwood, P. L. & Klein-Flügge, M. C. Computational modelling of social cognition and behaviour—a reinforcement learning primer. *Soc. Cogn. Affect. Neurosci.* (2021) doi:10.1093/scan/nsaa040.

25. Sutton, R. S. & Barto, A. G. *Reinforcement Learning, Second Edition: An Introduction*. (MIT Press, 2018).

26. Raab, H. A. & Hartley, C. A. Adolescents exhibit reduced Pavlovian biases on instrumental learning. *Sci. Rep.* **10**, 15770 (2020).

27. Eckstein, M. K., Wilbrecht, L. & Collins, A. G. What do reinforcement learning models measure? Interpreting model parameters in cognition and neuroscience. *Curr. Opin. Behav. Sci.* **41**, 128–137 (2021).
28. White, S. F. *et al.* Disrupted Expected Value and Prediction Error Signaling in Youths With Disruptive Behavior Disorders During a Passive Avoidance Task. *Am. J. Psychiatry* **170**, 315–323 (2013).
29. White, S. F. *et al.* Disrupted Expected Value Signaling in Youth With Disruptive Behavior Disorders to Environmental Reinforcers. *J. Am. Acad. Child Adolesc. Psychiatry* **53**, 579-588.e9 (2014).
30. Newman, J. P. & Kosson, D. S. Passive avoidance learning in psychopathic and nonpsychopathic offenders. *J. Abnorm. Psychol.* **95**, 252–256 (1986).
31. Blair, R. J. R. *et al.* Passive avoidance learning in individuals with psychopathy: modulation by reward but not by punishment. *Personal. Individ. Differ.* **37**, 1179–1192 (2004).
32. Finger, E. C. *et al.* Disrupted Reinforcement Signaling in the Orbitofrontal Cortex and Caudate in Youths With Conduct Disorder or Oppositional Defiant Disorder and a High Level of Psychopathic Traits. *Am. J. Psychiatry* **168**, 152–162 (2011).
33. Fanti, K. A., Kyranides, M. N., Lordos, A., Colins, O. F. & Andershed, H. Unique and Interactive Associations of Callous-Unemotional Traits, Impulsivity and Grandiosity with Child and Adolescent Conduct Disorder Symptoms. *J. Psychopathol. Behav. Assess.* **40**, 40–49 (2018).

34. Freitag, C. M. *et al.* Conduct disorder in adolescent females: current state of research and study design of the FemNAT-CD consortium. *Eur. Child Adolesc. Psychiatry* **27**, 1077–1093 (2018).

35. Kaufman, J. *et al.* 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 (1997).

36. Andershed, H. A., Kerr, M., Stattin, H. & Levander, S. Psychopathic traits in non-referred youths : a new assessment tool. in 131–158 (Elsevier, 2002).

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

38. Wechsler, D. *Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV)*. (Psychological Corporation, San Antonio, TX, 2008).

39. Petersen, A. C., Crockett, L., Richards, M. & Boxer, A. A self-report measure of pubertal status: Reliability, validity, and initial norms. *J. Youth Adolesc.* **17**, 117–133 (1988).

40. Wong, A. C.-N., Palmeri, T. J. & Gauthier, I. Conditions for Facelike Expertise With Objects: Becoming a Ziggerin Expert—but Which Type? *Psychol. Sci.* **20**, 1108–1117 (2009).

41. E-Prime. Psychology Software Tools, Inc.

42. RStudio Team. RStudio: Integrated Development for R. RStudio, PBC (2020).

43. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models using lme4. *ArXiv14065823 Stat* (2014).

44. Wagenmakers, E.-J. A practical solution to the pervasive problems of p values. *Psychon. Bull. Rev.* **14**, 779–804 (2007).
45. Jeffreys, H. *The Theory of Probability*. (OUP Oxford, 1998).
46. TheMathWorksInc. MATLAB 2019b. (2019).
47. Cutler, J. *et al.* Ageing is associated with disrupted reinforcement learning whilst learning to help others is preserved. *Nat. Commun.* **12**, 4440 (2021).
48. Huys, Q. J. M. *et al.* Disentangling the Roles of Approach, Activation and Valence in Instrumental and Pavlovian Responding. *PLOS Comput. Biol.* **7**, e1002028 (2011).
49. Lockwood, P. L., Klein-Flügge, M. C., Abdurahman, A. & Crockett, M. J. Model-free decision making is prioritized when learning to avoid harming others. *PNAS* **117**, 27719–27730 (2020).
50. Wittmann, M. K. *et al.* Global reward state affects learning and activity in raphe nucleus and anterior insula in monkeys. *Nat. Commun.* **11**, 3771 (2020).
51. MacKay, D. J. C. *Information Theory, Inference and Learning Algorithms*. (Cambridge University Press, 2003).
52. Blair, R. J. R. The roles of orbital frontal cortex in the modulation of antisocial behavior. in *Biosocial Theories of Crime* (Routledge, 2010).
53. Platje, E. *et al.* Executive Functioning, Reward/Punishment Sensitivity, and Conduct Problems in Boys With Callous-Unemotional Traits. *Int. J. Offender Ther. Comp. Criminol.* **62**, 4008–4023 (2018).

54. Riden, B. S., Kumm, S. & Maggin, D. M. Evidence-Based Behavior Management Strategies for Students With or at Risk of EBD: A Mega Review of the Literature. *Remedial Spec. Educ.* **43**, 255–269 (2022).

55. Eckstein, M. K., Master, S. L., Dahl, R. E., Wilbrecht, L. & Collins, A. G. E. Reinforcement learning and Bayesian inference provide complementary models for the unique advantage of adolescents in stochastic reversal. *Dev. Cogn. Neurosci.* **55**, 101106 (2022).