



Beyond hostility: exploring facial emotion recognition biases in youths with conduct disorder

Janine Bacher¹ · Beryll von Planta¹ · Anka Bernhard^{2,3} · Graeme Fairchild⁴ · Lucrez Jansen⁵ ·
Stephane A. De Brito⁶ · Christine M. Freitag³ · Kerstin Konrad^{7,8} · Christina Stadler¹ · Gregor Kohls² ·
Eva Unternaehrer¹

Received: 15 April 2025 / Accepted: 10 August 2025
© The Author(s) 2025

Abstract

Facial emotion recognition (FER) biases refer to systematic tendencies to recognize specific emotions when processing facial expressions. In youths with conduct disorder (CD), who are characterized by highly impairing antisocial behavior, research on FER biases has focused on hostile attribution biases. This work has shown that youths with CD perceive ambiguous social cues as angry. However, youths with CD may not only show biases towards anger, which is why we investigated FER biases in youths with CD towards the six basic emotions. Within the European FemNAT-CD study, we analyzed data from 610 youths with CD (60% female) and 818 typically developing controls (TDCs; 68% female), aged 9 to 18 years ($M=14.1$, $SD=2.41$ years). FER biases were assessed using the Emotion Hexagon Task by showing morphed emotional expressions and asking participants to choose the predominant emotion. Biases were calculated as tendency towards an emotion shown at 0%, 10%, 30%, or 50% intensity. Our findings from hierarchical linear modelling indicate that youths with CD exhibited stronger FER biases than TDCs across all emotions, meaning that they misclassified each emotion more often. However, this difference varied by intensity, with youths with CD displaying weaker biases at higher intensity levels and a smaller increase in bias with increasing intensity level. Our findings indicate that youths with CD not only show a hostile attribution bias but rather misclassify emotions as predominant when they are present at low intensity, regardless of type of emotion.

Keywords Emotion recognition bias · Facial emotion recognition · Conduct disorder · Antisocial behavior · Adolescent · FemNAT-CD

Gregor Kohls and Eva Unternaehrer contributed equally to this work and share senior authorship.

✉ Janine Bacher
janine.bacher@upk.ch

¹ Child and Adolescent Psychiatric Research Department, University Psychiatric Clinics Basel, Basel, Switzerland

² Department of Child and Adolescent Psychiatry, Medical Faculty, German Center for Child and Adolescent Health (DZKJ), TUD Dresden University of Technology, partner site Leipzig/Dresden, Dresden, Germany

³ Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital Frankfurt, Goethe University, Frankfurt am Main, Germany

⁴ Department of Psychology, University of Bath, Bath, UK

⁵ Department of Child and Adolescent Psychiatry, VU University Medical Center, Amsterdam, the Netherlands

⁶ Centre for Human Brain Health, School of Psychology, University of Birmingham, Birmingham, UK

⁷ Child Neuropsychology Section, Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, RWTH Aachen University, Aachen, Germany

⁸ JARA-Brain Institute II, Molecular Neuroscience and Neuroimaging, RWTH Aachen & Research Centre Juelich, Juelich, Germany

Introduction

Youths with conduct disorder (CD) show highly impairing patterns of aggressive and antisocial behavior [1]. They also display impairments in the ability to recognize facial expressions [2], which is crucial for understanding social interactions [3]. Facial expressions serve as nonverbal cues communicating emotional states and intentions, influencing how others respond [4]. Being able to identify these expressions supports accurate interpretation of social contexts and improves interpersonal interactions [5].

Previous studies investigating facial emotion recognition (FER) impairments in youths with CD have primarily focused on FER accuracy, that is, whether emotions are correctly identified [e.g., 6]. Rather than focusing solely on FER accuracy, biases in FER may, however, be more predictive of antisocial behavior [6]. Biases in FER can be defined as systematic tendencies towards recognizing specific emotions when processing facial expressions [7], which can lead to distorted social information processing and maladaptive behavioral responses [6]. This distinction is conceptually important, as categorization accuracy alone does not reveal what emotion is chosen instead, which is a key piece of information when trying to understand emotion recognition impairments in CD. Youths with CD might show a tendency towards perceiving ambiguous emotional expressions as angry (i.e., anger bias) [8]. Based on the social information-processing theory proposed by Dodge and Crick [9], this in return may cause them to respond aggressively (i.e., with anger) and increase the likelihood of being rejected by peers due to their aggressive behavior [10]. Youths with CD might also feel rejected if they show a tendency towards perceiving an ambiguous emotional expression as disgust (i.e., disgust bias), since perceiving facial expressions of disgust can trigger increased shame [11].

Importantly, FER biases are not the same as attributional biases, which are grounded in attribution theory [12]. They refer to how individuals explain the causes of others' behaviors [13] and have been studied in youths with CD in terms of a hostile attribution biases [14]. FER biases also differ from attentional biases, which describe the tendency to pay more attention to certain emotional stimuli over others and show difficulties disengaging from them [15]. Our construct of FER biases specifically refers to systematic tendencies towards specific emotions when processing and labelling facial expressions, without involving explanations of behaviors or the duration of attention towards emotional stimuli. Given this distinction, the present study concentrates on examining FER biases independently from attributional or attentional biases.

A central question is for which emotions youths with CD show a FER bias. In a study investigating FER, youths

with conduct problems aged 7 to 13 years showed a bias for the emotion anger when labelling facial expressions [8]. In contrast, Ciucci and colleagues [16] found no association between conduct problems in youths aged 11 to 14 years and FER bias towards anger. The scale they used in their study to assess conduct problems was reported to correlate highly with independent clinical diagnoses of conduct disorder [17]. But research so far has not only recognized an anger bias in youths with CD: Short and colleagues [18] found that youths with CD aged 12 to 18 years were more likely to classify neutral faces as fearful and therefore showed a fear bias rather than an anger bias. In related conditions like general disruptive behavior, youths between the ages of 8 and 18 showed FER biases for anger, happiness and surprise depending on aggression subtype and callous-unemotional traits [19]. Taken together, previous findings on FER biases in youths with CD are inconsistent, suggesting that biases may not be limited to anger but could also involve fear, happiness, or surprise.

The aim of this study was to examine FER biases in youths with CD for all six basic emotions compared to typically developing controls (TDCs) across different intensity levels, since analyzing FER biases across emotion intensities may offer deeper insights into the emotional processing mechanisms underlying antisocial behavior in youths with CD [20]. Overall, we expected differences in FER biases between youths with CD and TDCs based on prior research on youths with CD or related conditions [8, 18, 19]. We also expected that these differences would vary depending on emotion and intensity level. However, we investigated these FER biases in bi-directional hypotheses due to the lack of previous evidence.

Methods

Participants and procedure

This study included 1,428 participants from the European multi-site project *Neurobiology and Treatment of Adolescent Female Conduct Disorder: The Central Role of Emotion Processing* (FemNAT-CD) who completed the Emotion Hexagon Task and fulfilled our data quality criteria (see Online Resource 1). The sample consisted of 610 (60% female) youths diagnosed with CD and 818 (68% female) TDCs from Germany, the United Kingdom, Spain, the Netherlands, Greece, Switzerland, and Hungary. Participants were aged 9 to 18 years ($M=14.1$, $SD=2.41$ years), and estimated full-scale IQ ranged from 70 to 145 ($M=100.1$, $SD=12.6$). They were recruited through clinical referrals, youth offending services, schools, and community channels (e.g., social media). Inclusion criteria for

the CD group consisted of a current CD diagnosis according to DSM-IV-TR criteria [21]. Participants taking psychotropic medication were assessed while remaining on their prescribed medication (e.g., psychostimulants). Overall exclusion criteria were an IQ below 70, diagnosis of autism spectrum disorders, schizophrenia, bipolar disorder or mania, neurological disorders or genetic syndromes (e.g., Down Syndrome). For the TDCs group, exclusion criteria were any current DSM-IV-TR diagnosis or a history of CD, ODD, or ADHD.

Measures

Facial emotion recognition biases

FER biases were assessed using the Emotion Hexagon Task, which consists of identifying facial expressions based on the Ekman-Friesen Pictures of Facial Affect [22, 23]. Each stimulus is a computer-generated blend (morph) of two emotions at varying intensities (e.g., 90% anger and 10% disgust). After a practice block, participants completed five blocks, each containing 30 trials. In each trial, participants viewed a facial expression for 3 s and were asked to state which emotion best described the expression shown. They could choose between labels for the six basic emotions happiness, surprise, fear, sadness, disgust and anger. Blends included 10%:90%, 30%:70% and 50%:50% intensities. The specific emotion combinations were based on confusability norms established by Ekman and Friesen [23], with each emotion placed adjacent to the one it was most frequently confused with [22]. By connecting the ends of the sequence to form a hexagon, six emotion continua were created for the Emotion Hexagon Task by morphing images along each edge: happiness–surprise, surprise–fear, fear–sadness, sadness–disgust, disgust–anger, and anger–happiness. E-Prime was used to implement the task and present the stimuli in random order.

In the present study, FER biases were calculated for each of the six basic emotions and computed across four levels of emotional expression intensity (0%, 10%, 30%, and 50%). To account for expressions with maximum ambiguity, in accordance with Schönenberg and Jusyte [24], we also included as bias when the indicated emotion was

represented at 50% intensity in a blend, as such bias would still reflect a tendency towards a specific emotion. The computation procedure was identical for all FER biases; for illustrative purposes, the process is described in detail below using the example of the anger bias (Table 1). The dataset was first filtered to include only trials where participants selected *anger* as the emotion that best described the expression shown. In this filtered dataset, for the 0% intensity level, all trials were included where anger was not present in the stimulus (i.e., combinations of happiness–surprise, surprise–fear, fear–sadness, and sadness–disgust). For 10%, 30%, and 50% intensity levels, trials were included where *anger* was blended with either of its two neighboring emotions (*happiness* or *disgust*) at the respective proportions (e.g., 10% anger : 90% happiness, or 10% anger : 90% disgust). Therefore, the 0% intensity level biases included a maximum of 100 possible trials, while each of the other intensity levels included up to 10 trials. To account for this difference, we calculated percentage bias scores by dividing the number of *anger* responses by the total possible number of trials at each intensity level and multiplying by 100. Final bias scores were then computed as the mean percentage bias across all intensity levels for each emotion.

Demographic variables

In addition to age and biological sex, cognitive abilities (IQ) and socioeconomic status (SES) were assessed. IQ was measured using the vocabulary and matrix reasoning subtests of the Wechsler Intelligence Scale (WISC-III-R/WISC-IV for adolescents up to the age of 16, WAIS-III-R/WAIS-IV for adolescents aged 17–18) or the Wechsler Abbreviated Scale of Intelligence (WASI) at UK sites [25–28]. SES was scored based on the Medical History Questionnaire; a semi-structured interview specifically designed for this study which was used to gather data regarding the medical history of participant [29]. The interview included questions about parental income, education level and occupation and was used to score SES by principal component extraction standardized per country. To assess parental educational status, ISCED-classifications were used [30].

Table 1 Calculation of FER biases on the example of the anger bias

Emotion Bias	Intensity Level			
	0%	10%	30%	50%
Anger	All other combinations:	10% Anger :	30% Anger :	50% Anger :
	Happiness : Surprise ^b	90% Happiness ^a	70% Happiness ^a	50% Happiness ^a
	Surprise : Fear ^b	10% Anger :	30% Anger :	50% Anger :
	Fear : Sadness ^b	90% Disgust ^a	70% Disgust ^a	50% Disgust ^a
	Sadness : Disgust ^b			

^a5 trials

^b25 trials

Statistical analyses

All analyses were conducted using R Statistical Software version 4.4.2 [31] for Windows (for details of used packages see Online Resource 1). We tested for demographic differences between groups using chi-square tests for categorical variables (age, sex) and t-tests for continuous variables (IQ, SES). For all analyses, p-values below 0.05 were considered statistically significant. Multilevel models (MLMs) were used to account for dependent observations within participants. On the person level, we included the effect of group to explore group differences in FER biases between youths with CD and TDCs and controlled for age, sex, IQ and SES. On the observation level, we added emotion type as factor and intensity as a continuous variable to see whether these influenced FER biases. MLMs account for multiple comparisons through partial pooling, meaning that person- or observation-level estimates are “shrunk” toward the overall mean, which reduces the risk of false positives [32]. To support the validity of this approach, we report model fit comparisons of alternative models (see ‘Preliminary findings – model fit’ [33]. We gradually added effects on both levels and compared the model fits using Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and the Maximum Likelihood Ratio Test. We also calculated the marginal R^2 to estimate effect sizes.

Results

Descriptive statistics

The demographic variables per group and comorbidities of the CD group are presented in Table 2. There was a significant group difference in age, with youths with CD being older than TDCs, and sex, with the CD group having a lower proportion

of female participants. The groups significantly differed in IQ, with youths with CD having lower IQs than TDCs. Youths with CD also had a significantly lower SES than TDCs.

Multilevel analysis

Preliminary findings – model fit We hierarchically built multilevel models and compared model fit to find the best-fitting solution (see Online Resource; Supplementary Table 1). To ensure transparent model selection, we detail this process by reporting the fit comparisons of several alternative models. First, a baseline model (null model) was fitted, where FER biases were predicted by a fixed intercept. Then, a random intercept model was tested to account for clustering of observations within participants. However, a model comparison suggested that the random intercept did not significantly improve model fit compared to a fixed intercept model. Despite this, we continued with the random intercept model because of structural reasons as trials were nested within participants to account for possible dependency in the data and to allow for further model extensions. Several model comparisons showed that the final model included the fixed effect of type of emotion, a quadratic fixed effect of intensity of emotion, group, as well as the interaction terms group * emotion, group * intensity, the three-way interaction term group * emotion * intensity and the covariates age and sex. Adding IQ and SES as control variables did not improve model fit, neither did including group * age, group * sex or group * IQ as interaction terms. Marginal R^2 for the final model including all significant predictors, covariates and interactions was 0.52, meaning that the fixed effects in the model explained 52% of the variance in FER biases.

Main findings – final model The final model revealed that the CD group had overall significantly stronger FER biases than TDCs (Table 3). There was no significant interaction

Table 2 Descriptive statistics of demographic variables and comorbidities

	CD <i>n</i> =610	TDCs <i>n</i> =818	Group effect <i>p</i> -values ^a
Age (years) <i>M</i> (SD)	14.29 (2.3)	13.97 (2.47)	<0.05
Females <i>n</i> (%)	365 (59.8)	553 (67.6)	<0.01
IQ <i>M</i> (SD)	95.7 (11.87)	103.39 (12.17)	<0.01
SES <i>M</i> (SD) ^b	-0.29 (0.95)	0.29 (1.01)	<0.01
Major Depressive Disorder <i>n</i> (%)	85 (13.93)		
Attention Deficit Disorder <i>n</i> (%)	224 (36.72)		
Posttraumatic Stress Disorder <i>n</i> (%)	45 (7.38)		
Oppositional Defiant Disorder <i>n</i> (%)	491 (80.49)		
Generalized Anxiety Disorder <i>n</i> (%)	20 (3.28)		
Social Phobia <i>n</i> (%)	26 (4.26)		
Psychotropic Medication <i>n</i> (%)	190 (31.15)		

^a*p*-values are based on two-sample t-tests or χ^2 tests

^b principal component extracted from socioeconomic status standardized by country (*M*=0, *SD*=1)

CD Conduct Disorder, *TDCs* typically developing controls

Table 3 Final multilevel model: results predicting FER biases

	B	SE	df	t	p
(Intercept)	19.4	0.52	32,810	37.22	<0.001
Type of Emotion					
Anger	Reference				
Disgust	1.78	0.35	32,810	5.07	<0.001
Fear	-2.08	0.35	32,810	-5.9	<0.001
Happiness	-4.39	0.35	32,810	-12.48	<0.001
Sadness	1.21	0.35	32,810	3.43	<0.001
Surprise	1.94	0.79	32,810	5.5	<0.001
Intensity					
Linear	2451.13	46.1	32,810	53.17	<0.001
Quadratic	956.98	46.1	32,810	20.76	<0.001
Age	-0.25	0.03	1,424	-7.5	<0.001
Male (vs. Female)	1.01	0.16	1,424	6.25	<0.001
Group CD (vs. TDCs)	0.77	0.38	1,424	2.01	0.04
Group * Emotion					
CD * Anger	Reference				
CD * Disgust	-0.84	0.54	32,810	-1.56	0.12
CD * Fear	-0.65	0.54	32,810	-1.2	0.23
CD * Happiness	-0.02	0.54	32,810	-0.04	0.97
CD * Sadness	-0.51	0.54	32,810	-0.95	0.34
CD * Surprise	1.2	0.54	32,810	2.22	0.03
Group * Intensity					
CD * Linear	-342.07	70.53	32,810	-4.85	<0.001
CD * Quadratic	-378.2	70.53	32,810	-5.36	<0.001
Group * Intensity * Emotion					
TDCs * Linear * Anger	Reference				
TDCs * Linear * Disgust	103.26	65.2	32,810	1.58	0.11
TDCs * Linear * Fear	82.74	65.2	32,810	1.27	0.2
TDCs * Linear * Happiness	377.81	65.2	32,810	5.8	<0.001
TDCs * Linear * Sadness	817.86	65.2	32,810	12.54	<0.001
TDCs * Linear * Surprise	207.74	65.2	32,810	3.19	0.001
TDCs * Quadratic * Anger	Reference				
TDCs * Quadratic * Disgust	-171.74	65.2	31,382	-2.63	0.008
TDCs * Quadratic * Fear	221.81	65.2	31,382	3.4	<0.001
TDCs * Quadratic * Happiness	582.03	65.2	31,382	8.93	<0.001
TDCs * Quadratic * Sadness	233.07	65.2	31,382	3.57	<0.001
TDCs * Quadratic * Surprise	-161.29	65.2	31,382	-2.47	0.01
CD * Linear * Anger	Reference				
CD * Linear * Disgust	3.97	75.5	31,382	0.05	0.96
CD * Linear * Fear	-4.28	75.5	31,382	-0.06	0.95
CD * Linear * Happiness	652.16	75.5	31,382	8.64	<0.001
CD * Linear * Sadness	832.73	75.5	31,382	11.03	<0.001
CD * Linear * Surprise	354.75	75.5	31,382	4.7	<0.001
CD * Quadratic * Anger	Reference				
CD * Quadratic * Disgust	97.91	75.5	31,382	1.3	0.19
CD * Quadratic * Fear	290.55	75.5	31,382	3.85	<0.001
CD * Quadratic * Happiness	878.13	75.5	31,382	11.63	<0.001
CD * Quadratic * Sadness	413.22	75.5	31,382	5.47	<0.001
CD * Quadratic * Surprise	-26.56	75.5	31,382	-0.35	0.72

CD Conduct Disorder, TDCs typically developing controls, B unstandardized beta, SE standard error, df degrees of freedom, t t-statistic

between group and type of emotion, indicating that group differences in FER biases did not vary systematically across emotions. However, there was a significant interaction between group and intensity, indicating that the strength of FER biases across intensity levels differed between groups (Fig. 1).

Specifically, youths with CD showed stronger FER biases than TDCs at lower intensity levels, but a smaller increase with increasing intensity level. This pattern is indicated by the significant negative quadratic effect. The three-way interaction between group, type of emotion and intensity level further revealed that this difference in FER biases across intensity levels between groups also differed between emotions. We demonstrate this complex interaction by comparing two example emotions. For the emotion fear, the difference between groups at the 50% intensity level was greater than for the emotion surprise, whereas at the 30% intensity level, the difference between groups in FER biases was smaller for fear than for surprise. Overall, the pattern

of intensity-related changes in FER biases differed not only between groups but also across emotions.

Discussion

The most commonly studied aspect of FER biases in youths with CD is the anger bias, whereby ambiguous facial expressions are misinterpreted as displaying anger, which could be perceived as a threat that potentially leads to aggression [8]. However, FER biases in youths with CD might be broader beyond a bias towards anger (e.g., bias towards fear or happiness) [18, 19] and depend on the intensity level of the presented facial expression. This study therefore expands the current knowledge base by examining FER biases in youths with CD compared to TDCs for all six basic emotions considering both type and different intensity levels of emotion.

Our multilevel model showed that, overall, youths with CD had higher FER biases than TDCs. While some studies have reported stronger anger biases in youths with CD

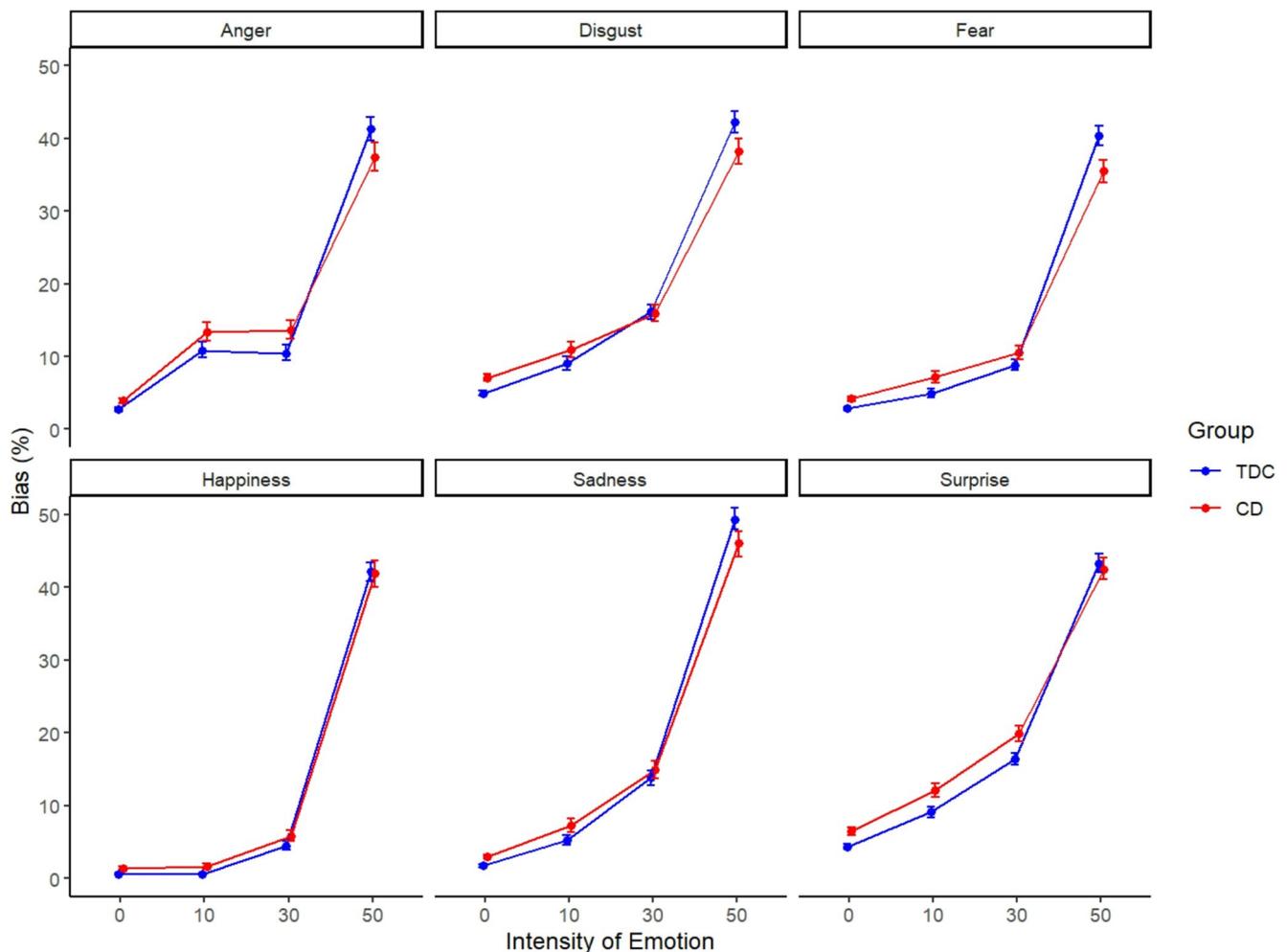


Fig. 1 Facial emotion recognition biases for each emotion and intensity level by group

[14, 34] and others have found no such effect [16], our findings therefore suggest that FER biases in CD extend beyond anger. In contrast to prior research highlighting increased biases for specific emotions - such as fear in the study by Short and colleagues [18] - youths with CD in our study exhibited stronger FER biases across all emotions. This is in line with FER accuracy research, showing that youths with CD perform worse in recognizing emotions in facial expressions [35]. In general, FER accuracy and FER biases overlap in the way that overall poorer emotion recognition accuracy also contributes to higher biases. However, FER accuracy alone does not indicate which incorrect emotions are chosen. For instance, youths with CD might show lower FER accuracy across all emotions but only exhibit a stronger bias toward anger. Alternatively, they could show lower FER accuracy across all emotions while having stronger biases for multiple emotions such as anger, sadness, and surprise. With our findings, we could show that lower FER accuracy in youths with CD is not driven by a bias toward a single emotion, but rather by generally increased FER biases. Due to the importance of FER for social interactions [4, 5], these findings seem to indicate a particular challenge in social situations for youths with CD. If these youths show increased tendencies to misinterpret emotional expressions, they are at risk to respond improperly to the other person and not react appropriately to a situation [5].

Our findings on FER biases also contribute to a novel understanding of how different emotions at different intensity levels are perceived: The difference between groups in FER biases did not vary with the type but rather with the intensity of emotions, such that the CD group had stronger biases at lower intensity levels compared to TDCs, and showed a smaller increase with increasing intensity level. It is theoretically expected that FER biases will get stronger as the intensity of the respective emotion increases, since higher intensity levels make emotional expressions clearer. The fact that the CD group showed a weaker increase in FER biases compared to TDCs and values below 50% for the 50% intensity level suggests that, under maximum ambiguity, they are more likely to select an emotion that is not present in the facial expression at all. While values below 50% for the 50% intensity level cannot be interpreted as a bias per se, the difference between youths with CD and TDCs at this intensity level might show that youths with CD make more random responses or errors. This pattern could, in turn, explain the higher FER biases of youths with CD observed at 0% intensity.

Our model showed a three-way interaction between group, emotion, and intensity level, indicating that group differences in FER biases across intensities varied by emotion. Notably, differences between youths with CD and TDCs in anger bias were more pronounced at the 10% and

30% intensity levels than at 0%. This is contradictory to previous research showing that, compared to other individuals with externalizing problems, such as ADHD, or CD and co-occurring ADHD, adolescents with CD only had a significantly stronger FER bias for the emotion anger when anger was not actually one of the expressions presented [12]. Additionally, the observed pattern in anger bias may reflect a heightened sensitivity to low-intensity representations of anger in youths with CD.

Our results contribute to and extend the sparse literature on CD and FER biases by incorporating emotion type as well as intensity level. This is important since investigating FER biases for all basic emotions might contribute to a better understanding of social behavior [6]. Furthermore, it might help to inform novel approaches for the treatment of CD by training emotion recognition skills with FER biases as treatment targets similar to Stoddard and colleagues [36]. These authors found that a computer-based training program, in which adolescents received positive feedback for rating ambiguous facial morphs as happy and negative feedback for rating them as angry, could shift anger biases toward more positive classifications of emotions. This training was also associated with reduced irritability and changes in brain activity in adolescents with Disruptive Mood Dysregulation Disorder (DMDD; which is characterized by frequent and intense anger and temper outbursts), indicating a potential novel treatment approach for CD youths with FER bias problems as well.

Strengths and limitations

This study is the first to investigate FER biases in youths with CD for all six basic emotions and across different intensity levels. The use of low intensity and more ambiguous facial expressions reveals more subtle differences between groups [37] as these stimuli are also more sensitive to biases than prototypical emotional expressions [37]. In addition, multilevel models used are excellently suited to analyzing observations nested in individuals and investigating the influence of intensity and emotion type through their hierarchical structure [38]. Including group as a fixed effect and sex as a control variable in the final multilevel model ensures that observed group effects are not confounded by the differences between groups in sex distribution.

Despite the methodological strengths of this study, several limitations should be considered. First, the cross-sectional design of the study limits the ability to draw conclusions about the direction or existence of a causal relationship between group status (i.e., having CD) and FER biases. This leaves it unclear whether FER biases might serve as a predisposing risk factor for CD or whether CD itself might contribute to the development of FER biases. To examine

whether FER biases predict future clinical outcomes (e.g., persistence or resistance of CD over time) or certain types of behaviors (e.g., aggressive symptoms), future longitudinal studies could measure changes in CD symptoms and FER biases over time. Our approach also does not fully capture the heterogeneity of CD and youths with CD showing different comorbidities, which could possibly influence FER biases. Furthermore, real-world social interactions involve dynamic, context-dependent emotional expressions, which are not fully captured in static FER tasks like the Emotion Hexagon Task. Future studies should consider incorporating more dynamic expressions, for example video-based, to improve ecological validity and better understand FER biases in CD.

Although our multilevel models accounted for individual differences and the hierarchical data structure, the inclusion of a random slope was not possible due to model convergence issues. This may limit the generalizability of the findings, as individual variations in response patterns could not be fully accounted for. Also, possible site effects were not included in the model. Adding a varying intercept did not lead to significant model improvement, which is surprising in view of the heterogeneity of neurocognitive functions in youths with CD [29]. Although significant, adding group as a fixed effect did not appear to make a large contribution to model improvement in both cases. Comparing the marginal variance for all multilevel models, the inclusion of emotion intensity seemed to improve model fit the most. Lastly, given the substantial overlap between accuracy and FER biases, we refrained from statistically controlling for overall accuracy to avoid removing meaningful variance from the bias scores. However, future studies might explore alternative methodological approaches to disentangle the influence of general accuracy on FER bias scores.

Conclusion

To our knowledge, this study is the first to examine differences in FER biases across all basic emotions in youths with CD compared to TDCs. Our findings suggest that youths with CD show increased FER biases compared to TDCs irrespective of type of emotion. However, the strength of these differences changes with emotion intensity, with the CD group showing a smaller increase in FER biases as intensity increases. These results underscore the importance of further investigating the potential neuropsychological mechanisms underlying such FER biases in youths with CD and examining whether FER training can reduce FER biases in CD populations. Given that FER biases are linked to social difficulties and antisocial behavior, interventions aimed at improving specific emotion recognition skills are a valuable component of treatment approaches for CD.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00787-025-02846-y>.

Acknowledgements The authors thank all participants and their families for taking part in this study, all institutions that were willing to collaborate with us for this study, and all members of the FemNAT-CD consortium.

Author contributions J.B. conducted all analyses and wrote the first draft of the manuscript. E.U. and G.K. contributed substantially to the manuscript, provided guidance and oversaw all aspects of the analyses. All other authors were involved in the 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 of the final manuscript.

Funding Open access funding provided by University of Basel. This study was funded by the European Commission's Seventh Framework Programme (FP7/2007–2013) under Grant Agreement no. 602407 (FemNAT-CD, coordinator: C.M.F.). G.K. was supported by a 2023 NARSAD Young Investigator Grant from the Brain & Behavior Research Foundation (BBRF; Grant No. 30849). This work was also supported by the Federal Ministry of Education and Research (BMBF) as part of the German Center for Child and Adolescent Health (DZKJ) under the funding code 01GL2405B.

Data availability The data that support the findings of this study are not openly available. However, interested researchers can apply to the FemNAT-CD steering committee (chaired by C.M.F.) for data access as part of a collaboration.

Declarations

Ethics approval 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 Ethische 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). The study was performed in accordance with the ethical standards outlined in the 1964 Declaration of Helsinki and its later amendments.

Consent to participate Informed consent was obtained from all participants aged 16 years and older. For participants under the age of 16, assent was obtained from the participants themselves, and informed consent was provided by their legal guardian(s).

Consent to publish All participants, as well as their legal guardian(s), signed informed consent regarding publishing their data.

Competing interests C.S. receives royalties for a book on aggression. C.M.F. receives royalties for books on ADHD, ASD and Depression. She has served as a consultant to Desitin and Roche. K.K. has received

speaker fees from Shire Pharmaceuticals and Medice. S.A.D.B. has received speaker fees from the Child Mental Health Centre and the Centre for Integrated Molecular Brain Imaging. The remaining authors have declared that they have no competing or potential conflicts of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- American Psychiatric Association (2022) Disruptive, Impulse-Control, and conduct disorders. In: Diagnostic and Statistical Manual of Mental Disorders, text rev
- a. Martin-Key N, w. Graf E, j. Adams W, Fairchild G (2018) Facial emotion recognition and eye movement behaviour in conduct disorder. *J Child Psychol Psychiatry* 59:247–257. <https://doi.org/10.1111/jcpp.12795>
- Camras LA, Halberstadt AG (2017) Emotional development through the lens of affective social competence. *Curr Opin Psychol* 17:113–117. <https://doi.org/10.1016/j.copsyc.2017.07.003>
- Horstmann G (2003) What do facial expressions convey: feeling states, behavioral intentions, or actions requests? *Emotion* 3:150–166. <https://doi.org/10.1037/1528-3542.3.2.150>
- de Paiva-Silva AI, Pontes MK, Aguiar JSR, de Souza WC (2016) How do we evaluate facial emotion recognition? *Psychol Neurosci* 9:153–175. <https://doi.org/10.1037/pne0000047>
- Barth JM, Bastiani A (1997) A longitudinal study of emotion recognition and preschool children's social behavior. *Merrill-Palmer Q* 43:107–128
- Vikhanova A, Mareschal I, Tibber M (2022) Emotion recognition bias depends on stimulus morphing strategy. *Atten Percept Psychophys* 84:2051–2059. <https://doi.org/10.3758/s13414-022-02532-0>
- Cadesky EB, Mota VL, Schachar RJ (2000) Beyond words: how do children with ADHD and/or conduct problems process non-verbal information about affect?? *J Am Acad Child Adolesc Psychiatry* 39:1160–1167. <https://doi.org/10.1097/00004583-200009000-00016>
- Social Information-Processing Bases of Aggressive Behavior in Children -, Kenneth A, Dodge NR, Crick (1990) <https://journals.sagepub.com/doi/abs/10.1177/0146167290161002>. Accessed 5 Aug 2025
- Favre CA, Garrote A, Kassir W et al (2024) The cycle of violence: effects of violence experience, behavior, and attitudes on adolescents' peer rejection networks. *Front Educ*. <https://doi.org/10.3389/educ.2024.1359558>
- Giner-Sorolla R, Espinosa P (2011) Social cuing of guilt by anger and of shame by disgust. *Psychol Sci* 22:49–53. <https://doi.org/10.1177/0956797610392925>
- Harvey J, Weary G (1984) Current issues in attribution theory and research. *Annu Rev Psychol* 35:427–459. <https://doi.org/10.1146/annurev.ps.35.020184.002235>
- Korn CW, Rosenblau G, Buritica JMR, Heekeren HR (2016) Performance feedback processing is positively biased as predicted by attribution theory. *PLoS One* 11:e0148581. <https://doi.org/10.1371/journal.pone.0148581>
- Kou H, Luo W, Li X et al (2022) Cognitive deficits for facial emotions among male adolescent delinquents with conduct disorder. *Front Psychiatry*. <https://doi.org/10.3389/fpsy.2022.937754>
- Pool E, Brosch T, Delplanque S, Sander D (2016) Attentional bias for positive emotional stimuli: a meta-analytic investigation. *Psychol Bull* 142:79–106. <https://doi.org/10.1037/bul0000026>
- Ciucci E, Baroncelli A, Facci C et al (2024) Callous–unemotional traits and emotion perception accuracy and bias in youths. *Children (Basel)* 11:419. <https://doi.org/10.3390/children11040419>
- Goodman R, Renfrew D, Mullick M (2000) Predicting type of psychiatric disorder from strengths and difficulties questionnaire (SDQ) scores in child mental health clinics in London and Dhaka. *Eur Child Adolesc Psychiatry* 9:129–134. <https://doi.org/10.1007/s007870050008>
- Short RML, Sonuga-Barke EJS, Adams WJ, Fairchild G (2016) Does comorbid anxiety counteract emotion recognition deficits in conduct disorder? *J Child Psychol Psychiatry* 57:917–926. <https://doi.org/10.1111/jcpp.12544>
- Kleine Deters R, Naaijen J, Holz NE et al (2023) Emotion recognition profiles in clusters of youth based on levels of callous-unemotional traits and reactive and proactive aggression. *Eur Child Adolesc Psychiatry* 32:2415–2425. <https://doi.org/10.1007/s00787-022-02079-3>
- Bowen KL, Morgan JE, Moore SC, van Goozen SHM (2014) Young offenders' emotion recognition dysfunction across emotion intensities: explaining variation using psychopathic traits, conduct disorder and offense severity. *J Psychopathol Behav Assess* 36:60–73. <https://doi.org/10.1007/s10862-013-9368-z>
- American Psychiatric Association (2000) Diagnostic and Statistical Manual of Mental Disorders, 4th ed, Text Revision. Washington, DC
- Calder AJ (1996) Facial emotion recognition after bilateral amygdala damage: differentially severe impairment of fear. *Cogn Neurosci* 13:699–745. <https://doi.org/10.1080/026432996381890>
- Ekman P, Friesen WV (1976) Pictures of Facial Affect
- Schönenberg M, Jusyte A (2014) Investigation of the hostile attribution bias toward ambiguous facial cues in antisocial violent offenders. *Eur Arch Psychiatry Clin Neurosci* 264:61–69. <https://doi.org/10.1007/s00406-013-0440-1>
- Wechsler D (1999) Wechsler Abbreviated Scale of Intelligence
- Wechsler D (1997) Wechsler Adult Intelligence Scale—Third Edition
- Wechsler D (2003) Wechsler intelligence scale for children—Fourth edition. WISC-IV
- Wechsler D (2008) Wechsler Adult Intelligence Scale—Fourth Edition
- Kohls G, Fairchild G, Bernhard A et al (2020) Neuropsychological subgroups of emotion processing in youths with conduct disorder. <https://doi.org/10.3389/fpsy.2020.585052>. *Front Psychiatry* 11:
- Organisation for Economic Co-Operation and Development (1999) Classifying Educational Programmes – Manual for ISCED-97 Implementation in OECD Countries – 1999 Edition
- R Core Team (2021) R: The R Project for Statistical Computing. In: R Lang. Environ. Stat. Comput. R Found. Stat. Comput. Vienna Austria. <https://www.r-project.org/>. Accessed 14 Apr 2025
- Gelman A, Hill Jennifer, Yajima M (2012) Why we (Usually) don't have to worry about multiple comparisons. *J Res Educ Eff* 5:189–211. <https://doi.org/10.1080/19345747.2011.618213>

33. Bell A, Holman D, Jones K (2019) Using shrinkage in multilevel models to understand intersectionality. *Methodology*. <https://doi.org/10.1027/1614-2241/a000167>
34. Hartmann D, Ueno K, Schwenck C (2020) Attributional and attentional bias in children with conduct problems and callous-unemotional traits: a case-control study. *Child Adolesc Psychiatry Ment Health* 14:9. <https://doi.org/10.1186/s13034-020-00315-9>
35. Kohls G, Baumann S, Gundlach M 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. <https://doi.org/10.1016/j.jaac.2019.04.003>
36. Stoddard J, Sharif-Askary B, Harkins EA et al (2016) An open pilot study of training hostile interpretation bias to treat disruptive mood dysregulation disorder. *J Child Adolesc Psychopharmacol* 26:49–57. <https://doi.org/10.1089/cap.2015.0100>
37. Adolphs R, Tranel D (2004) Impaired judgments of sadness but not happiness following bilateral amygdala damage. *J Cogn Neurosci* 16:453–462. <https://doi.org/10.1162/089892904322926782>
38. Bauer DJ, McNeish DM, Baldwin SA, Curran PJ (2020) Analyzing nested data: multilevel modeling and alternative approaches. In: Wright AGC, Hallquist MN (eds) *The Cambridge handbook of research methods in clinical psychology*. Cambridge University Press, Cambridge, pp 426–443

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.