



Neuropsychological Subgroups of Emotion Processing in Youths With Conduct Disorder

Gregor Kohls¹, Graeme Fairchild², Anka Bernhard³, Anne Martinelli³, Areti Smaragdi⁴, Karen Gonzalez-Madruga⁵, Amy Wells⁶, Jack C. Rogers⁷, Ruth Pauli⁸, Helena Oldenhof⁹, Lucres Jansen⁹, Arthur van Rhijn⁹, Linda Kersten¹⁰, Janine Alfano¹⁰, Sarah Baumann¹¹, Beate Herpertz-Dahlmann¹¹, Agnes Vetro¹², Helen Lazaratou¹³, Amaia Hervas¹⁴, Aranzazu Fernández-Rivas¹⁵, Arne Popma⁹, Christina Stadler¹⁰, Stephane A. De Brito⁸, Christine M. Freitag³ and Kerstin Konrad^{1,16*}

OPEN ACCESS

Edited by:

Carla Sharp,
University of Houston, United States

Reviewed by:

Isabelle Esther Bauer,
University of Texas Health Science
Center at Houston, United States
Vivek Agarwal,
King George's Medical
University, India
Spencer C. Evans,
Harvard University, United States

*Correspondence:

Kerstin Konrad
k.konrad@fz-juelich.de

Specialty section:

This article was submitted to
Child and Adolescent Psychiatry,
a section of the journal
Frontiers in Psychiatry

Received: 19 July 2020

Accepted: 01 December 2020

Published: 22 December 2020

Citation:

Kohls G, Fairchild G, Bernhard A,
Martinelli A, Smaragdi A,
Gonzalez-Madruga K, Wells A,
Rogers JC, Pauli R, Oldenhof H,
Jansen L, Rhijn Av, Kersten L,
Alfano J, Baumann S,
Herpertz-Dahlmann B, Vetro A,
Lazaratou H, Hervas A,
Fernández-Rivas A, Popma A,
Stadler C, De Brito SA, Freitag CM
and Konrad K (2020)
Neuropsychological Subgroups of
Emotion Processing in Youths With
Conduct Disorder.
Front. Psychiatry 11:585052.
doi: 10.3389/fpsy.2020.585052

¹ Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, Child Neuropsychology Section, University Hospital RWTH Aachen, Aachen, Germany, ² Department of Psychology, University of Bath, Bath, United Kingdom, ³ Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital Frankfurt, Frankfurt am Main, Germany, ⁴ Centre for Addiction and Mental Health, Toronto, ON, Canada, ⁵ King's College London, London, United Kingdom, ⁶ School of Psychology, Cardiff University, Cardiff, United Kingdom, ⁷ School of Psychology, Institute for Mental Health, University of Birmingham, Birmingham, United Kingdom, ⁸ Centre for Human Brain Health, School of Psychology, University of Birmingham, Birmingham, United Kingdom, ⁹ Department of Child and Adolescent Psychiatry, Amsterdam Public Health – Mental Health, Vrije Universiteit Amsterdam, Amsterdam, Netherlands, ¹⁰ Psychiatric University Clinics, University of Basel, Basel, Switzerland, ¹¹ Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital RWTH Aachen, Aachen, Germany, ¹² Child and Adolescent Psychiatry Department, Pediatrics and Child Health Center, University of Szeged, Szeged, Hungary, ¹³ Child and Adolescent Unit of the 1st Department of Psychiatry, National and Kapodistrian University of Athens, Athens, Greece, ¹⁴ Child and Adolescent Mental Health Unit, University Hospital Mutua Terrassa, Barcelona, Spain, ¹⁵ Psychiatric Service, Basurto University Hospital, Bilbao, Spain, ¹⁶ JARA-Brain Institute II, Molecular Neuroscience and Neuroimaging, RWTH Aachen and Research Centre Juelich, Juelich, Germany

Background: At the group level, youths with conduct disorder (CD) show deficient emotion processing across various tasks compared to typically developing controls (TDC). But little is known about neuropsychological subgroups within the CD population, the clinical correlates of emotion processing deficits [for instance, with regard to the presence or absence of the DSM-5 Limited Prosocial Emotions (LPE) specifier], and associated risk factors.

Methods: 542 children and adolescents with CD (317 girls) and 710 TDCs (479 girls), aged 9–18 years, were included from the FemNAT-CD multisite study. All participants completed three neuropsychological tasks assessing emotion recognition, emotion learning, and emotion regulation. We used a self-report measure of callous-unemotional traits to create a proxy for the LPE specifier.

Results: Relative to TDCs, youths with CD as a group performed worse in all three emotion domains. But using clinically based cut-off scores, we found poor emotion recognition skills in only 23% of the participants with CD, followed by emotion regulation deficits in 18%, and emotion learning deficits in 13% of the CD group. Critically, the majority of youths with CD (~56%) did not demonstrate any meaningful neuropsychological deficit, and only a very small proportion showed pervasive deficits across all three domains (~1%). Further analyses indicate that established DSM-5

subtypes of CD are not tightly linked to neurocognitive deficits in one particular emotion domain over another (i.e., emotion recognition deficits in CD+LPE vs. emotion regulation deficits in CD–LPE).

Conclusions: Findings from this large-scale data set suggest substantial neuropsychological diversity in emotion processing in the CD population and, consequently, only a subgroup of youths with CD are likely to benefit from additional behavioral interventions specifically targeting emotion processing mechanisms.

Keywords: conduct disorder (CD), callous-unemotional (CU) traits, limited prosocial emotions specifier, emotion recognition, emotion learning, emotion regulation, neuropsychology, heterogeneity

INTRODUCTION

Conduct disorder (CD) is one of the most prevalent externalizing disorders in childhood and adolescence (1). It is a leading cause of referral to mental health and youth welfare services and incurs enormous healthcare and societal costs (2). Paradoxically, though, CD is one of the least studied, funded, and understood psychiatric disorders in youth (3). Children and adolescents with CD are characterized by severe antisocial and aggressive behaviors that violate age-appropriate societal norms and the rights of others (1). Empirical data emphasize that CD is a highly heterogeneous condition in terms of clinical phenotype (including different subtypes and psychiatric comorbidities), clinical course (i.e., persistent vs. desisting symptomatology), psychosocial outcomes throughout the lifespan, and contributing environmental and dispositional risk factors [see (3) for a comprehensive overview]. Regarding the latter, accumulating evidence suggests that deficits in different emotion processing domains, such as emotion recognition (e.g., difficulties in identifying facial expressions), emotion learning (e.g., difficulties in learning from punishment), and emotion regulation (e.g., difficulties in inhibiting impulsive responses to emotional cues), may offer a particularly powerful basis for explaining potentially different presentations and trajectories of CD behaviors, including aggression (4–8). For instance, deficits in the recognition of distress cues, such as emotional expressions of fear, sadness or pain, but also of other facial expressions, such as happiness, appear to be most pronounced in individuals with CD who have high levels of callous-unemotional (CU) traits (i.e., reduced guilt and empathy, callousness, and uncaring attitudes) (9). Individuals with CD who present with at least two of these CU traits fulfill criteria for the Limited Prosocial Emotions (LPE) specifier in DSM-5 (1). This subtype of CD is considered particularly severe as affected individuals typically present with an earlier age-of-onset and a more serious and stable set of symptoms, including proactive aggression, placing them at increased risk for poor treatment outcomes (10) [but see also (11, 12)] and for developing mental health problems in adulthood, such as antisocial personality disorder (ASPD) (6, 13, 14). In contrast, individuals with CD but without the LPE specifier (i.e., those showing subclinical levels of CU traits) are thought to show emotion regulation deficits, such as an inability to maintain behavioral control when confronted with acute

emotional stimuli (e.g., visual threats), which may contribute to impulsive acts of reactive aggression and an increased risk for anxiety and depression (6). Finally, emotional learning deficits, such as a failure to learn how to avoid choices that lead to punishment rather than reward, occur more broadly in youths with CD irrespective of their LPE status (15).

However, most prior work on emotion functioning in CD, and its clinically defined subtypes, has been limited by relying on relatively small samples with varying selection criteria and neurocognitive tasks (16), including mixed samples of youths with CD or oppositional defiant disorder (ODD), or focusing on a single subdomain of emotion dysfunction instead of all three domains linked to CD, including emotion recognition, learning, and regulation (4, 5, 7). Thus, studies to date have largely been unsuited or underpowered for testing within-CD, individual variability of the underlying neurocognitive disease mechanism(s), including emotion dysfunction. To address the above-mentioned research gaps, we initiated the largest study to date to comprehensively investigate emotion recognition, emotion learning, and emotion regulation using a broad neuropsychological test battery within a single sample of youths with CD ($n = 542$) compared to typical controls ($n = 710$) (17). As traditionally done in this line of research, we first compared the group of youths diagnosed with CD with the typical controls, and based on statistically significant group differences or the lack thereof, we determined whether a CD-related neurocognitive deficit was present or not. As expected, we found that emotion deficits in the CD group spanned across the three neurocognitive domains. However, we also noted that the significant group differences between CD and controls in task performance had effect sizes in the small to very small range (i.e., Cohen's d s < 0.29). As this is in line with previous meta-analytic findings (18, 19), these results indeed suggest: (i) substantial distributional overlap *between* the CD and non-CD samples in terms of performance on emotion processing tasks; and (ii) substantial variation in emotion processing abilities *within* the CD population. Hence, it is reasonable to assume that the significant CD-vs.-control effects for emotion processing tasks—or any other neuropsychological measure reported in the literature [e.g., (20)], may either be truly small effects driven by the entire CD sample or, which appears more likely, they are driven by only a subset of youths with CD who have emotion processing deficits (21).

In fact, this notion of diversity in emotion processing is emphasized by current neurocognitive models of CD etiology [e.g., (6, 7)]. These models suggest that dysfunction in distinct emotion processing domains are associated with different subtypes of CD and related symptom sets (22). For instance, given the assumption that youths with CD with the LPE specifier show difficulties in perceiving other people's emotions, particularly distress and happiness, one might predict that performance in this neurocognitive domain would be disproportionately deficient in this subgroup, whereas the subgroup of youths with CD but without the LPE specifier would show specific difficulties with emotion regulation (4, 6).

In addition, emotion dysfunction might serve as an “intermediate phenotype”—i.e., developmental neurocognitive mechanism (23)—linking risk for psychopathology with the emergence of clinical symptomatology, including clinical subtypes of CD (24). There are multiple dispositional and contextual risk factors that have repeatedly been implicated in CD, such as birth complications, maladaptive parenting, or low socioeconomic status [reviewed in (3)]. Data from epidemiological and at-risk samples suggests that particular risk factors appear to have closer associations with a specific domain of emotion dysfunction (25). For instance, children exposed to physical violence or abuse exhibit altered emotion recognition processes (26), including an altered ability to identify and discriminate specific emotions (i.e., anger) contributing to ‘hostile attribution biases’ (i.e., misinterpreting neutral or ambiguous facial expressions as threatening), which, in turn, predict the emergence of CD behaviors, such as aggression (27, 28). In slight contrast, exposure to numerous adversities, such as poverty (incl. low socioeconomic status), deprivation (incl. institutional rearing), maltreatment, or pre- and perinatal influences (incl. maternal smoking during pregnancy, or birth complications), appears to be related to emotion regulation and learning difficulties predicting the onset of both externalizing and internalizing problems (29). Although the literature is far from being conclusive in linking adversity factors with specific neurocognitive processes as intermediate phenotypes of conduct problems, studying neuropsychologically defined subtypes of emotion dysfunction in CD may provide novel insights into mechanisms that presumably underlie the complex developmental pathways from risk for psychopathology to different clinical expressions of the disorder (30).

Thus, the primary aim of the current study was to adopt a clinically motivated, person-centered (rather than variable-centered) bottom-up analytic approach to explore the neurocognitive diversity of emotion functioning in CD. We accomplished this by re-analyzing the neuropsychological task performance data from our large sample of girls and boys with CD who were comprehensively clinically assessed and reliably diagnosed using standardized, semi-structured interviews (17, 31). For each of the three emotion processing tasks that assessed emotion recognition, emotion learning, and emotion regulation, respectively, we defined deficit as task performance within the bottom 10% of an age-matched control group (equivalent to approximately 1.3 standard deviations below the mean), following the common-metric approach usually

applied in pediatric neuropsychology and as previously used in neurocognitive studies in ADHD (21, 32–35).

To our knowledge, this study is the first to investigate neuropsychological subgroups within the CD population by exploring the proportion of youths with CD who do vs. those who do not have deficits in emotion processing, including emotion recognition, emotion learning, and emotion regulation (but see (36, 37) for similar approaches focusing on other neurocognitive domains, such as verbal skills, mental flexibility, or memory, in smaller-scale studies). Our sample is particularly suited to investigate neurocognitive diversity within CD as it is one of the largest, most representative and clinically well-characterized cohorts of girls and boys with CD (vs. typical controls) recruited from a variety of sources, including the community, specialist schools, mental health clinics, welfare institutions, and youth offending services in different European countries (31). Given the magnitude of the effect sizes observed in our previous study (17), we expected to find subgroups of youths with CD without deficit in any domain vs. those who have deficits in only one domain, two domains, or across all three domains. Most importantly, we tested the extent to which neuropsychologically defined subgroups would map clinically onto the CD subtypes described in the DSM-5, including CD with vs. without the LPE specifier, and as a secondary aim, we explored whether the neuropsychological subgroups would be associated with specific CD-related risk factors.

MATERIALS AND METHODS

Participants

As part of the European multi-site project entitled “Neurobiology and Treatment of Adolescent Female Conduct Disorder: The Central Role of Emotion Processing” (FemNAT-CD; <https://cordis.europa.eu/project/id/602407/reporting>), we reanalyzed the neuropsychological data obtained from our large sample of youths with CD ($n = 542$, 317 girls) and TDCs ($n = 710$, 479 girls), aged 9–18 years (see (17) for details on recruitment, clinical assessments, and sample characteristics). In brief, we used data from participants who provided a complete neuropsychological dataset which included facial emotion recognition (*Emotion Hexagon task*), emotion learning (*Passive Avoidance Learning task*), and emotion regulation skills (*Emotional Go/Nogo task*). Participants were recruited through community outreach (e.g., mainstream schools) as well as from mental health clinics, welfare institutions, and youth offending services at 10 sites across Europe (**Supplementary Table 1**) (31). Overall exclusion criteria were IQ < 70, autism spectrum disorders, schizophrenia, bipolar disorder or mania, neurological disorders, and genetic syndromes. Individuals with CD were diagnosed according to DSM-IV-TR criteria (38). Youths with “only” ODD who did not fulfill the diagnostic criteria for CD were excluded from the current analysis. TDCs were free of current psychiatric diagnoses and lifetime diagnoses of CD, ODD, and ADHD. We excluded TDCs with lifetime histories of and/or current disruptive behavior disorders, such as ADHD, ODD, and CD, in order to rule out the influence of any subclinical or precursor symptoms that are potentially linked to CD. Written informed consent

TABLE 1 | Demographic and clinical characteristics.

	CD <i>n</i> =542	TDC <i>n</i> =710	Group effect <i>p</i> -values ^a
Age (years) <i>M</i> (SD)	14.4 (2.3)	14.0 (2.5)	0.001
Females (%)	58.5	67.5	0.001
Estimated IQ <i>M</i> (SD)	94.9 (12.4)	103.5 (12.2)	<0.001
SES <i>M</i> (SD)	−0.29 (0.93)	0.28 (1.03)	<0.001
CD total symptoms <i>M</i> (SD)	5.45 (2.34)	0.05 (0.23)	<0.001
Average age-of-onset of CD (years) <i>M</i> (SD)	10.3 (3.8)	N/A	
CD age-of-onset subtype (%)			
Childhood	43.0	N/A	
Adolescence	53.3	N/A	
Unspecified	3.7	N/A	
CD severity (%)			
Mild	24.3	N/A	
Moderate	52.8	N/A	
Severe	20.5	N/A	
Unknown	2.4	N/A	
Impairment caused by current CD (%)			
With peers	63.7	N/A	
With family	85.9	N/A	
With school	78.9	N/A	
Unknown	1.2	N/A	
LPE specifier (%)	43.7	18.3	<0.001
Current comorbidities <i>n</i> (%)			
ODD	78.2	N/A	
ADHD	38.4	N/A	
SUD	17.4	N/A	
MDD	14.8	N/A	
PTSD	6.7	N/A	
GAD	3.0	N/A	
Psychotropic meds (%)	30.2	N/A	

Diagnoses and CD symptoms were based on the Schedule for Affective Disorders and Schizophrenia for School-Age Children–Present and Lifetime version (K-SADS-PL). For TDC, any current psychiatric diagnosis as well as a history of ADHD, ODD, or CD was exclusionary. ADHD, attention deficit hyperactivity disorder; CD, conduct disorder; GAD, generalized anxiety disorder; IQ, estimated intelligence quotient; LPE, limited prosocial emotions specifier [see (39)]; MDD, major depressive disorder; Meds, on psychotropic medications; ODD, oppositional defiant disorder; PTSD, post-traumatic stress disorder; SES, socioeconomic status (SES was based on parental income, education level, and occupation (40)); SUD, substance use disorder (including substance abuse and dependence); TDC, typically developing controls.

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

was obtained for all participants, and local ethics committees approved the study protocol. **Table 1** summarizes the sample's main demographic and clinical characteristics.

All individuals were clinically assessed with the Kiddie-Schedule for Affective Disorders and Schizophrenia–Present and Lifetime version [K-SADS-PL (41)]. The K-SADS-PL is a semi-structured clinical interview that is administered separately to caregivers and participants by trained staff members to assess

current and lifetime psychiatric diagnoses, disorder severity, and age-of-onset and duration of a disorder. Additionally, where available, information from medical or case files was used. Summary ratings were derived from the clinical judgment using all sources. The items of the K-SADS-PL are scored on a scale from 0 to 3. A rating of 0 indicates no (insufficient) information, a score of 1 indicates a given symptom is not present, 2 indicates a subclinical expression, while a score of 3 is given when a symptom is present and clinically significant. Scores were recoded, so that a clinical rating of “not present” is represented by 0, a subclinical rating by a score of 1, and a clinically significant rating by a score of 2. Inter-rater reliability (IRR; *N* = 75, i.e., *n* = 5–8 per site) of CD was high (Cohen's κ = 0.91), with an agreement rate of 94.7%. IRR of other disorders, including ADHD, ODD, major depressive disorder (MDD), and generalized anxiety disorder (GAD), was also high (Cohen's κ s $\geq$ 0.84, agreement rates $\geq$ 92%), which is in line with the reliability data reported by Kaufman et al. (41). These authors also report data which support the concurrent validity of the diagnoses generated with this instrument. Youths who met criteria for a specific disorder (e.g., behavior disorder) scored significantly higher than undiagnosed youths on rating scales assessing related symptom sets. Using the K-SADS-PL, we also determined the CD-onset type [i.e., childhood-onset (CO-CD): presence of at least one CD symptom and impairment prior to age 10; adolescent-onset (AO-CD): CD symptoms only emerge after age 10] (1).

Full-scale IQs were estimated using the vocabulary and matrix reasoning subtests of the Wechsler Intelligence Scale for Children–Fourth Edition (42), the Wechsler Adult Intelligence Scale–Fourth Edition (43), or the Wechsler Abbreviated Scale of Intelligence (44). The vocabulary subtest consists of 31 items, and youths are required to verbally define and/or describe a word or concept that is orally presented to them. Each item is scored on a 0-, 1-, or 2-point basis according to the manual. In the matrix reasoning subtest 30 visually-depicted incomplete matrices are presented, and youths are required choose one item from a selection of five options that correctly completes the matrix. Each correct item receives 1 point. The *T* and standard scores for each subtest were transformed into *z*-scores and then combined to yield estimates of full-scale IQ. For the two-subtest short form (FSIQ-2) internal consistency and test-retest reliability were reported to be excellent (>0.90). Estimated full-scale IQ scores were highly correlated with scores on tests purported to measure similar constructs. Correlations between the short form and the original tests were reported to be acceptable (0.71) to excellent (0.92) (45).

CU traits scores were derived from the Youth Psychopathic traits Inventory (YPI) (46). The YPI is a 50-item self-report measure of psychopathic traits. Each item is answered on a 4-point Likert scale ranging from “does not apply at all” (1) to “applies very well” (4). Higher scores indicate higher levels of psychopathy. CU traits scores were calculated using the total score for the subscales ‘remorselessness’ (e.g., “To feel guilt and regret when you have done something wrong is a waste of time”), ‘unemotionality’ (e.g., “I usually feel calm when other people are

scared”), and ‘callousness’ (e.g., “*I think that crying is a sign of weakness, even if no one sees you*”). The CU traits dimension showed good internal consistency (Cronbach’s $\alpha = 0.81$). Test-retest reliability of the YPI over a 6-month period was reported to be adequate ($ICC = 0.76$) (47). Convergent and divergent validity was supported in a sample of 360 youths from the general population (e.g., CU traits scores correlated positively with narcissism, but negatively with empathy measures). We also used the three CU traits subscales of the YPI to create a proxy for the LPE specifier, following the procedure developed by Colins and Vermeiren (39). A participant was considered to meet criteria for one of the CU traits when she/he reported that at least one item on the corresponding subscale applied “very well” to her/him [i.e., a score of 4 on a 4-point Likert scale, ranging from “Does not apply at all” (1) to “Applies very well” (4)]. Participants were considered to meet criteria for the LPE specifier if two or more CU traits were endorsed to threshold.

Participants reported on their own aggressive behaviors using the Reactive-Proactive aggression Questionnaire (RPQ) (48), which includes 11 items related to ‘reactive aggression’ (e.g., “*I have damaged things because I felt mad*”), and 12 items related to ‘proactive aggression’ (e.g., “*I have had fights to show that I was on top*”). Each item is rated on a 3-point Likert scale ranging from 0 (“never”) to 2 (“often”). The proactive and reactive aggression scales are sum scores of the respective items. Internal consistency for the two subscales was good (Cronbach’s $\alpha = 0.75$ and 0.88 , respectively). Raine et al. (48) also report data which support the validity of the two subscales.

In addition to gender/sex (i.e., male) and general cognitive abilities (i.e., low IQ), numerous longitudinal studies (e.g., (49, 50)) have identified several risk factors linked to CD (reviewed in (3)) of which the following were assessed in the present study: maternal smoking during pregnancy, parental maladaptive behavior (i.e., repeated delinquency of mother/father), and socioeconomic status [i.e., SES (40)]. These three variables were extracted from the Medical History Questionnaire which is a semi-structured interview for parents/caregivers specifically designed for this study with items included based on evidence about CD-related risk factors derived from epidemiological studies (31). Additionally, childhood exposure to parental violence/abuse/neglect, and deviant peer affiliations were evaluated with the Childhood Experience of Care and Abuse Questionnaire (CECA-Q, i.e., total sum of the subscale scores for “antipathy mother/father,” “neglect mother/father,” and “physical abuse mother/father,” Cronbach’s $\alpha = 0.78$) (51), and the Social and Health Assessment (SAHA, i.e., “affiliation with delinquent peers” subscale) (52), respectively. The CECA-Q is a self-report questionnaire to assess lack of parental care (neglect and antipathy), parental physical abuse, and sexual abuse from any adult (not used in this study) before age 17. Satisfactory reliability and validity have been reported by Bifulco et al. (51). The SAHA subscale consists of nine items, and youths are asked about how many of their close friends are involved in different types of risk-taking and delinquent behavior (“None”; “A few”; “Some”; or “Most or all”): e.g., dropping out of school, smoking cigarettes, drinking alcohol, or using marijuana. The summed score could range from 9 to 36

where higher scores indicate greater association with delinquent peers. The internal consistency was high (Cronbach’s $\alpha = 0.91$), and published data support the validity of this subscale (53).

Neuropsychological Test Battery

We used the *Emotion Hexagon task* to assess the accuracy of facial emotion recognition (54). Participants were asked to label morphed facial expressions as either happy, sad, angry, fearful, disgusted, or surprised (i.e., the six “basic” emotions). Morphs were created from six expression pairings: happy-surprised, surprised-fearful, fearful-sad, sad-disgusted, disgusted-angry, and angry-happy. Each pair included two prototype expressions in proportions 90:10, 70:30, 50:50, 30:70, and 10:90 (i.e., 10% happy and 90% surprised for the happy-surprised continuum). Morphed expressions were presented individually and randomly on a computer monitor for a maximum of 3 s, and participants were asked to select by mouse-click one of the six emotion labels that best described the expression shown. Participants were given as long as necessary to make their selection and were not given feedback about their performance accuracy. Participants completed one practice block, followed by five blocks that each displayed all 30 morphed expressions once (6 pairs x 5 morphs). The total score for incorrect recognition per expression ranged from 20 (100% error rate) to 0 (0% error rate), with 50:50 morphs not being scored or analyzed.

We administered a modified *Passive Avoidance Learning task* to assess the accuracy of emotional learning (55). The task involves assigning reward and punishment values to novel stimuli (“ziggerins” (56)). Novel stimuli were chosen to tap into pure learning effects without the bias of stimulus familiarity. Participants were instructed to learn by trial-and-error to respond through button press to four different reward stimuli (gaining 1, 700, 1,400, or 2,000 points, respectively; non-responses were counted as omission errors in %) and to avoid responding to four different punishment stimuli (losing 1, 700, 1,400, or 2,000 points, respectively; responses to these stimuli were counted as avoidance errors in %). Each stimulus was shown once within a block of 8 trials, with 10 blocks overall (including one practice block). Stimuli were displayed on a computer monitor for a maximum of 3 s, followed by performance feedback (i.e., amount of points won, or lost, as well as the running total points). Participants started the task with 10,000 points.

We administered the *Emotional Go/Nogo task* to assess the accuracy of emotion regulation defined as the ability to maintain cognitive control when confronted with interfering emotional information, including positive and negative facial expressions (57, 58). Participants were instructed to press a response button as quickly and accurately as possible whenever a named facial expression appeared on the screen (go trials) and not to press for any other expression (nogo trials). The task included six randomly presented blocks of go-nogo pairings: neutral-happy, neutral-fearful, happy-neutral, fearful-neutral, happy-fearful, and fearful-happy. Each block included 35 go (73%) and 13 nogo (27%) stimuli. The go trials occurred more frequently in order to create a pre-potent tendency for the participant to respond. Stimuli consisted of gray-scaled fearful, happy, and

neutral expressions from six male and six female adults, with four African-American, Asian, and Caucasian individuals for each expression type, respectively (NimStim set numbers: 6, 8, 11, 14, 15, 16, 27, 36, 39, 43, 44, and 45). Stimulus duration was 500 ms with 1 s interstimulus intervals. False alarm error rates in % for nogo trials indexed emotion regulation, with higher rates reflecting worse performance (59).

Order of tasks was pseudorandomized separately across group (CD, TDC), sex (female, male), and age brackets (9–12, 13–15, and 16–18 years). The extracted performance variables of the three tasks (see **Table 1**) had acceptable to good reliabilities (Cronbach's $\alpha \geq 0.70$). Details on the test battery and procedures are provided in **Figure 1** [for more details, see also (17)]. We chose this particular test battery based on influential models of emotion dysfunction in CD (see (17) for details), and because the three tasks have widely been used in neuropsychological research of emotion functioning in developmental psychopathology, including CD, ADHD, and internalizing disorders. Thus, the validity of the test battery comes from its proven usefulness to distinguish between clinical groups and controls in previous research (4). Available psychometric data further support both reliability and validity of all three neuropsychological measures (59–61). Standard operating procedures (SOP) ensured consistency of data collection, handling, and analysis across all data collecting sites.

Statistical Analyses

First, all raw scores for the performance variables of interest from our neuropsychological battery (see above, and **Table 2** below) were age-, IQ-, and sex-adjusted using standard regression procedures, resulting in *z*-scores as the dependent variables in the following factor analysis (conducted in SPSS v25, IBM Corp., Armonk, NY). We ran a confirmatory principal component (PCA) “factor” analysis, using varimax rotation with Kaiser normalization, as we had *a priori* expectations about the number of factors that would be associated with the measured dependent variables (62), i.e., separate factors for emotion recognition, emotion learning, and emotion regulation, respectively. All participants were included in the PCA in order to maximize statistical power as well as to create a common metric of performance scores by which youths with CD and TDCs could be compared. Per component, factor scores were extracted for each participant using the Anderson-Rubin method ($M = 0$, $SD = 1$) to avoid multicollinearity. We then explored case-control differences for the three emotion domains within a repeated-measures analysis of variance model using the factor scores as the dependent variables. Effect sizes were calculated using partial eta squared (η^2_p), where 0.01, 0.06, and 0.14 represent small, medium and large effects, respectively (63). We additionally report 95% confidence intervals (CIs) for all effect size measures. Finally, and most importantly, in keeping with the clinically motivated subgrouping approach that has been very informative in the ADHD field (21, 32–35), we subdivided the individuals with CD into subgroups on the basis of their neuropsychological task performance. For each emotion processing domain identified in the factor analysis, participants were classified as “deficient” or “intact,” with deficit being defined as performance (i.e., factor scores) within the bottom 10% of

an age-matched control group using the following age brackets to recognize that test performance typically improves with age (64): 9–12, 13–15, and 16–18 years. We then investigated the proportion of youths with CD who had one or multiple deficits vs. no deficit at all. We had three reasons for choosing the 10th percentile criterion as a reasonable cutoff for deficit: First, we wanted to use a clinically useful definition of performance deficit that is established in the field of clinically oriented neuropsychological assessments as this is the core interest of person-centered approaches; second, to apply a threshold which is easily reproducible and comparable between studies within and across disorders (in contrast, for instance, to more elaborate analytic approaches based on machine learning algorithms which usually require specialized expertise that is not available in all research labs or clinics (65)); and third, any other reasonable cutoff (e.g., 5th percentile) would still classify many youths with CD as “intact” or conversely classify a substantial number of typically developing controls (TDC) as “deficient” (i.e., the positive predictive value increases, but the negative predictive value decreases with more stringent cutoffs) (21). Regression analyses, chi-square tests, and *t*-tests were then used to explore differences in clinical correlates and risk factors between the neuropsychologically defined CD subgroups.

RESULTS

Correlational and Principal Component Analyses

As expected, correlations between the dependent variables (i.e., age-, IQ-, and sex-adjusted performance scores) were larger within each of the three emotion domains (mean $r_{\text{Olkin \& Pratt}} = 0.37$, 95% CI: 0.34, 0.44) than between the domains (mean $r_{\text{Olkin \& Pratt}} = 0.12$, 95% CI: 0.07, 0.17; Fisher's $z = 7.28$, $p < 0.001$), indicating that our test battery did indeed capture emotion processing as a multifaceted construct rather than a unitary one. For the PCA, the Kaiser-Meyer-Olkin (*KMO*) measure confirmed that the sample size was adequate for the analysis [$KMO = 0.86$, which is a sufficiently high value (66)], and the *KMO* values for all dependent variables were substantially higher than 0.5 (i.e., $KMOs \geq 0.68$). Bartlett's Test of Sphericity indicated that correlations between variables were large enough for PCA [$\chi^2(190) = 7514.62$, $p < 0.001$]. The dependent variables from the three tasks loaded onto the predicted three components, accounting for 48.1% of the total explained variance in performance scores. Components 1, 2, and 3 represented emotion regulation (17.4% variance), emotion learning (17.2% variance), and emotion recognition (13.5% variance), respectively (**Table 1**). Note: Using other factor analytic procedures, such as the Maximum Likelihood method or Principal Axis Factoring, yielded similar results which is in line with prior work (67) (data available on request).

Comparing Dimensionally the CD and TDC Groups in Factor Scores

We analyzed the factor scores for the three emotion processing domains using a three (domain: emotion regulation vs. emotion learning vs. emotion recognition) by two (group: CD vs. TDC)

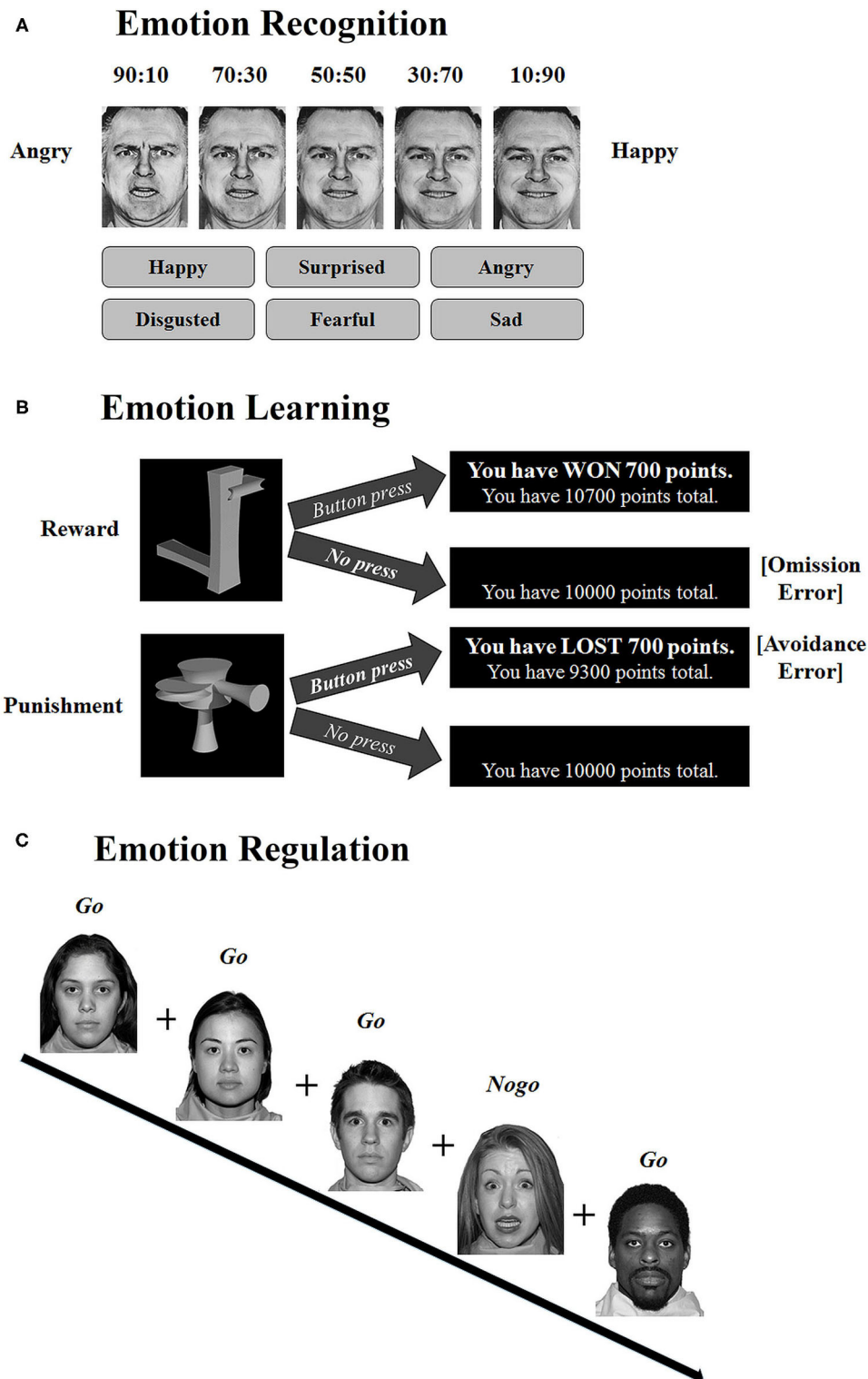


FIGURE 1 | Illustration of the model based neuropsychological test battery used to assess **(A)** emotion recognition, **(B)** emotion learning, and **(C)** emotion regulation, respectively. **(A)** As an example, the angry-happy facial expression continuum from the *Emotion Hexagon* task is depicted, including the five different morphs from this continuum as well as the six emotion labels used in the task. Only one facial expression is displayed in each trial. **(B)** Examples from the *Passive Avoidance Learning* task, depicting one stimulus associated with reward (e.g., gaining 700 points by button press), and one stimulus associated with punishment (e.g., losing 700 points by button press). **(C)** Example layout of the emotion regulation condition from the *Emotional Go/Nogo* task, including neutral expressions as the “Go” targets and fearful expressions as the “Nogo” non-targets. This Figure was republished from Kohls et al. (17) with permission from ELSEVIER.

TABLE 2 | Results from the confirmatory principal component analysis.

Variables	Rotated “factor” loadings		
	Emotion regulation	Emotion learning	Emotion recognition
False alarm error rate in % (Go/Gogo pairings)			
Neutral/Happy	0.783		
Neutral/Fearful	0.764		
Fearful/Happy	0.746		
Happy/Neutral	0.744		
Fearful/Neutral	0.726		
Happy/Fearful	0.723		
Error rate in % (Punishment, and Reward conditions)			
Losing 1,400 points		0.724	
Losing 700 points		0.712	
Losing 1 point		0.703	
Losing 2,000 points		0.670	
Gaining 1 point		−0.686	
Gaining 1,400 points		−0.659	
Gaining 700 points		−0.535	
Gaining 2,000 points		−0.476	
Error rate in % (Emotion expression)			
Happiness			0.718
Surprise			0.704
Fear			0.667
Sadness			0.632
Disgust			0.605
Anger			0.590
Eigenvalues	3.48	3.44	2.71

repeated-measures analysis of variance model, followed by *post-hoc* pairwise comparisons in cases where significant main or interaction effects emerged (Bonferroni-corrected for multiple comparisons). This analysis revealed a small but significant group by domain interaction effect [$F_{(2,500)} = 3.34$, $p = 0.036$, $\eta^2_p = 0.003$, 95% CI: 0.0001, 0.008], and a significant main effect of group with a medium effect size [$F_{(1,250)} = 72.98$, $p < 0.001$, $\eta^2_p = 0.055$, 95% CI: 0.033, 0.081]. Compared to TDCs, youths with CD as a group performed worse in all emotion domains (in line with our prior analyses examining task performance separately for each neuropsychological task (17)), but the largest case-control differences were found for the emotion recognition domain ($\eta^2_p = 0.035$, 95% CI: 0.018, 0.058), followed by the emotion regulation domain ($\eta^2_p = 0.018$, 95% CI: 0.006, 0.035), with the smallest difference observed for the emotion learning domain ($\eta^2_p = 0.007$, 95% CI: 0.0009, 0.019) (Figure 2). We also explored the extent to which group membership (i.e., CD or TDC) was predicted by the three emotion domain scores using a logistic regression analysis. This analysis revealed significant effects of all three emotion domains (*Wald* χ^2 s > 9.43 , $ps \leq 0.002$). However, the final model [$\chi^2(3) = 77.78$,

$p < 0.001$; Hosmer-Lemeshow-Test: *ns*] successfully predicted group membership of 83% (95% CI: 79.8, 85.5%) of the TDCs, i.e., relatively high specificity, but only 36% (95% CI: 31.8, 40.0%) of the CD cases, i.e., relatively low sensitivity, again pointing to substantial variability in emotion processing skills within the CD group.

Neuropsychologically Based Subgroup Analysis

Proportion of Deficit

Figure 3 presents a Venn diagram showing the proportion of CD cases who exceeded the threshold for one or multiple deficits in the emotion recognition, emotion learning, and/or emotion regulation domains, with deficit being defined as performance (i.e., factor scores) within the bottom 10% of their respective age-matched control group (which equals ≤ -1.3 standard deviations from the mean of performance data by TDCs). In the CD group, 43.7% of the participants were deficient in at least one domain of emotion processing. Deficits in emotion recognition were the most common deficit, followed by emotion regulation, and emotion learning was the least common deficit. Overlap between the different deficits was rare, with only $\sim 1\%$ of the youths with CD displaying a pervasive deficit across all three domains. Compared to TDCs, “deficit” was significantly more frequent among the youths with CD for the emotion recognition and emotion regulation domains ($ps < 0.001$), but not for the emotion learning domain ($p = 0.13$). Notably, a substantial subgroup of CD youths (56.3%) showed no deficit in their emotion processing abilities across the three domains. Among the TDCs, $\sim 27\%$ showed at least one emotion deficit (i.e., 24.6% qualified for only one, 2.4% for two, and 0.1% for three deficits), leaving $\sim 73\%$ of the TDCs who performed normally across the three emotion domains (Note: If the three neuropsychological tasks were completely independent, one would expect to find $\geq 30\%$ of TDCs to be deficient in at least one emotion domain given the 10th percentile criterion as our cutoff for deficit per task).

Clinical Correlates

Introducing the three emotion deficit domains (i.e., deficit: yes = “1,” or no = “0”) as predictors into multiple linear or logistic regression models—i.e., running one model with the three predictors for each dependent variable separately—did not reveal any significant associations with the main clinical variables of interest among the youths with CD [i.e., CD symptom severity, CD age-of-onset subtype, CU traits/LPE specifier, or the presence of major comorbidities, including ADHD, ODD, major depressive disorder (MDD), generalized anxiety disorder (GAD), post-traumatic stress disorder (PTSD), and substance use disorder (SUD)]. More specifically, the subgroup of youths with CD with an emotion recognition deficit did not differ significantly from those without such deficit regarding: (i) the presence of the LPE specifier (44.1% vs. 43.6%, $p = 0.92$), (ii) CD age-of-onset subtype (CO-CD/AO-CD: 43.3%/52.8% vs. 42.9%/53.5%, $p = 0.98$), (iii) number of CD symptoms (K-SADS-PL CD symptom count: 5.6 ± 2.4 vs. 5.4 ± 2.3 , $p = 0.57$); or (iv) proactive aggression (RPQ proactive aggression

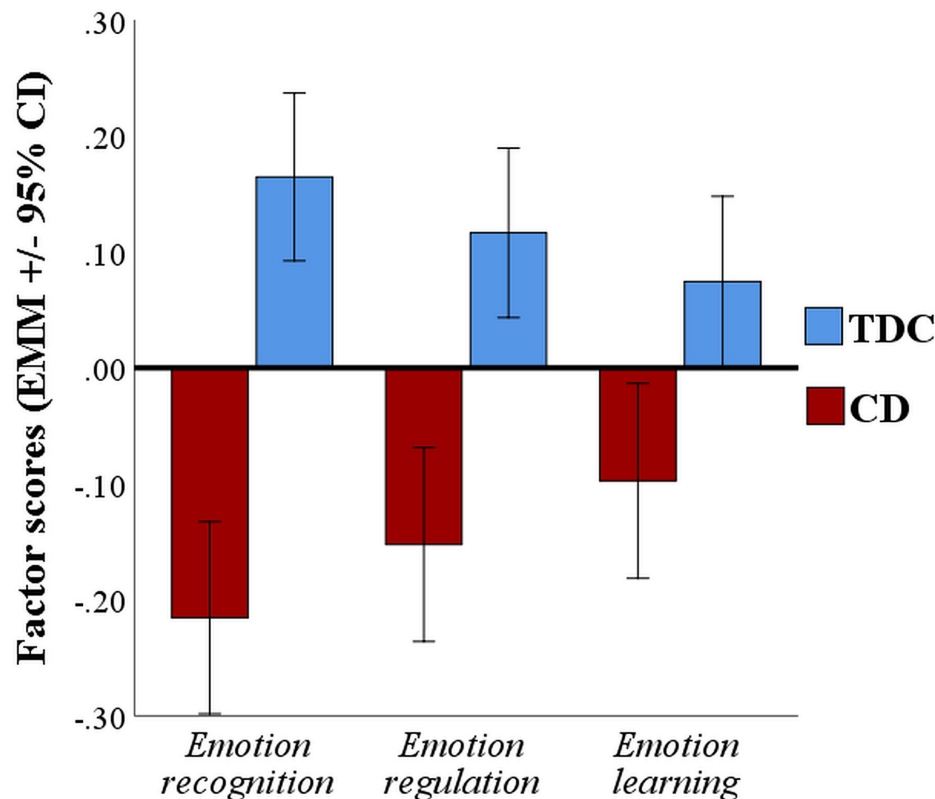


FIGURE 2 | Factor scores for youths with CD vs. TDCs for the three emotion processing domains investigated in this study. Relative to TDCs, youths with CD showed the expected deficits in all three domains, with the greatest deficits in the emotion recognition domain, followed by the emotion regulation and emotion learning domains (EMM, Estimated Marginal Means; 95% CI = error bars represent 95% Confidence Intervals).

subscale score: 5.0 ± 5.2 vs. 4.7 ± 4.6 , $p = 0.56$). Moreover, the subgroup of youths with CD with an emotion regulation deficit did not differ significantly from those without such deficit in terms of: (i) the LPE specifier (38.8% vs. 44.8%, $p = 0.31$), (ii) the presence of comorbid anxiety disorders (16.2% vs. 11.7%, $p = 0.24$), or (iii) scores on the RPQ reactive aggression subscale (12.0 ± 4.7 vs. 12.1 ± 5.0 , $p = 0.86$). Taken together, these findings suggest that deficits in emotion processing related to CD (4) do not map neatly onto established DSM-5 subtypes, such as CD+LPE or childhood-onset CD.

Associated Risk Factors

As a secondary aim, we explored associations between the three emotion processing domains and established risk factors for CD within three logistic regression models, i.e., one for each neurocognitive domain (deficit: yes = “1,” or no = “0”) using the following predictors: gender/sex, general cognitive abilities (i.e., low IQ), maternal smoking during pregnancy, socioeconomic status (i.e., low SES), parental maladaptive behavior (i.e., delinquency), childhood exposure to parental violence/abuse/neglect, and deviant peer affiliations. For emotion recognition, boys with CD were 2.1 times more likely to have a deficit than girls with CD ($\text{Wald } \chi^2 = 4.19$, $p = 0.04$), and

a lower IQ among the youths with CD was also significantly associated with an increased likelihood of exhibiting a deficit in this domain ($\text{Wald } \chi^2 = 6.10$, $p = 0.013$). For emotional learning, youths with CD whose mothers smoked during pregnancy were 3.2 times more likely to have a deficit in this domain than youths with CD whose mothers did not smoke during pregnancy ($\text{Wald } \chi^2 = 5.37$, $p = 0.021$). None of the assessed risk factors predicted an emotion regulation deficit. Please note that none of the risk factors were associated with task performance in the TDCs.

DISCUSSION

The primary aim of this study was to investigate neuropsychological subgroups across three emotion processing domains linked to CD and test whether specific subgroups would map onto existing DSM-5 defined CD subtypes, including CD with vs. without the LPE specifier (4). Our dimensional analysis showed that, relative to TDCs, the CD group performed worse in all three emotion domains, but the largest case-control difference was found for the emotion recognition domain, followed by the emotion regulation domain, with the smallest difference observed for the emotion learning domain; this supports the

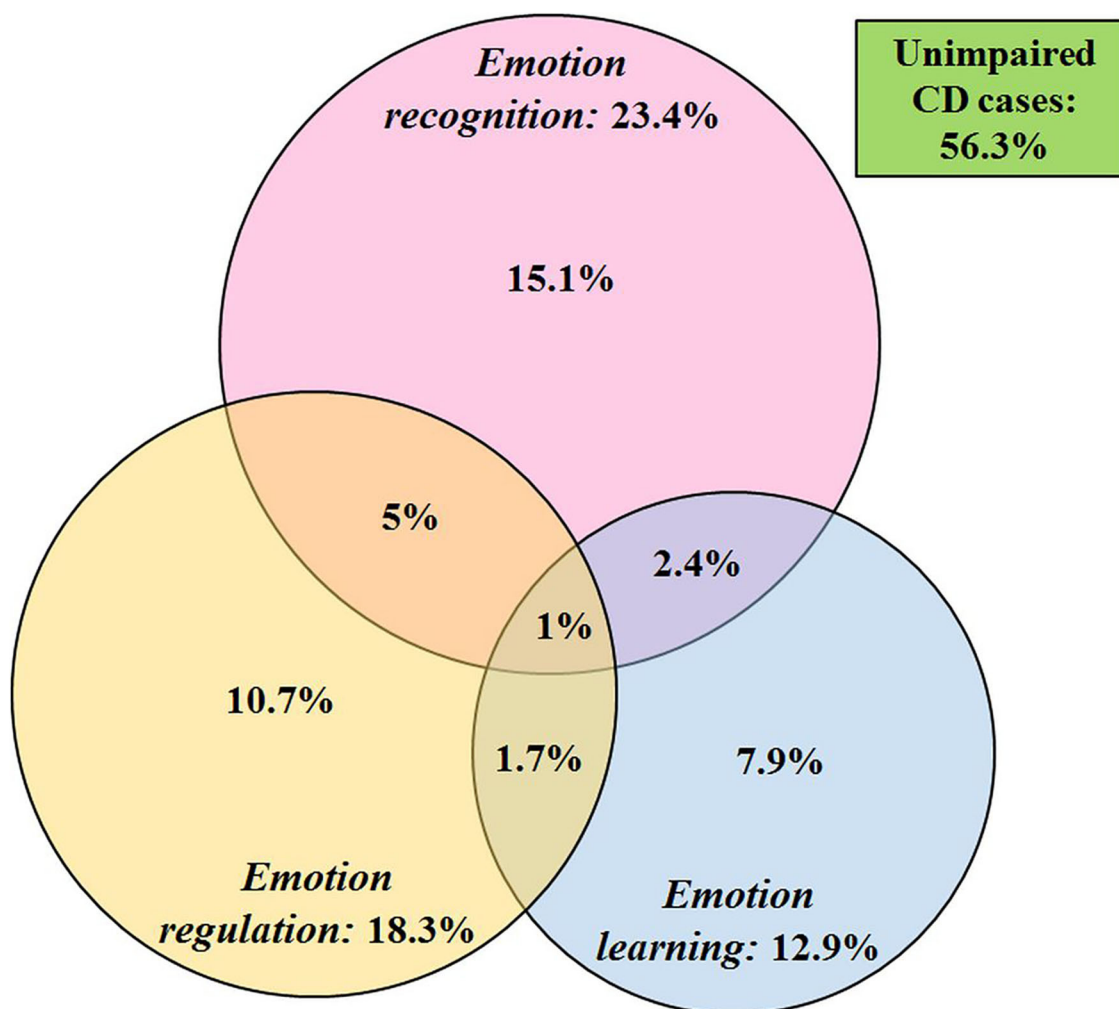


FIGURE 3 | Proportion of CD cases ($n = 542$) with deficits in emotion recognition, regulation, and learning, and their degree of overlap. The numbers reported next to the domain labels reflect the sum of those with a deficit in that particular domain, including those with single, or multiple deficit(s).

findings of our previous report (17). Our categorical subgroup analysis substantiated this finding by revealing that deficient emotion recognition skills were the most common deficit in CD (~23%), followed by emotion regulation (~18%), with emotion learning the least common deficit (~13%). Critically, though, we also found that the majority of youths with CD (~56%) did not demonstrate meaningful deficits in any of the three emotion domains, and only a very small proportion of the CD sample showed pervasive deficits across all domains (i.e., deficits in all three emotion functions: ~1%). Overall, emotion processing deficits, if existent in youths with CD, appear to be unrelated to main phenotypic characteristics, such as age-of-onset of CD, symptom severity, or co-occurring psychiatric disorders. Moreover, contrary to our assumptions, the present data do not support notions that established clinical subtypes of CD, most importantly CD with vs. without the LPE specifier (1), are tightly linked to neurocognitive deficits in one particular emotion domain over another (i.e., emotion recognition deficits

in CD+LPE vs. emotion regulation deficits in CD-LPE) (4, 6). Notably, being male and having a lower IQ increased the likelihood for showing a deficit in the emotion recognition domain, whereas maternal smoking during pregnancy increased the likelihood for having an emotion learning deficit. Please note, though, that our analytic approach to subdivide the CD sample into subgroups based on their neurocognitive performance (i.e., emotion domain deficit: “yes” or “no”) likely attenuated statistical power to detect associations with clinical correlates and associated risk factors.

Although the present findings support recent theoretical models emphasizing that CD is a remarkably heterogeneous condition, with different individuals being affected to different degrees in different domains of emotion functioning (4, 6, 7), our data also suggest that these influential neurocognitive models need to be modified to account for the substantial proportion of youths with CD who performed normally across all emotion processing domains studied here. Thus, future

studies will need to (i) assess a broader range of emotion and non-emotion processes implicated in CD, such as executive function, decision-making, social cognition, and language skills (68, 69); (ii) investigate their performance profiles regarding neuropsychological heterogeneity; and then (iii) test their respective role in distinguishing between different developmental pathways of established and potentially novel clinical subtypes of CD, including testing the ability of each of these domains to predict the emergence or desistance of CD in high-risk groups as part of longitudinal studies (70).

It is, however, also conceivable that in fact emotion processing deficits only account for a proportion of youths with CD and that other biopsychosocial factors contribute to CD in those without such deficits (71). This idea is comparable to Moffitt's developmental taxonomy theory of CD suggesting that the adolescence-limited subtype does not suffer from neuropsychological deficits, whereas the early-onset and life-course-persistent subtype does (72) (see also (73) for a recent review). We note, however, that our data do not support this theory as age-of-onset of CD (although defined only retrospectively according to information from the K-SADS-PL interviews) was unrelated to the presence or absence of emotion processing deficits in this sample of youths with CD. This is consistent with previous smaller-scale studies showing no differences in emotion processing between CO-CD and AO-CD subtypes, including facial emotion recognition (74) and emotional learning (75), whereas both groups were deficient across these neuropsychological domains compared to TDCs. It should be stressed, though, that the empirical foundation of Moffitt's theory is not directly concerned with emotion functioning, but focuses instead on neurocognitive domains, such as reading, memory, vocabulary, and IQ (76). Therefore, more work is needed to determine the extent to which emotion functioning, in parallel to other neurocognitive mechanisms, contributes to different clinical manifestations and pathways within the CD population (including those with and without the LPE specifier) (77).

Both our dimensional and categorical data suggest that emotion recognition is the most consistently deficient neurocognitive domain in this sample of youths with CD. We can only speculate whether deficits in emotion recognition are more pivotal in the etiology of CD compared to both emotion learning and emotion regulation, or whether the specific task used to assess emotion recognition was simply more sensitive in detecting case-control differences than the two tasks that assessed the remaining emotion domains. Future studies might try to address this point, for instance, by creating tasks that are of comparable complexity and difficulty across various neurocognitive functions [e.g., emotion learning task (78), or emotion regulation task (79)].

Notably, our results neither confirm that youths with CD with the LPE specifier were disproportionately deficient in emotion recognition nor show that those without the LPE specifier displayed difficulties specifically in emotion regulation (4, 6). We did, however, find that being male and having a lower IQ—both well-documented risk factors for CD (50)—were associated with deficient emotion recognition skills [see

for related findings (80, 81)], whereas maternal smoking during pregnancy—another well-known CD risk factor—increased the odds for deficient emotional learning; none of the risk factors tested here predicted an emotion regulation deficit. Although we are cautious in interpreting these results since we measured most of the risk factors retrospectively, the findings suggest that some biopsychosocial dispositions may increase the probability of developing specific types of emotion dysfunction, while others do not (or might do so only in interaction with other risk variables) (82). This idea is worth pursuing in future prospective longitudinal studies that examine a wider range of CD-related risk factors, including genetic and epigenetic processes (83–85), their complex interactions as well as their unique contribution to emotion dysfunctions as potential intermediate phenotypes at both the behavioral and brain level in CD.

This study had some additional limitations (see (17) for strengths and limitations regarding the sample composition): First, each emotion domain was assessed using only one experimental task which makes our neurocognitive battery less representative, thereby limiting the generalizability of the results. Follow-up studies should preferably apply more comprehensive test batteries including several tasks measuring overlapping emotion domains so as to replicate and extend our findings and, thus, obtain a richer understanding of emotion functioning in CD. Second, we stratified youths with CD as “deficient” or “intact” in terms of task performance, following a clinically motivated, person-centered subgrouping approach that is typical for neuropsychological assessments, highly relevant for clinical decision-making, and easily applicable by practitioners (86). However, the present procedure—and other “traditional” statistical data clustering methods aiming to fractionate clinical groups on the basis of neuropsychological scores—have been criticized for several reasons which go beyond the current study and cannot be discussed in depth here [but see (65)]. Notably, the most serious issue is that these clustering methods divide the data arbitrarily into a pre-specified number of severity classes regardless of the underlying data distribution (i.e., two classes here: “deficient” vs. “intact” individuals), and they always produce a result. This sometimes leads to even unmanageably small subgroups, such as the 1% of youths with CD who were found to show deficits across all three emotion domains in the present study. Thus, alternatively, more elaborate machine-learning approaches, such as “normative modeling” (87), have been proposed to map neuropsychological variation with the advantage of not making strong *a priori* assumptions about the existence or number of subgroups with abnormal task performance (i.e., defined as extreme value, or “outlier,” from the normative range). It will be interesting to see how these novel subgrouping algorithms complement traditional clustering methods in identifying distinct neurocognitive subtypes of CD with potentially unique clinical profiles and underlying biology.

In conclusion, the current findings provide first evidence that youths with CD display strikingly diverse profiles in neuropsychological performance across three domains of emotion processing that have previously been linked to CD etiology, including emotion recognition, emotion learning, and emotion regulation (4). Similar to findings in ADHD (88), we

were able to reveal different neurocognitive subgroups in which emotion functioning was deficient to varying degrees, with a sizable subgroup of CD cases showing no meaningful emotion processing deficits at all. Clearly, deficits within a specific emotion processing domain may be clinically important for only a subgroup of patients, but not for the entire CD population. Consequently, treatments targeting emotion processing may be beneficial for some, but not all, individuals with CD (12, 89, 90). Whether, and if so which, neuropsychological interventions in non-emotion domains may be required to help patients with CD who have intact emotion processing skills needs to be evaluated in future studies. Importantly, while the current study should be regarded as exploratory and illustrative since we used, for the first time, a common-metric analytic approach from pediatric neuropsychology in order to subgroup youths with CD based on their emotion processing performance, novel classification algorithms based on machine learning are warranted to assist in identifying and validating distinct and meaningful neurocognitive phenotypes of CD, ideally replicating and thus substantiating the findings of this study.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on reasonable request in line with the data sharing policy of the FemNAT-CD consortium.

ETHICS STATEMENT

Written informed consent was obtained for all participants after they were informed about the experimental procedures and the aims of the study. Local ethics committees at all sites of the FemNAT-CD consortium reviewed and approved the study protocol. The study was carried out in accordance with the

recommendations of good clinical practice and in accordance with the Declaration of Helsinki and national legislation.

AUTHOR CONTRIBUTIONS

GK, GF, and KK conceived the current study. AB, AM, AS, KG-M, AW, JR, RP, HO, LJ, AR, LK, JA, and SB recruited participants and carried out assessments and data collection. GF, BH-D, AV, HL, AH, AF-R, AP, CS, SD, CF, and KK supervised the study. GK performed the statistical analyses and drafted the manuscript. GF, CS, SD, CF, and KK contributed to writing parts of the manuscript. CF coordinated the project which this study was part of. All authors read, commented on drafts of the paper, and approved the final manuscript and its submission.

FUNDING

This study was funded by the European Commission's Seventh Framework Programme (FP7/2007-2013) under Grant Agreement no. 602407 (FemNAT-CD, coordinator: CF, University Hospital Frankfurt, Germany).

ACKNOWLEDGMENTS

The authors thank all participants and their families for taking part in this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsy.2020.585052/full#supplementary-material>

REFERENCES

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5)*. American Psychiatric Association (2013).
2. Rivenbark JG, Odgers CL, Caspi A, Harrington H, Hogan S, Houts RM, et al. The high societal costs of childhood conduct problems: evidence from administrative records up to age 38 in a longitudinal birth cohort. *J Child Psychol Psychiatry*. (2018) 59:703–10. doi: 10.1111/jcpp.12850
3. Fairchild G, Hawes DJ, Frick PJ, Copeland WE, Odgers CL, Franke B, et al. Conduct disorder. *Nat Rev Dis Primers*. (2019) 5:43. doi: 10.1038/s41572-019-0095-y
4. Blair RJR, Leibenluft E, Pine DS. Conduct disorder and callous-unemotional traits in youth. *N Engl J Med*. (2014) 371:2207–16. doi: 10.1056/NEJMra1315612
5. Blair RJR. The neurobiology of psychopathic traits in youths. *Nat Rev Neurosci*. (2013) 14:786–99. doi: 10.1038/nrn3577
6. Frick PJ, Viding E. Antisocial behavior from a developmental psychopathology perspective. *Dev Psychopathol*. (2009) 21:1111–31. doi: 10.1017/S0954579409990071
7. Blair RJR, Veroude K, Buitelaar JK. Neuro-cognitive system dysfunction and symptom sets: a review of fMRI studies in youth with conduct problems. *Neurosci Biobehav Rev*. (2018) 91:69–90. doi: 10.1016/j.neubiorev.2016.10.022
8. Rubia K. “Cool” inferior frontostriatal dysfunction in attention-deficit/hyperactivity disorder vs. “hot” ventromedial orbitofrontal-limbic dysfunction in conduct disorder: a review. *Biol Psychiatry*. (2011) 69:e69–87. doi: 10.1016/j.biopsych.2010.09.023
9. Blair RJR. Psychopathic traits from an RDoC perspective. *Curr Opin Neurobiol*. (2015) 30:79–84. doi: 10.1016/j.conb.2014.09.011
10. Hawes DJ, Price MJ, Dadds MR. Callous-unemotional traits and the treatment of conduct problems in childhood and adolescence: a comprehensive review. *Clin Child Fam Psychol Rev*. (2014) 17:248–67. doi: 10.1007/s10567-014-0167-1
11. Wilkinson S, Waller R, Viding E. Practitioner review: involving young people with callous unemotional traits in treatment—does it work? A systematic review. *J Child Psychol Psychiatry*. (2016) 57:552–65. doi: 10.1111/jcpp.12494
12. Pisano S, Muratori P, Gorga C, Levantini V, Iuliano R, Catone G, et al. Conduct disorders and psychopathy in children and adolescents: aetiology, clinical presentation and treatment strategies of callous-unemotional traits. *Ital J Pediatr*. (2017) 43:84. doi: 10.1186/s13052-017-0404-6
13. Frick PJ, Ray JV, Thornton LC, Kahn RE. Can callous-unemotional traits enhance the understanding, diagnosis, and treatment of serious conduct problems in children and adolescents? A comprehensive review. *Psychol Bull*. (2014) 140:1–57. doi: 10.1037/a0033076
14. Viding E, Fontaine NMG, McCrory EJ. Antisocial behaviour in children with and without callous-unemotional traits. *J R Soc Med*. (2012) 105:195–200. doi: 10.1258/jrsm.2011.110223
15. White SE, Pope K, Sinclair S, Fowler KA, Brislin SJ, Williams WC, et al. Disrupted expected value and prediction error signaling in youths with

- disruptive behavior disorders during a passive avoidance task. *Am J Psychiatry*. (2013) 170:315–23. doi: 10.1176/appi.ajp.2012.12060840
16. Herpers PCM, Scheepers FE, Bons DMA, Buitelaar JK, Rommelse NNJ. The cognitive and neural correlates of psychopathy and especially callous-unemotional traits in youths: a systematic review of the evidence. *Dev Psychopathol*. (2014) 26:245–73. doi: 10.1017/S0954579413000527
 17. Kohls G, Baumann S, Gundlach M, Scharke W, Bernhard A, Martinelli A, et al. Investigating sex differences in emotion recognition, learning, and regulation among youths with conduct disorder. *J Am Acad Child Adolesc Psychiatry*. (2020) 59:263–73. doi: 10.1016/j.jaac.2019.04.003
 18. Dawel A, O'Kearney R, McKone E, Palermo R. Not just fear and sadness: meta-analytic evidence of pervasive emotion recognition deficits for facial and vocal expressions in psychopathy. *Neurosci Biobehav Rev*. (2012) 36:2288–304. doi: 10.1016/j.neubiorev.2012.08.006
 19. Smith SF, Lilienfeld SO. The response modulation hypothesis of psychopathy: a meta-analytic and narrative analysis. *Psychol Bull*. (2015) 141:1145–77. doi: 10.1037/bul0000024
 20. Morgan AB, Lilienfeld SO. A meta-analytic review of the relation between antisocial behavior and neuropsychological measures of executive function. *Clin Psychol Rev*. (2000) 20:113–36. doi: 10.1016/S0272-7358(98)00096-8
 21. Nigg JT, Willcutt EG, Doyle AE, Sonuga-Barke EJS. Causal heterogeneity in attention-deficit/hyperactivity disorder: do we need neuropsychologically impaired subtypes? *Biol Psychiatry*. (2005) 57:1224–30. doi: 10.1016/j.biopsych.2004.08.025
 22. Fanti KA, Panayiotou G, Lazarou C, Michael R, Georgiou G. The better of two evils? Evidence that children exhibiting continuous conduct problems high or low on callous-unemotional traits score on opposite directions on physiological and behavioral measures of fear. *Dev Psychopathol*. (2016) 28:185–98. doi: 10.1017/S0954579415000371
 23. Nolen-Hoeksema S, Watkins ER. A heuristic for developing transdiagnostic models of psychopathology: explaining multifinality and divergent trajectories. *Perspect Psychol Sci*. (2011) 6:589–609. doi: 10.1177/1745691611419672
 24. McLaughlin KA. Future directions in childhood adversity and youth psychopathology. *J Clin & Adolesc Psychol*. (2016) 45:361–82. doi: 10.1080/15374416.2015.1110823
 25. Bevilacqua L, Goldman D. Genetics of emotion. *Trends Cogn Sci*. (2011) 15:401–8. doi: 10.1016/j.tics.2011.07.009
 26. Pollak SD, Kistler DJ. Early experience is associated with the development of categorical representations for facial expressions of emotion. *Proc Natl Acad Sci USA*. (2002) 99:9072–6. doi: 10.1073/pnas.142165999
 27. Dodge KA, Pettit GS, Bates JE, Valente E. Social information-processing patterns partially mediate the effect of early physical abuse on later conduct problems. *J Abnorm Psychol*. (1995) 104:632–43. doi: 10.1037/0021-843X.104.4.632
 28. Shackman JE, Pollak SD. Impact of physical maltreatment on the regulation of negative affect and aggression. *Dev Psychopathol*. (2014) 26:1021–33. doi: 10.1017/S0954579414000546
 29. McLaughlin KA, DeCross SN, Jovanovic T, Tottenham N. Mechanisms linking childhood adversity with psychopathology: learning as an intervention target. *Behav Res Ther*. (2019) 118:101–9. doi: 10.1016/j.brat.2019.04.008
 30. Meyer-Lindenberg A, Weinberger DR. Intermediate phenotypes and genetic mechanisms of psychiatric disorders. *Nat Rev Neurosci*. (2006) 7:818–27. doi: 10.1038/nrn1993
 31. Freitag CM, Konrad K, Stadler C, De Brito SA, Popma A, Herpertz SC, et al. Conduct disorder in adolescent females: current state of research and study design of the FemNAT-CD consortium. *Eur Child Adolesc Psychiatry*. (2018) 27:1077–93. doi: 10.1007/s00787-018-1172-6
 32. Sonuga-Barke E, Bitsakou P, Thompson M. Beyond the dual pathway model: evidence for the dissociation of timing, inhibitory, and delay-related impairments in attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry*. (2010) 49:345–55. doi: 10.1016/j.jaac.2009.12.018
 33. Wählstedt C, Thorell LB, Bohlin G. Heterogeneity in ADHD: neuropsychological pathways, comorbidity and symptom domains. *J Abnorm Child Psychol*. (2009) 37:551–64. doi: 10.1007/s10802-008-9286-9
 34. Coghill DR, Seth S, Matthews K. A comprehensive assessment of memory, delay aversion, timing, inhibition, decision making and variability in attention deficit hyperactivity disorder: advancing beyond the three-pathway models. *Psychol Med*. (2014) 44:1989–2001. doi: 10.1017/S003329713002547
 35. Sjöwall D, Roth L, Lindqvist S, Thorell LB. Multiple deficits in ADHD: executive dysfunction, delay aversion, reaction time variability, and emotional deficits. *J Child Psychol Psychiatry*. (2013) 54:619–27. doi: 10.1111/jcpp.12006
 36. Moffitt TE, Lynam DR, Silva PA. Neuropsychological tests predicting persistent male delinquency. *Criminology*. (1994) 32:277–300. doi: 10.1111/j.1745-9125.1994.tb01155.x
 37. Frost LA, Moffitt TE, McGee R. Neuropsychological correlates of psychopathology in an unselected cohort of young adolescents. *J Abnorm Psychol*. (1989) 98:307–13. doi: 10.1037/0021-843X.98.3.307
 38. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR)*. American Psychiatric Association (2000).
 39. Colins OF, Vermeiren RRJ. The usefulness of DSM-IV and DSM-5 conduct disorder subtyping in detained adolescents. *J Nerv Ment Dis*. (2013) 201:736–43. doi: 10.1097/NMD.0b013e3182a20e94
 40. Pauli R, Tino P, Rogers JC, Baker R, Clanton R, Birch P Pauli R, et al. Positive and negative parenting in conduct disorder with high vs. low levels of callous-unemotional traits. *Dev Psychopathol*. (2020). doi: 10.1017/S0954579420000279. [Epub ahead of print].
 41. Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, 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*. (1997) 36:980–8. doi: 10.1097/00004583-199707000-00021
 42. Wechsler D. *Wechsler Intelligence Scale for Children*. 4th ed (WISC-IV). San Antonio, TX: Psychological Corporation (2003).
 43. Wechsler D. *Wechsler Adult Intelligence Scale*. 4th ed (WAIS-IV). San Antonio, TX: Psychological Corporation (2008).
 44. Wechsler D. *Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio, TX: Psychological Corporation (1999).
 45. McCrimmon AW, Smith AD. Review of the Wechsler abbreviated scale of intelligence, second edition (WASI-II). *J Psychoeduc Assess*. (2013) 31:337–41. doi: 10.1177/0734282912467756
 46. Andershed H, Kerr M, Stattin H, Levander S. Psychopathic traits in non-referred youths: a new assessment tool. In: Blaauw E, Sheridan L, editors. *Psychopaths: Current International Perspectives*. The Hague: Elsevier (2002). p. 131–58.
 47. van Baardewijk Y, Stegge H, Andershed H, Thomaes S, Scholte E, Vermeiren R. Measuring psychopathic traits in children through self-report. The development of the Youth psychopathic traits inventory-child version. *Int J Law Psychiatry*. (2008) 31:199–209. doi: 10.1016/j.jltp.2008.04.004
 48. Raine A, Dodge K, Loeber R, Gatzke-Kopp L, Lynam D, Reynolds C, et al. The reactive-proactive aggression questionnaire: differential correlates of reactive and proactive aggression in adolescent boys. *Aggress Behav*. (2006) 32:159–71. doi: 10.1002/ab.20115
 49. Boden JM, Fergusson DM, Horwood LJ. Risk factors for conduct disorder and oppositional/defiant disorder: evidence from a New Zealand birth cohort. *J Am Acad Child Adolesc Psychiatry*. (2010) 49:1125–33. doi: 10.1097/00004583-201011000-00006
 50. Moffitt TE, Caspi A, Rutter M, Silva PA. *Sex Differences in Antisocial Behaviour: Conduct Disorder, Delinquency, and Violence in the Dunedin Longitudinal Study*. Cambridge: Cambridge University Press (2001).
 51. Bifulco A, Bernazzani O, Moran PM, Jacobs C. The childhood experience of care and abuse questionnaire (CECAQ): validation in a community series. *Br J Clin Psychol*. (2005) 44:563–81. doi: 10.1348/014466505X35344
 52. Ruchkin VV, Schwab-Stone M, Vermeiren R. *Social and Health Assessment (SAHA) Psychometric Development Summary*. Yale, NH: University (2004)
 53. Isaksson J, Sjöblom S, Schwab-Stone M, Stickley A, Ruchkin V. risk factors associated with alcohol use in early adolescence among American inner-city youth: a longitudinal study. *Subst Use Misuse*. (2020) 55:358–66. doi: 10.1080/10826084.2019.1671867
 54. Calder AJ. Facial emotion recognition after bilateral amygdala damage: differentially severe impairment of fear. *Cogn Neuropsychol*. (1996) 13:699–745. doi: 10.1080/026432996381890

55. Newman JP, Kosson DS. Passive avoidance learning in psychopathic and nonpsychopathic offenders. *J Abnorm Psychol.* (1986) 95:252–6. doi: 10.1037/0021-843X.95.3.252
56. Wong AC-N, Palmeri TJ, Gauthier I. Conditions for face-like expertise with objects: becoming a ziggerin expert - but which type? *Psychol Sci.* (2009) 20:1108–17. doi: 10.1111/j.1467-9280.2009.02430.x
57. Hare TA, Tottenham N, Davidson MC, Glover GH, Casey BJ. Contributions of amygdala and striatal activity in emotion regulation. *Biol Psychiatry.* (2005) 57:624–32. doi: 10.1016/j.biopsych.2004.12.038
58. Tottenham N, Hare TA, Casey BJ. Behavioral assessment of emotion discrimination, emotion regulation, and cognitive control in childhood, adolescence, and adulthood. *Front Psychol.* (2011) 2:39. doi: 10.3389/fpsyg.2011.00039
59. Schulz KP, Fan J, Magidina O, Marks DJ, Hahn B, Halperin JM. Does the emotional go/no-go task really measure behavioral inhibition? Convergence with measures on a non-emotional analog. *Arch Clin Neuropsychol.* (2007) 22:151–60. doi: 10.1016/j.acn.2006.12.001
60. Palermo R, O'Connor KB, Davis JM, Irons J, McKone E. New tests to measure individual differences in matching and labelling facial expressions of emotion, and their association with ability to recognise vocal emotions and facial identity. *PLoS ONE.* (2013) 8:e0068126. doi: 10.1371/journal.pone.0068126
61. Epstein MK, Poythress NG, Brandon KO. The self-report psychopathy scale and passive avoidance learning: a validation study of race and gender effects. *Assessment.* (2006) 13:197–207. doi: 10.1177/1073191105284992
62. Stevens JP. *Applied Multivariate Statistics for the Social Sciences.* 5th ed. New York, NY: Routledge (2009).
63. Cohen J. *Statistical Power Analysis for the Behavioral Sciences.* New York, NY: Academic Press (1988).
64. Herba CM, Landau S, Russell T, Ecker C, Phillips ML. The development of emotion-processing in children: effects of age, emotion, and intensity. *J Child Psychol Psychiatry.* (2006) 47:1098–106. doi: 10.1111/j.1469-7610.2006.01652.x
65. Marquand AF, Wolfers T, Mennes M, Buitelaar J, Beckmann CF. Beyond lumping and splitting: a review of computational approaches for stratifying psychiatric disorders. *Biol Psychiatry Cogn Neurosci Neuroimaging.* (2016) 1:433–47. doi: 10.1016/j.bpsc.2016.04.002
66. Kaiser HF. An index of factorial simplicity. *Psychometrika.* (1974) 39:31–6. doi: 10.1007/BF02291575
67. Velicer WF, Peacock AC, Jackson DN. A comparison of component and factor patterns: a Monte Carlo approach. *Multivariate Behav Res.* (1982) 17:371–88. doi: 10.1207/s15327906mbr1703_5
68. Lynam D, Henry B. The role of neuropsychological deficits in conduct disorders. In: Hill J, Maughan B, editors. *Conduct Disorders in Childhood and Adolescence.* Cambridge: Cambridge University Press (2000). p. 235–63.
69. Moffitt TE. The neuropsychology of conduct disorder. *Dev Psychopathol.* (1993) 5:135–51. doi: 10.1017/S0954579400004302
70. Fanti KA. Understanding heterogeneity in conduct disorder: a review of psychophysiological studies. *Neurosci Biobehav Rev.* (2018) 91:4–20. doi: 10.1016/j.neubiorev.2016.09.022
71. Raine A. Biosocial studies of antisocial and violent behavior in children and adults: a review. *J Abnorm Child Psychol.* (2002) 30:311–26. doi: 10.1023/A:1015754122318
72. Moffitt TE. Adolescence-limited and life-course-persistent antisocial behavior: a developmental taxonomy. *Psychol Rev.* (1993) 100:674–701. doi: 10.1037/0033-295X.100.4.674
73. Moffitt TE. Male antisocial behaviour in adolescence and beyond. *Nat Hum Behav.* (2018) 2:177–86. doi: 10.1038/s41562-018-0309-4
74. Fairchild G, Van Goozen SHM, Calder AJ, Stollery SJ, Goodyer IM. Deficits in facial expression recognition in male adolescents with early-onset or adolescence-onset conduct disorder. *J Child Psychol Psychiatry.* (2009) 50:627–36. doi: 10.1111/j.1469-7610.2008.02020.x
75. Fairchild G, Van Goozen SH, Stollery SJ, Goodyer IM. Fear conditioning and affective modulation of the startle reflex in male adolescents with early-onset or adolescence-onset conduct disorder and healthy control subjects. *Biol Psychiatry.* (2008) 63:279–85. doi: 10.1016/j.biopsych.2007.06.019
76. Moffitt TE, Caspi A. Childhood predictors differentiate life-course persistent and adolescence-limited antisocial pathways among males and females. *Dev Psychopathol.* (2001) 13:355–75. doi: 10.1017/S0954579401002097
77. Krol N, Morton J, De Bruyn E. Theories of conduct disorder: a causal modelling analysis. *J Child Psychol Psychiatry.* (2004) 45:727–42. doi: 10.1111/j.1469-7610.2004.00267.x
78. Finger EC, Marsh AA, Mitchell DG, Reid ME, Sims C, Budhani S, et al. Abnormal ventromedial prefrontal cortex function in children with psychopathic traits during reversal learning. *Arch Gen Psychiatry.* (2008) 65:586–94. doi: 10.1001/archpsyc.65.5.586
79. Raschle NM, Fehlbaum LV, Menks WM, Martinelli A, Prätzlich M, Bernhard A, et al. Atypical dorsolateral prefrontal activity in female adolescents with conduct disorder during effortful emotion regulation. *Biol Psychiatry Cogn Neurosci Neuroimaging.* (2019) 4:984–94. doi: 10.1016/j.bpsc.2019.05.003
80. Buitelaar JK, van der Wees M, Swaab-Barneveld H, van der Gaag RJ. Verbal memory and performance IQ predict theory of mind and emotion recognition ability in children with autistic spectrum disorders and in psychiatric control children. *J Child Psychol Psychiatry.* (1999) 40:869–81. doi: 10.1111/1469-7610.00505
81. Lawrence K, Campbell R, Skuse D. Age, gender, and puberty influence the development of facial emotion recognition. *Front Psychol.* (2015) 6:761. doi: 10.3389/fpsyg.2015.00761
82. Dodge K, Coie L, Lynam D. Aggression and antisocial behavior in youth. In: Eisenberg N, editor. *Handbook of Child Psychology: Social, Emotional and Personality Development.* Vol. 3. 6th ed. New York, NY: John Wiley and Sons (2006). p. 719–85.
83. Loeber R, Farrington DP. Young children who commit crime: epidemiology, developmental origins, risk factors, early interventions, and policy implications. *Dev Psychopathol.* (2000) 12:737–62. doi: 10.1017/S0954579400004107
84. Murray J, Farrington DP. Risk factors for conduct disorder and delinquency: key findings from longitudinal studies. *Can J Psychiatry.* (2010) 55:633–42. doi: 10.1177/070674371005501003
85. Hill J. Biological, psychological and social processes in the conduct disorders. *J Child Psychol Psychiatry.* (2002) 43:133–64. doi: 10.1111/1469-7610.00007
86. Tsatsanis KD. Psychological and neuropsychological assessment of children. In: Martin A, Volkmar FR, editors. *Lewis's Child and Adolescent Psychiatry. A Comprehensive Textbook.* 4th ed. Philadelphia, PA: Wolters Kluwer (2007). p. 357–70.
87. Marquand AF, Rezek I, Buitelaar J, Beckmann CF. Understanding heterogeneity in clinical cohorts using normative models: beyond case-control studies. *Biol Psychiatry.* (2016) 80:552–61. doi: 10.1016/j.biopsych.2015.12.023
88. Nigg JT, Karalunas SL, Feczko E, Fair DA. Toward a revised nosology for attention-deficit/hyperactivity disorder heterogeneity. *Biol Psychiatry Cogn Neurosci Neuroimaging.* (2020) 5:726–37. doi: 10.1016/j.bpsc.2020.02.005
89. Dadds MR, Cauchi AJ, Wimalaweera S, Hawes DJ, Brennan J. Outcomes, moderators, and mediators of empathic-emotion recognition training for complex conduct problems in childhood. *Psychiatry Res.* (2012) 199:201–7. doi: 10.1016/j.psychres.2012.04.033
90. Hubble K, Bowen KL, Moore SC, van Goozen SHM. Improving negative emotion recognition in young offenders reduces subsequent crime. *PLoS ONE.* (2015) 10:e0132035. doi: 10.1371/journal.pone.0132035

Conflict of Interest: KK has received speaker fees from Shire Pharmaceuticals and Medice. CS receives royalties for a book on aggression. CF receives royalties for books on ADHD, ASD, and Depression. SD has received speaker fees from the Child Mental Health Centre and the Centre for Integrated Molecular Brain Imaging.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Kohls, Fairchild, Bernhard, Martinelli, Smaragdi, Gonzalez-Madruga, Wells, Rogers, Pauli, Oldenhof, Jansen, Rhijn, Kersten, Alfano, Baumann, Herpertz-Dahlmann, Vetro, Lazaratou, Hervas, Fernández-Rivas, Popma, Stadler, De Brito, Freitag and Konrad. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.