

Facial Emotion Recognition and Eye Movement Behaviour in Conduct Disorder

Journal:	<i>Journal of Child Psychology and Psychiatry</i>
Manuscript ID:	JCPP-OA-2016-00614.R2
Manuscript Type:	Original Article
Date Submitted by the Author:	30-Jun-2017
Complete List of Authors:	Martin-Key, Nayra; University of Southampton, Psychology Graf, Erich; University of Southampton, Psychology Adams, Wendy; University of Southampton, Psychology Fairchild, Graeme; University of Bath, Psychology
Key Words:	Conduct disorder, Callous unemotional traits, Emotion recognition, Eye movement, Sex differences

Facial Emotion Recognition and Eye Movement Behaviour in Conduct Disorder

N. A. Martin-Key¹, E. W. Graf¹, W. J. Adams^{1*} and G. Fairchild^{2*}

¹Department of Psychology, University of Southampton, UK

²Department of Psychology, University of Bath, UK

*Equal contribution.

Declaration of interest:

Nayra Martin-Key was supported by Kids Company and a Vice-Chancellor’s Scholarship from the University of Southampton. No conflicts of interest declared.

Abbreviated title: Emotion Recognition in Conduct Disorder

Total word count: 7304

Abstract: 248

3 Figures and 2 Tables; Figure S1 and Tables S1-3

*Emotion Recognition in Conduct Disorder***Abstract**

Background: Conduct Disorder (CD) is associated with impairments in facial emotion recognition. However, it is unclear whether such deficits are explained by a failure to attend to emotionally informative face regions, such as the eyes, or by problems in the appraisal of emotional cues.

Method: Male and female adolescents with CD and varying levels of callous-unemotional (CU) traits and age- and sex-matched typically-developing (TD) controls (aged 13-18) categorised the emotion of dynamic and morphed static faces. Concurrent eye tracking was used to relate categorisation performance to participants' allocation of overt attention.

Results: Adolescents with CD were worse at emotion recognition than TD controls, with deficits observed across static and dynamic expressions. In addition, the CD group fixated less on the eyes when viewing fearful and sad expressions. Across all participants, higher levels of CU traits were associated with fear recognition deficits and reduced attention to the eyes of surprised faces. Within the CD group, however, higher CU traits were associated with better fear recognition. Overall, males were worse at recognising emotions than females and displayed a reduced tendency to fixate the eyes.

Discussion: Adolescents with CD, and particularly males, showed deficits in emotion recognition and fixated less on the eyes when viewing emotional faces. Individual differences in fixation behaviour predicted modest variations in emotion categorisation. However, group differences in fixation were small and did not explain the much larger group differences in categorisation performance, suggesting that CD-related deficits in emotion recognition were not mediated by abnormal fixation patterns.

Keywords: Conduct disorder; callous-unemotional traits; emotion recognition; eye tracking; sex differences

Introduction

Conduct disorder (CD) is a pervasive and persistent form of disruptive behaviour that is characterised by the violation of other people’s rights or age-appropriate societal norms (American Psychiatric Association, 2013). CD is more common in males than in females (Moffitt et al., 2001), and reflecting this, most research on this disorder has been conducted using male-only samples. However, CD is associated with similar neuropsychological impairments in males and females, including lower verbal IQ (Lynam, Moffitt, & Stouthamer-Loeber, 1993; Moffitt & Silva, 1988; Pajer et al., 2008), deficits in autonomic fear conditioning (Fairchild, Stobbe, van Goozen, Calder, & Goodyer, 2010; Fairchild, Van Goozen, Stollery, & Goodyer, 2008), and reduced eye-blink startle responses (Fairchild et al., 2010, 2008). While research has demonstrated that CD is associated with impairments in facial emotion recognition in males and females (Fairchild et al., 2010; Fairchild, Van Goozen, Calder, Stollery, & Goodyer, 2009; Schwenck et al., 2012, 2014; Short, Sonuga-Barke, Adams, & Fairchild, 2016; Sully, Sonuga-Barke, & Fairchild, 2015), the underlying cause(s) of these deficits are not well understood. They could reflect difficulties with attention (e.g., impaired orienting to the eye region of the face) and/or appraisal (interpretation of stimuli that have been successfully encoded).

There is accumulating evidence that young people with callous-unemotional (CU) traits, i.e., a lack of concern for other people’s feelings, superficial affect, and a reduced ability to feel guilt (Pardini & Frick, 2013), are impaired in recognising fearful facial expressions (Marsh & Blair, 2008). This has been associated with a failure to attend to emotionally relevant regions of the face, such as the eyes (Dadds et al., 2006; Dadds, El Masry, Wimalaweera, & Guastella, 2008). Healthy individuals, and particularly females, show a preference for the eyes when viewing faces (Hall, Hutton, & Morgan, 2010). Moreover, this preference has been positively associated with facial expression recognition

Emotion Recognition in Conduct Disorder

accuracy (Hall et al., 2010); attending to the eyes appears to be important for recognition of all emotions (Spezio, Adolphs, Hurley, & Piven, 2007).

To date, only one study has examined eye movements in male adolescents with CU traits. The authors found selective deficits in fear recognition in those who were high in CU traits (Dadds et al., 2008). Notably, this group fixated less on the eyes than the low CU traits group. Interestingly, when instructed to look at the eye region, the deficit in fear recognition in the high CU traits group was ameliorated (Dadds et al., 2008), suggesting underlying problems with attentional allocation, rather than appraisal. Notably, this pattern of orienting is similar to that displayed by patients with amygdala damage: these individuals also demonstrate fear-specific recognition impairments that are improved by asking them to attend to the eyes (Adolphs et al., 2005).

The existing literature raises a number of important questions. First, given that adolescents with clinically-diagnosed CD have generally been reported to show global impairments in facial emotion recognition (Fairchild et al., 2010, 2009; Short et al., 2016; Sully et al., 2015), rather than specific difficulties with fear, it is unclear whether these broader difficulties are also explained by attentional deficits. It is therefore of interest to examine whether adolescents with CD show atypical eye movements across a wider range of emotional expressions, and whether atypical fixation patterns explain the more global deficits that have been reported in CD populations. Second, while research has shown similar emotion recognition impairments in males and females with CD (although see Pajer, Leininger, & Gardner, 2010), no studies have included males and females in the same experiment to examine whether the relationship between CD and emotion recognition differs by gender, and it is unclear whether comparable deficits in attention and/or appraisal are present in both sexes. Finally, the majority of studies in this area have used high intensity static facial stimuli that do not resemble the subtlety of expressions that we encounter in

Emotion Recognition in Conduct Disorder

everyday life (although see Bowen, Morgan, Moore, & van Goozen, 2013; Schwenck et al., 2012, 2014). By presenting static stimuli across a range of emotional intensities, as well as *dynamic* facial expressions of emotion, we can examine whether individuals with CD show impairments when processing more naturalistic stimuli.

Accordingly, the present study investigated facial emotion recognition in male and female adolescents with CD and varying levels of CU traits, and typically-developing (TD) controls. Static facial expressions morphed to display varying emotional intensities and dynamic, full-intensity facial expressions were included. Importantly, the emotion categorisation task was paired with eye tracking methods to assess whether CD-related deficits in emotion recognition could be explained by atypical fixation patterns (e.g., fixating less on the eyes). A secondary aim was to investigate whether CU traits **are associated with impaired** emotion recognition performance and attention to the eyes, in line with Dadds and colleagues' (2008) findings, and whether atypical fixation patterns mediate fear recognition deficits in those with high levels of CU traits.

We predicted that adolescents with CD would show deficits in emotion recognition, particularly for negative emotions such as anger, sadness, and fear, and lower intensity static expressions. In addition, we hypothesised that CD individuals with elevated CU traits would show particularly marked impairments for fear. We predicted that, overall, having CD would be related to a reduced tendency to fixate the eyes and that those with CD and elevated CU traits would show the most pronounced deficits in attending to the eyes, given Dadds et al.'s (2008) findings. Lastly, we hypothesised that deficits in facial emotion recognition in the CD group would be mediated by a failure to fixate the eyes, and that this relationship would be strongest for expressions that are principally communicated via the eyes, such as fear.

Method

Participants

One hundred and twenty-eight adolescents aged 13-18 years were recruited through Youth Offending Services and pupil referral units across Southampton and Hampshire, via referrals from caseworkers, and through mainstream schools and colleges in Southampton via mail-shots. Of these 128, five did not meet the inclusion criteria (see below), six TD and four CD participants could not be eye tracked (either due to technical difficulties or issues with eye tracker calibration), and two TD and 10 CD participants opted not to take part in the laboratory experiment after being interviewed. Thus, the final sample of 101 participants consisted of 50 adolescents (26 male) who met diagnostic criteria for CD, and 51 TD adolescents (26 male). All participants and the parents of those aged below 16 provided written informed consent to participate in the study, which was approved by the University Ethics Committee and the Hampshire County Council Children's Services Research Governance Committee.

Inclusion criteria for the study were: (i) fluency in English; (ii) no hard contact lenses or bi/tri-focal glasses; (iii) an Intelligence Quotient (IQ) ≥ 70 (assessed via the Wechsler Abbreviated Scale of Intelligence (Wechsler, 1999)); and (iv) being free of psychosis or Autism Spectrum Disorder (ASD). All participants were assessed for CD, Oppositional Defiant Disorder (ODD), Attention-Deficit/Hyperactivity Disorder (ADHD), Major Depressive Disorder (MDD), Generalised Anxiety Disorder (GAD), Obsessive Compulsive Disorder (OCD), Psychosis, Post-Traumatic Stress Disorder (PTSD), and Alcohol and Substance Use Disorders using the Schedule of Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime version (K-SADS-PL; Kaufman et al., 1997), a semi-structured diagnostic interview based on DSM-IV criteria. ASDs were assessed using the ASD module of the unpublished DSM-5 version of the K-SADS-PL.

Emotion Recognition in Conduct Disorder

Following extensive training and shadowing by experienced staff members, postgraduate students carried out separate diagnostic interviews with participants and caregivers. Regular case conferences were held to discuss difficult or borderline cases and we assessed the inter-rater reliability of CD and other disorders (Cohen’s kappas ranged from .87-1.00). Data were combined across informants such that a symptom was considered present if endorsed by either informant, as suggested by Kaufman and colleagues (1997). CU traits were assessed using the self-report Inventory of Callous-Unemotional traits (ICU; Frick, 2003; Cronbach’s alphas = .81 (entire sample), .82 (CD group)).

Procedure

Emotional face categorisation.

We assessed participants’ ability to categorise dynamic and static facial expressions of emotion. Our dynamic stimuli were drawn from the Amsterdam Dynamic Facial Expression Set (ADFES; van der Schalk, Hawk, Fischer, & Doosje, 2011). Fifty-six stimuli (8 models (4 males) expressing 7 emotions (anger, sadness, fear, happiness, surprise, disgust, and neutral)) were selected from the full set. The original ADFES sequences are approximately 6000ms in duration, with the final ~5000ms consisting of the actor holding the full emotion expression. We removed this final section such that our dynamic stimuli depicted a neutral to full emotion transition lasting 1000ms, broadly consistent with prior reports relating to the timing of emotion expressions from neutral poses (Yoshikawa & Sato, 2008).

In preparing the static stimuli, a bespoke morphing algorithm (Adams, Gray, Garner, & Graf, 2010) combined each full expression static stimulus from the ADFES with the same model’s neutral face. In total, 100 static stimuli were used (4 models (2 males, 2 females) × 6 facial expressions (plus neutral) × 4 emotion intensities (30%, 50%, 70%, 100%)). The images were scaled, aligned, masked (to hide external features, e.g., hair) and matched for

Emotion Recognition in Conduct Disorder

mean luminance and root-mean-square contrast. Stimuli were displayed on a 1024×768 monitor and subtended 10.5° of visual angle at a viewing distance of 60cm.

Participants completed four experimental blocks of 39 randomly interleaved trials, taking breaks between blocks as necessary. Each trial began with a 500ms fixation cross. Following a 1000ms stimulus presentation, participants were presented with seven emotion labels (anger, sadness, happiness, fear, surprise, disgust, and neutral) and used a mouse to select the emotion that best described the displayed expression (see Figure 1B). Participants were given an unlimited time to respond, but were instructed to be as quick and accurate as possible.

[Insert Figure 1 here]

Eye tracking.

Eye position was recorded using an EyeLink 1000 eye tracker (SR Research Ltd, Canada), with a monocular sampling rate of 1000Hz and mean spatial accuracy of ~ .25-.50°. A chin and forehead rest stabilised the head. For each participant, an initial tracker calibration was performed, in which participants sequentially fixated nine target points on the screen. In addition, during the experiment, each trial was preceded by a drift correction to ensure that the accuracy of the calibration parameters was maintained. This procedure was repeated when the drift error exceeded $> 1^\circ$, and the trial was only continued after a further calibration.

Data analytic strategy

Demographic characteristics.

Socioeconomic status (SES) was categorised according to the profession of the parent(s) using UK Office for National Statistics (ONS) guidelines (2010). Participants whose parents' professions fell under 'high or intermediate' ONS categories were classified

Emotion Recognition in Conduct Disorder

as high SES, whilst those categorised as ‘routine, manual, or unemployed’ were classified as low SES. Due to the limited variation in ethnicity in the sample, participants were categorised as Caucasian or non-Caucasian. Group differences in continuous variables were explored using one-factor ANOVAs, whereas comparisons of binary variables were conducted using Chi-Square tests.

Emotion categorisation data.

We employed linear mixed-effects model (LMM) analyses¹ in Matlab 8.5.0 (The MathWorksInc. Natick, MA, USA) using the ‘*fitlme*’ function to examine the effects of CD status, gender, CU traits, IQ, SES, stimulus type (dynamic vs. static), emotional intensity, and two-way interactions between these variables, on (i) overall categorisation accuracy, and (ii) categorisation accuracy for each emotion. Initial models also included subject, psychiatric comorbidity², and age as random factors.

When assessing categorisation accuracy across all emotions, emotional expression (anger, sadness, fear, happiness, surprise, disgust, and neutral) was also included as a random factor. This provided the maximum power to explore the effects of our fixed variables, while accounting for variance introduced by subject and emotional expression. This was followed by separate analyses of categorisation performance for each individual emotion. The significance of each predictor (and hence its inclusion in the final model) was defined by likelihood ratio tests comparing the models with and without each predictor (see Data Analytic Strategy in Supplementary Materials, available online, for further details). Simple effect sizes are quantified in the main text using Cohen’s *d* (small $\geq .20$, medium $\geq .50$, large $\geq .80$; Cohen, 1988). In the context of a multi-predictor model, Cohen’s f^2 can be used to

¹This approach was chosen over the traditional ANOVA approach as it provides a principled way to include both continuous and categorical variables, whilst modelling individual subject- and stimulus-related sources of variation. By more precisely modelling random variation, power to detect effects of interest is increased, without inflating the probability of Type II errors (see Baayen, Davidson, & Bates, 2008; Bates, Kliegl, Vasishth, & Baayen, 2015).

²PTSD, alcohol and substance abuse were not included as random factors as they were present in only one subject, and therefore were subsumed by the subject factor.

Emotion Recognition in Conduct Disorder

quantify the degree of variance explained by a single predictor, when accounting for all other variables. This additional effect size measure (which compliments the likelihood ratio tests described above) is reported in Supplementary Materials (Table S3, available online).

Eye tracking data.

In line with Dadds and colleagues (2008), we performed a region of interest (ROI) analysis on the eye position data. Bespoke software was used to manually select three ROIs for each of our face stimuli: two around the eyes and one around the mouth (see Figure S1 in Supplementary Materials), regions that have been shown to be important in conveying emotional expressions (Adolphs et al., 2002). At each 1ms time-point, the participant's eye position was evaluated against these ROIs.

We used these data to create two preference scores: (i) *initial eye preference*: the percentage of trials in which the participant moved from the first fixation point (a cross in the centre of the face) to fixate on the eye region first, minus the percentage of trials in which they moved to fixate on the mouth first, and (ii) *total eye preference*: the percentage of overall trial time spent fixating the eye region, minus that spent fixating the mouth. In both cases, the first fixation was disregarded, as fixation on the central cross was necessary in order to initiate the trial. LMM analyses were used to assess whether: (i) initial eye preference, or (ii) total eye preference were predicted by CD status, gender, CU traits, IQ, SES, stimulus type, emotional intensity, or their interactions.

Relating eye tracking to behavioural data.

Finally, we examined whether differences in fixation behaviour could explain (i.e., mediate) the relationship between participant characteristics (e.g., CD status, gender, CU traits) and emotion categorisation performance. Here, we investigated whether adding initial and/or total eye preference to the best-fitting models of emotion categorisation resulted in

Emotion Recognition in Conduct Disorder

significant improvements in the predictive power of the models that would imply that the associations were explained by fixation behaviour.

Results

Demographic and clinical characteristics

Table 1 presents information about the characteristics of each of the four groups, with statistical comparisons. The groups did not differ significantly in age or ethnicity. However, the CD groups had significantly lower IQ scores than the TD groups. In addition, CD males were more likely to come from low SES backgrounds than the TD groups. Finally, both CD groups had higher levels of CU traits than their sex-matched control groups.

[Insert Table 1 here]

Overall categorisation accuracy

Figure 2 shows emotion categorisation as a function of CD status, gender, emotional intensity (morph strength), and stimulus type (static vs. dynamic), with the results of the LMM analyses presented in Table 2³.

[Insert Figure 2 here]

Having CD and being male were both independently associated with significantly lower emotion categorisation accuracy across all emotions (CD status: $d = .72$, gender: $d = .97$). As expected, categorisation accuracy for static expressions increased with emotional intensity ($r^2 = .52$), but this effect was independent of CD status and gender.

[Insert Table 2 here]

³For simplicity, Table 2 only includes predictors of interest; i.e., those relating to our *a priori* hypotheses (see Table S1 in Supplementary Materials, available online, for full model summaries).

Categorisation accuracy for individual emotions

Next, we investigated categorisation accuracy for each emotion separately (see Table 2). Participants with CD showed significantly poorer fear recognition relative to TD participants ($d = .75$). In addition, higher levels of CU traits were associated with poorer fear recognition across the whole sample ($r^2 = .30$). However, there was also a significant interaction between CD status and CU traits such that within the CD group, those with higher levels of CU traits showed *better* fear recognition (see Figure S2 in Supplementary Materials, available online). Having CD ($d = .78$) and being male ($d = .45$) were both independently associated with poorer recognition of angry faces. Being male was also associated with poorer recognition of disgust, sadness, surprise, and neutral expressions ($ds = .41-.58$).

For all emotions, categorisation accuracy increased with stimulus intensity, and for all emotions except fear and disgust, categorisation was better for dynamic than static stimuli (see Table 2). These effects were independent of CD group status or gender, with one exception: for disgust recognition, CD status significantly interacted with emotional intensity, such that the TD group showed greater improvements in performance than the CD group as emotional intensity increased (see Table 2).

Eye movements: initial and total eye preference

Figure 3 shows eye movement behaviour as a function of CD status and gender when viewing both static and dynamic facial expressions.

On average, participants made 3.44 fixations per 1000ms trial (across all conditions), with an average fixation duration of 303.8ms; previous research suggests that the first two fixations are most predictive of facial emotion recognition (Shurgin et al., 2014) and facial emotion recognition typically occurs within 1000ms (De Sonnevile et al., 2002).

Importantly, the CD and TD groups did not differ in: (i) the percentage of trials in which the

Emotion Recognition in Conduct Disorder

first fixations were outside the ROIs, or (ii) the total time spent fixating outside the ROIs, and neither CD status nor gender interacted with emotion, or intensity for these measures (as assessed via separate ANOVAs for static and dynamic stimuli). Thus, the analyses below focus on the relative preference for eyes and mouth regions, which capture the critical group differences in eye movement behaviour.

[Insert Figure 3 here]

We first considered whether eye movement behaviour was related to CD status, gender, CU traits, IQ, SES, stimulus type, and emotional intensity across all emotional expressions (Table 2). Females showed increased initial eye preferences relative to males ($d = .51$). However, neither CD status nor CU traits were significantly associated with initial or total eye preferences across all emotional expressions.

Next, we performed separate eye movement analyses for each emotional expression. Participants with CD showed a reduced tendency to fixate the eyes first when viewing sad expressions ($d = .76$). In addition, this initial eye preference was lower in males, relative to females, and lower for dynamic, relative to static, stimuli (see Table 2). Having CD ($d = .41$) and being male ($d = .49$) were both associated with a lower total eye preference when viewing fearful expressions. Higher levels of CU traits were associated with a reduced tendency to fixate the eyes first when viewing surprised faces ($r^2 = .20$). On the other hand, CU traits and emotional intensity interacted to predict initial eye preference for surprise: this increased with emotional intensity, and this effect was larger in individuals with higher levels of CU traits (see Table 2).

Finally, relative to males, females displayed a stronger tendency to fixate the eyes first for angry, fearful, surprised, and neutral expressions ($ds = .48-.57$). With the exception of

Emotion Recognition in Conduct Disorder

disgust, females **also** showed greater total eye preferences than males for all emotions ($d_s = .43-.66$).

Eye movement behaviour as a predictor of categorisation accuracy

The preceding analyses revealed: (i) that CD status, gender, and CU traits (for fear alone) predicted emotion categorisation performance, and (ii) that CD status, gender, and CU traits (for surprise alone) **were related to** eye fixation patterns. This raises an important question – does a reduced tendency to fixate the eyes *explain* the emotion recognition deficits observed in **participants with** CD?

If eye fixation patterns influence the categorisation of emotional faces, then we would expect our models of categorisation to *improve* when measures of eye movement behaviour are added as predictors. However, if these fixation patterns *mediate* the emotion recognition deficits associated with CD and/or CU traits, then we would expect an additional effect to emerge when eye movement behaviours are added to our categorisation models: the magnitude of effects related to CD status or CU traits will be diminished (i.e., coefficients will become smaller), **because** the CD- or CU traits-related effects on categorisation performance are explained by abnormal fixation patterns.

Indeed, including initial and total eye preference measures significantly improved the best-fitting model of overall emotion categorisation (see Table S2 in Supplementary Materials, available online). However, these improvements were very small in the context of the overall effect sizes (see Table S3 for model improvements, as quantified via Cohen's f^2). Increased total eye preference was associated with improvements in overall emotion categorisation across all participants, irrespective of CD status or gender. This was partially counteracted by a negative relationship between initial eye preference and categorisation performance, in the context of the full, multi-predictor model.

Emotion Recognition in Conduct Disorder

In addition, eye movement behaviour added predictive power to the individual models of categorisation for anger, fear, happiness, and sadness (see Table S2 in Supplementary Materials). However, the model improvements were again very small, in terms of effect size (see Table S3). Participants' categorisation of anger improved with increasing total eye preference, irrespective of CD status or gender. This effect was weaker for dynamic and lower intensity static stimuli. Increased initial eye preference was associated with better fear categorisation, but only for high intensity expressions (the reverse was found for low intensity expressions). Higher total eye preference scores were associated with enhanced happiness recognition, but only for lower intensity expressions. Finally, initial eye preference scores were negatively associated with recognition of lower intensity sad expressions, across all participants.

The addition of eye movement measures led to very small, but significant, improvements in the models of categorisation across emotional expressions, and for fear, happiness, and sadness recognition considered separately. Importantly, however, the addition of these predictors had little effect on the importance of CD, gender or CU traits as independent predictors in the model – the average change in these predictors was -2%. In addition, we note that the effects of eye movement behaviour on emotion recognition did not vary as a function of CD status, CU traits or gender. In addition, irrespective of CD status, eye movement behaviour was only associated with modest effects on categorisation performance. In summary, there were only small differences between the CD and TD groups in fixation patterns, and these differences did not explain the much larger group differences in emotion categorisation. These analyses provide strong evidence that CD-related deficits in emotion categorisation are not mediated by atypical fixation patterns. Thus, difficulties interpreting emotional cues, rather than attentional issues, are likely to underlie the emotion recognition deficits observed in the CD group.

Discussion

The present study explored male and female typically-developing (TD) and Conduct Disorder (CD) adolescents' ability to categorise dynamic and morphed static facial expressions. We also assessed the effects of callous-unemotional (CU) traits on emotion recognition performance. Overall, having CD and being male had independent and detrimental effects on the recognition of facial expressions across multiple emotions. This pattern of findings held regardless of whether the stimuli were static or dynamic and also across different emotional intensities, and the associated effect sizes were in the medium-to-large range.

When considering each emotion separately, having CD and being male were independently related to poorer anger recognition. CD was also associated with impaired fear recognition. These findings support previous research demonstrating impaired recognition of anger and fear in adolescents with CD (Bowen et al., 2013; Fairchild et al., 2010, 2009; Short et al., 2016; Sully et al., 2015). Critically, our work extends these previous findings by demonstrating that anger and fear recognition deficits are still present when the stimuli are dynamic or presented at lower intensity. Furthermore, CU traits were associated with impaired fear recognition when considering the *entire* sample. Contrary to expectations, however, elevated CU traits were related to *enhanced* recognition of fearful expressions in the CD group.

Although the former finding for the overall sample is consistent with past research demonstrating an association between psychopathic traits and difficulties identifying fearful expressions (Blair, Colledge, Murray, & Mitchell, 2001; Fairchild et al., 2009), we note that the latter finding may be considered surprising given previous findings that impaired fear recognition is more pronounced in those with CD and elevated CU traits relative to those with lower levels of CU traits (Fairchild et al., 2009; Marsh & Blair, 2008), and theories

Emotion Recognition in Conduct Disorder

proposing that psychopathy is linked to selective deficits in processing distress cues (Blair, 1995, 2003). However, not all studies have found an association between CU traits and fear recognition impairments in CD populations. In fact, some researchers found no differences between individuals with CD and high versus low CU traits in emotion recognition performance (Sully et al., 2015) and others have shown *superior* fear recognition in children with conduct problems and elevated CU traits (Woodworth & Waschbusch, 2008).

Considering the eye tracking data, males were less likely to fixate the eye region first, relative to females, across all emotions. This confirms previous findings in adults showing sex differences in orienting towards the eyes (Hall et al., 2010). When considering each emotion separately, having CD and being male were both independently related to a reduced tendency to fixate the eyes for sad and fearful expressions. CU traits predicted reduced attention to the eyes, but this was only true for *surprised* expressions. This finding is partly in line with Dadds and colleagues' (2008) finding of reduced attention to the eyes in children high in CU traits. Critically, having CD and being male appear to be stronger predictors of reduced overt attention to the eyes than having elevated CU traits.

However, while eye movement behaviours differed according to both group status and gender, these differences were comparatively small and did not explain the larger group differences in emotion recognition performance. In fact, although individual differences in eye movement behaviour were associated with variation in emotion recognition accuracy, the addition of the eye movement **variables** to the models of categorisation accuracy provided only marginal benefits in their predictive power to explain categorisation performance.

Importantly, including these variables did not reduce the importance of CD status, CU traits or gender in explaining emotion recognition. Therefore, our findings suggest that although individuals with CD, and particularly males with CD, exhibit abnormal eye fixation patterns, these differences do not explain the association between CD status and emotion recognition

Emotion Recognition in Conduct Disorder

performance. Taken together, our findings suggest that problems interpreting emotional cues, rather than attentional issues, are likely to underlie the emotion recognition deficits observed in adolescents with CD.

To our knowledge, this is the first study to examine recognition of and attention to static and dynamic emotional faces in both male and female subjects with and without CD. Using both dynamic and lower intensity stimuli increased the study's ecological validity, and the fact that adolescents with CD showed impaired recognition of facial expressions across stimulus types suggests that previously reported deficits were not merely artefacts of using simplified, static stimuli. In addition, the CD and TD groups were well-characterised from a clinical perspective, using reliable diagnostic measures.

The study also had several limitations. First, the sample size was only moderate, reflecting the difficulties in recruiting adolescents, and particularly females, with CD. Furthermore, relative to the TD group, the CD group had lower IQs and came from lower socioeconomic status (SES) backgrounds. It is important to note, however, that CD has been repeatedly linked to lower IQ and SES (Piotrowska, Stride, Croft, & Rowe, 2015). For this reason, matching groups on these variables has the potential to reduce the representativeness of the groups, and therefore the generalisability of our findings. Additionally, we included both IQ and SES as fixed variables in all models, meaning that we accounted for the variance introduced by group differences in these variables. Critically, because our design is cross-sectional, we cannot infer that the presence of CD is causally related to deficits in emotion recognition. When analysing the behavioural data for individual emotional expressions, we did not apply a correction for multiple outcome measures: these data arose from separate subsets of trials. More generally, there is no consensus regarding the appropriateness or necessity of adjusting p values to account for multiple outcome measures; here, we provide effect sizes as a less contentious measure of statistical importance (see Feise, 2002). Finally,

Emotion Recognition in Conduct Disorder

it is worth noting that CU traits were assessed using the self-report version of the Inventory of Callous-Unemotional traits, which may be influenced by social desirability effects.

Conclusion

Adolescents with CD showed impairments in emotion recognition, and a reduced tendency to fixate the eye region of the face when viewing fearful and sad expressions. Interestingly, we found that having CD and being male had additive, detrimental effects on emotion recognition and attention to the eyes. As such, interventions seeking to improve emotion recognition in adolescents with CD may need to be tailored according to gender, with males being likely to require more comprehensive training programmes than females. Critically, our analyses suggest that CD-related deficits in facial emotion recognition might be better conceptualised as resulting from problems in the appraisal of emotional cues, rather than abnormal eye fixation patterns.

Legends for Figures

Figure 1. Task design and examples of facial expression stimuli used in the study. Panel A shows examples of static fearful stimuli, with intensities of 30%, 50%, 70%, and 100%, respectively. Panel B depicts a trial sequence. Participants viewed facial expressions for 1000ms and were then asked to label the emotion.

Figure 2. Emotion categorisation accuracy data for each group, as a function of emotion, intensity and stimulus type. Categorisation accuracy scores for static stimuli, split by emotion intensity (30%, 50%, 70%, 100%), and dynamic stimuli (full intensity; 100%) per emotion are shown in panels A-F. *Note:* error bars show +/-standard error. CD, conduct disorder; TD, typically-developing.

Figure 3. Eye movement data for each group, split by emotion and stimulus type. Initial and total eye preference scores for dynamic stimuli are shown in panels A-B, whereas eye preference scores for static stimuli (collapsed across emotional intensity, for ease of plotting) are presented in panels C-D. *Note:* error bars show +/-standard error. CD, conduct disorder; TD, typically-developing.

Acknowledgements:

Emotion Recognition in Conduct Disorder

We thank our participants and their parents for taking part in the study. We also thank the schools, pupil referral units, and the Southampton and Hampshire Youth Offending Services for helping with participant recruitment.

Correspondence:

Graeme Fairchild, Department of Psychology, University of Bath, Bath, BA2 7AH, UK.

E-mail: g.fairchild@bath.ac.uk

References

Emotion Recognition in Conduct Disorder

- Adams, W. J., Gray, K. L. H., Garner, M., & Graf, E. W. (2010). High-Level Face Adaptation Without Awareness. *Psychological Science*, 21(2), 205–210. doi: 10.1177/0956797609359508
- Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature*, 433(7021), 68–72. doi:10.1038/nature03086
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*. Washington, D.C.: American Psychiatric Association.
- Baayen, R., Davidson, D., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(Special Issue: Emerging Data Analysis), 390–412. doi:10.1016/j.jml.2007.12.005
- Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious Mixed Models. ArXiv e-prints.
- Blair, R. J. R. (1995). A cognitive developmental approach to morality: investigating the psychopath. *Cognition*, 57(1), 1–29. doi: 10.1016/0010-0277(95)00676-P
- Blair, R. J. R. (2003). Neurobiological basis of psychopathy. *The British Journal of Psychiatry*, 182(1), 5–7. doi: 10.1192/bjp.182.1.5
- Blair, R. J. R., Colledge, E., Murray, L., & Mitchell, D. G. V. (2001). A selective impairment in the processing of sad and fearful expressions in children with psychopathic tendencies. *Journal of Abnormal Child Psychology*, 29(6), 491–498. doi: 10.1023/A:1012225108281
- Bowen, K. L., Morgan, J. E., Moore, S. C., & van Goozen, S. H. M. (2013). Young

Emotion Recognition in Conduct Disorder

Offenders' Emotion Recognition Dysfunction Across Emotion Intensities: Explaining Variation Using Psychopathic Traits, Conduct Disorder and Offense Severity. *Journal of Psychopathology and Behavioral Assessment*, 36(1), 60–73. doi: 10.1007/s10862-013-9368-z

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, N.J.: L. Erlbaum Associates.

Dadds, M. R., Perry, Y., Hawes, D. J., Merz, S., Riddell, A. C., Haines, D. J., & ... Abeygunawardane, A. I. (2006). Attention to the eyes and fear-recognition deficits in child psychopathy. *British Journal of Psychiatry*, 189(9), 280–281. doi: 10.1192/bjp.bp.105.018150

Dadds, M. R., El Masry, Y., Wimalaweera, S., & Guastella, A. J. (2008). Reduced Eye Gaze Explains “Fear Blindness” in Childhood Psychopathic Traits. *Journal of the American Academy of Child & Adolescent Psychiatry*, 47(4), 455–463. doi: 10.1097/chi.0b013e31816407f1

De Sonnevile, L. M. J., Verschoor, C. A., Njokiktjien, C., Op het Veld, V., Toorenaar, N., & Vranken, M. (2002). Facial identity and facial emotions: speed, accuracy, and processing strategies in children and adults. *Journal of Clinical and Experimental Neuropsychology*, 24(2), 200–213. doi: 10.1076/jcen.24.2.200.989

Fairchild, G., Stobbe, Y., van Goozen, S. H. M., Calder, A. J., & Goodyer, I. M. (2010). Facial Expression Recognition, Fear Conditioning, and Startle Modulation in Female Subjects with Conduct Disorder. *Biological Psychiatry*, 68(3), 272–279. doi: 10.1016/j.biopsych.2010.02.019

Fairchild, G., Van Goozen, S. H. M., Calder, A. J., Stollery, S. J., & Goodyer, I. M. (2009).

Emotion Recognition in Conduct Disorder

Deficits in facial expression recognition in male adolescents with early-onset or adolescence-onset conduct disorder. *Journal of Child Psychology and Psychiatry*, 50(5), 627–636. doi: 10.1111/j.1469-7610.2008.02020.x

Fairchild, G., Van Goozen, S. H., Stollery, S. J., & Goodyer, I. M. (2008). Fear Conditioning and Affective Modulation of the Startle Reflex in Male Adolescents with Early-Onset or Adolescence-Onset Conduct Disorder and Healthy Control Subjects. *Biological Psychiatry*, 63(3), 279–285. doi: 10.1016/j.biopsych.2007.06.019

Feise R. J. (2002). Do multiple outcome measures require p-value adjustment? *BioMed Central Medical Research Methodology*, 2(1), 8. doi:10.1186/1471-2288-2-8

Frick, P. J. (2003). *The Inventory of Callous-Unemotional Traits*. Unpublished rating scale. The University of New Orleans.

Hall, J. K., Hutton, S. B., & Morgan, M. J. (2010). Sex differences in scanning faces: Does attention to the eyes explain female superiority in facial expression recognition? *Cognition & Emotion*, 24(4), 629. doi: 10.1080/02699930902906882

Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., & ... Ryan, N. D. (1997). Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. *Journal of the American Academy of Child and Adolescent Psychiatry*, 36(7), 980–988. doi: 10.1097/00004583-199707000-00021

Lynam, D., Moffitt, T., & Stouthamer-Loeber, M. (1993). Explaining the relation between IQ and delinquency: Class, race, test motivation, school failure, or self-control? *Journal of Abnormal Psychology*, 102(2), 187–196. doi: 10.1037/0021-843x.102.2.187

Marsh, A. A., & Blair, R. J. R. (2008). Deficits in facial affect recognition among antisocial

Emotion Recognition in Conduct Disorder

populations: A meta-analysis. *Neuroscience & Biobehavioral Reviews*, 32(3), 454–465.
doi: 10.1016/j.neubiorev.2007.08.003

Moffitt, T. E., Caspi, A., Rutter, M., & Silva, P. A. (2001). *Sex differences in antisocial behaviour: conduct disorder, delinquency, and violence in the Dunedin longitudinal study*. Cambridge: Cambridge University Press.

Moffitt, T. E., & Silva, P. A. (1988). IQ and delinquency: A direct test of the differential detection hypothesis. *Journal of Abnormal Psychology*, 97(3), 330–333. doi: 10.1037/0021-843x.97.3.330

Office for National Statistics (2010). *Standard occupational classification 2010*. Basingstoke, UK: Palgrave Macmillan.

Pajer, K. A., Stein, S., Tritt, K., Chang, C.-N., Wang, W., & Gardner, W. (2008). Conduct disorder in girls: neighborhoods, family characteristics, and parenting behaviors. *Child and Adolescent Psychiatry and Mental Health*, 2(10), 1–28. doi: 10.1186/1753-2000-2-28

Pajer, K., Leininger, L., & Gardner, W. (2010). Recognition of facial affect in girls with conduct disorder. *Psychiatry Research*, 175(3), 244–51. doi: 10.1016/j.psychres.2009.06.003

Pardini, D. & Frick, P. J. (2013). Multiple developmental pathways to conduct disorder: Current conceptualizations and clinical implications. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, 22(1), 20–25.

Piotrowska, P. J., Stride, C. B., Croft, S. E., & Rowe, R. (2015). Socioeconomic status and antisocial behaviour among children and adolescents: A systematic review and meta-analysis. *Clinical Psychology Review*, 35(2), 47-55. doi:10.1016/j.cpr.2014.11.003

Emotion Recognition in Conduct Disorder

- Schwenck, C., Gensthaler, A., Romanos, M., Freitag, C. M., Schneider, W., & Taurines, R. (2014). Emotion recognition in girls with conduct problems. *European Child and Adolescent Psychiatry*, 23(1), 13–22. doi: 10.1007/s00787-013-0416-8
- Schwenck, C., Mergenthaler, J., Keller, K., Zech, J., Salehi, S., Taurines, R., & ... Freitag, C. M. (2012). Empathy in children with autism and conduct disorder: Group-specific profiles and developmental aspects. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 53(6), 651–659. doi: 10.1111/j.1469-7610.2011.02499.x
- Schurgin, M. W., Nelson, J., Iida, S., Ohira, H., Chiao, J. Y., & Franconeri, S. L. (2014). Eye movements during emotion recognition in faces. *Journal of Vision*, 14(13), doi:10.1167/14.13.14
- Short, R. M. L., Sonuga-Barke, E. J. S., Adams, W. J., & Fairchild, G. (2016). Does comorbid anxiety counteract emotion recognition deficits in conduct disorder? *Journal of Child Psychology and Psychiatry*, 57(8), 917–926. doi: 10.1111/jcpp.12544
- Spezio, M. L., Adolphs, R., Hurley, R. S. E., & Piven, J. (2007). Abnormal use of facial information in high-functioning autism. *Journal of Autism and Developmental Disorders*, 37(5), 929–939. doi: 10.1007/s10803-006-0232-9
- Sully, K., Sonuga-Barke, E. J. S., & Fairchild, G. (2015). The familial basis of facial emotion recognition deficits in adolescents with conduct disorder and their unaffected relatives. *Psychological Medicine*, 45(09), 1965–1975. doi: 10.1017/s0033291714003080
- van der Schalk, J., Hawk, S. T., Fischer, A. H., & Doosje, B. (2011). Moving faces, looking places: Validation of the Amsterdam Dynamic Facial Expression Set (ADFES). *Emotion*, 11(4), 907–920. doi: 10.1037/a0023853
- Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence*. New York, N.Y.: The

Emotion Recognition in Conduct Disorder

Psychological Corporation: Harcourt Brace & Company.

Woodworth, M., & Waschbusch, D. (2008). Emotional processing in children with conduct problems and callous/unemotional traits. *Child: Care, Health and Development*, 34(2), 234–244. doi: 10.1111/j.1365-2214.2007.00792.x

Yoshikawa, S., & Sato, W. (2008). Dynamic facial expressions of emotion induce representational momentum. *Cognitive, Affective and Behavioral Neuroscience*, 8(1), 25-31. doi:10.3758/CABN.8.1.25

Key Points

Emotion Recognition in Conduct Disorder

- Adolescents with Conduct Disorder (CD) are reported to show deficits in facial emotion recognition, but it is unclear whether these impairments reflect difficulties with attention and/or appraisal of emotional content.
- We investigated recognition of, and attention to, static and dynamic facial expressions in male and female adolescents with CD and typically-developing controls using eye tracking methods.
- Participants with CD, and males in particular, exhibited emotion recognition deficits and fixated less on the eyes when viewing facial expressions than controls.
- Across the whole sample, higher callous-unemotional traits were related to poorer fear recognition, but within the CD group, higher callous-unemotional traits were associated with better fear recognition.
- Although fixation on the eyes was positively associated with recognition performance, abnormal fixation patterns did not explain the emotion recognition deficits associated with CD.

Emotion Recognition in Conduct Disorder

Table 1. Demographic and clinical characteristics of the sample

	TD Males ¹ (n = 26)	CD Males ² (n = 26)	TD Females ³ (n = 25)	CD Females ⁴ (n = 24)		Post-hocs
	<i>M</i> (SD)				<i>F</i>	
Age (years)	16.22 (1.45)	15.94 (1.98)	16.40 (1.53)	16.21 (1.69)	.33	-
IQ	104.65 (11.39)	87.12 (7.33)	100.40 (12.64)	91.42 (16.70)	10.87***	1, 3 > 2, 4
CU traits	22.88 (6.08)	30.38 (7.56)	17.96 (6.45)	26.83 (8.95)	12.78***	1, 3 < 2; 3 < 4
	<i>n</i> (%)				χ^2	
High SES	17 (65)	5 (19)	14 (56)	10 (42)	10.77*	1, 3 > 2
Low SES	4 (15)	12 (46)	8 (32)	9 (38)		
Caucasian	21 (81)	24 (92)	24 (96)	23 (96)	4.87	-
ADHD	-	11 (42)	-	6 (25)	1.67	-
MDD	-	5 (19)	-	4 (17)	.06	-
Anxiety	-	1 (4)	-	4 (17)	2.28	-
Substance abuse	-	0 (0)	-	1 (4)	-	
Alcohol abuse	-	0 (0)	-	1 (4)	-	
PTSD	-	0 (0)	-	1 (4)	-	

Note: The presence of a current psychiatric disorder was an exclusion criterion for the TD group. Key: ADHD, attention-deficit/hyperactivity disorder; CD, Conduct Disorder; CU, callous-unemotional; IQ, intelligence quotient; MDD, major depressive disorder; PTSD, Post-Traumatic Stress Disorder; SD, standard deviation; SES, socioeconomic status; TD, typically-developing. **p* < .05, ****p* < .001.

*Emotion Recognition in Conduct Disorder*Table 2. *Simplified models for emotion categorisation and eye movement behaviour, across all emotions and for individual emotions.*

		Predictors (B)							
		CD	Gender	CU	Stimulus type	Emotional intensity	CD*CU	CD* Emotional intensity	CU* Emotional intensity
Categorisation Accuracy	All	-5.51***	5.67***	-	-	.68***	-	-	-
	Anger	-7.18*	6.83*	-	8.40**	1.01***	-	-	-
	Fear	-29.08**	-	-	-	.70***	1.00*	-	-
	Disgust	-	13.22***	-	-10.31***	.42***	-	-.14**	-
	Happiness	-	-	-	7.52***	.52***	-	-	-
	Sadness	-	6.46**	-	17.80***	.71***	-	-	-
	Surprise	-	3.47*	-	8.61***	.77***	-	-	-
	Neutral	-	7.00*	-	-	-	-	-	-
Initial Eye Preference	All	-	21.96*	-	-	-	-	-	-
	Anger	-	29.56**	-	13.57**	-	-	-	-
	Fear	-	34.41**	-	-	-	-	-	-
	Disgust	-	-	-	-11.77*	-.18**	-	-	-
	Happiness	-	-	-	-	-	-	-	-
	Sadness	-11.97**	45.00**	-	33.48***	-.33***	-	-	-
	Surprise	-	63.57***	-1.32**	58.65***	-	-	-	.01*
	Neutral	-	31.38**	-	-	-	-	-	-
Total Eye Preference	All	-	-	-	4.78***	-	-	-	-
	Anger	-	8.91*	-	4.88***	-	-	-	-
	Fear	-8.84*	9.39*	-	6.55***	.05**	-	-	-
	Disgust	-	-	-	4.68***	-	-	-	-
	Happiness	-	8.60*	-	4.32***	-	-	-	-
	Sadness	-	12.31**	-	3.92**	-.05*	-	-	-
	Surprise	-	9.13*	-	6.84***	.15***	-	-	-
	Neutral	-	9.41*	-	-	-	-	-	-

Key: *B*, unstandardised beta; CD, Conduct Disorder; CU, callous-unemotional; Initial eye preference, the percentage of trials in which the participant fixated on the eye region first, minus the percentage of trials in which they fixated the mouth region first; Total eye preference, the percentage of overall trial time spent fixating the eye region, minus that spent fixating the mouth. * $p < .05$, ** $p < .01$, *** $p < .001$.

Figure 1.

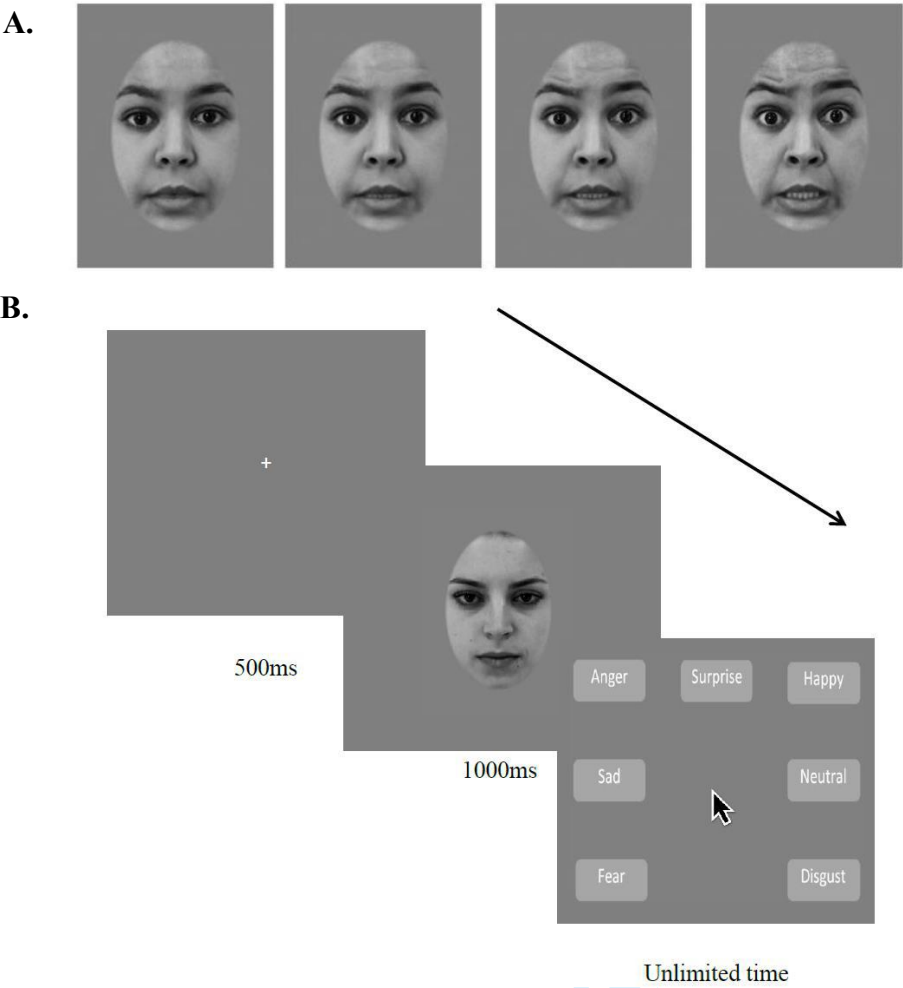


Figure 2.

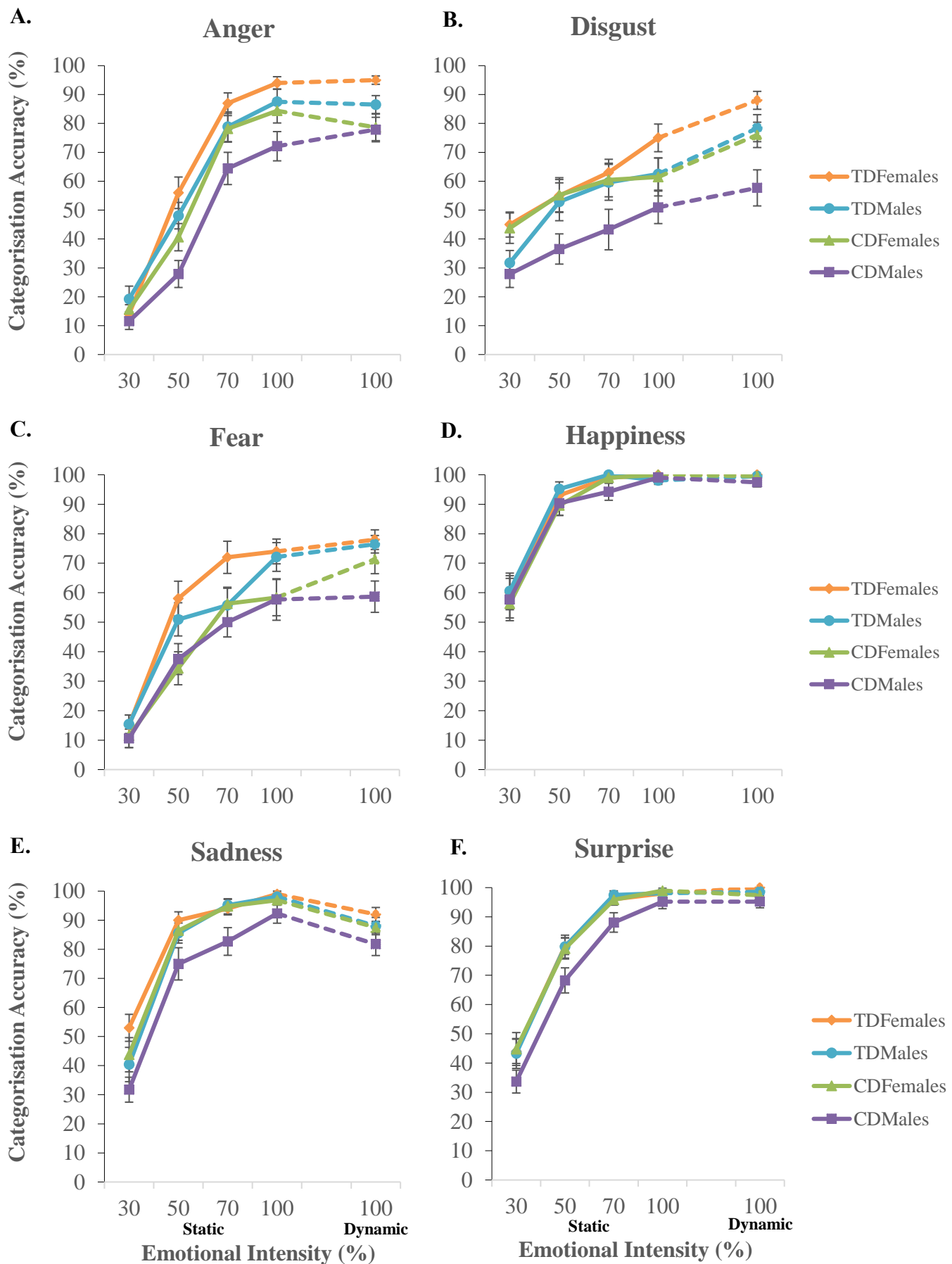
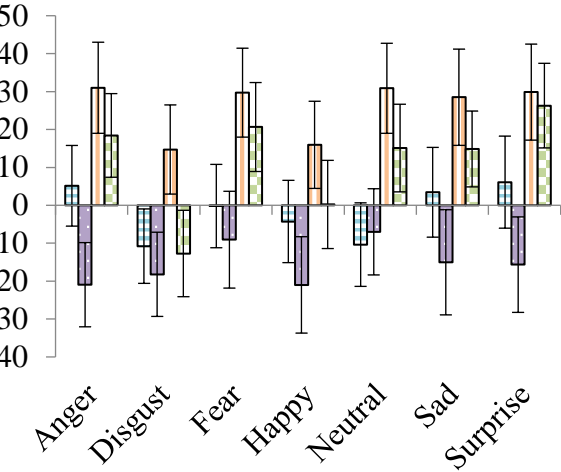
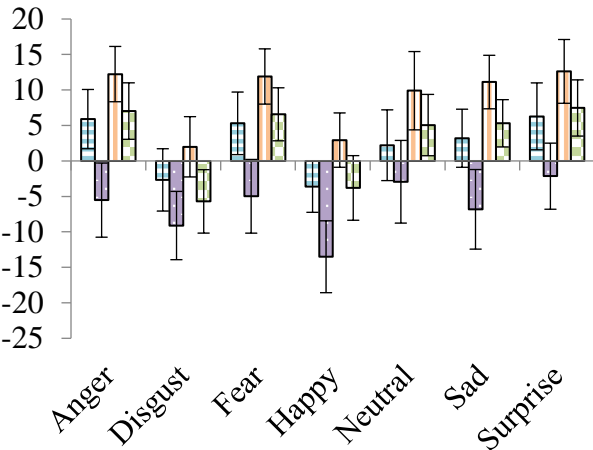


Figure 3.

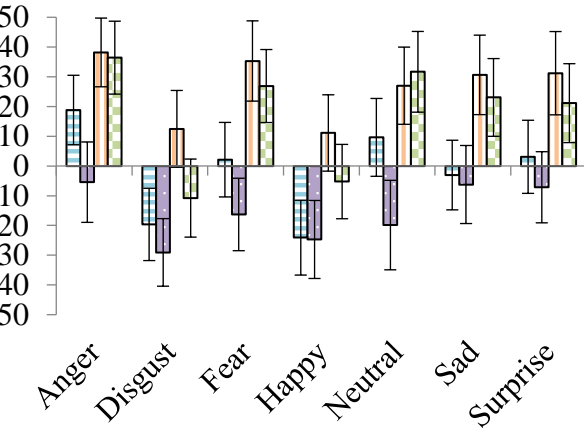
A. Dynamic Expressions: Initial Eye Preference



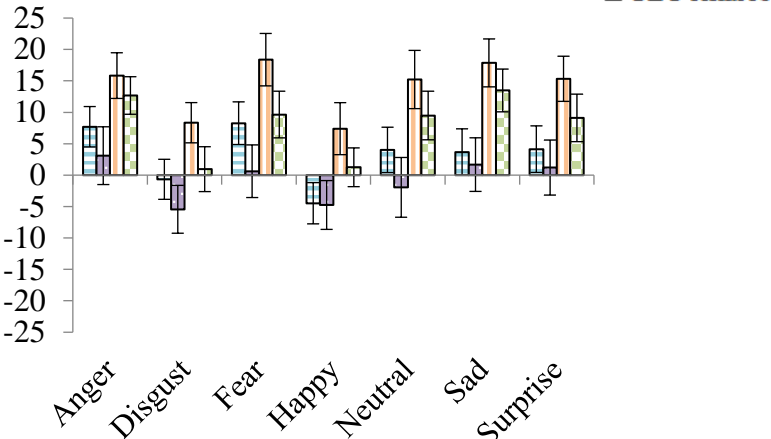
B. Dynamic Expressions: Total Eye Preference



C. Static Expressions: Initial Eye Preference



D. Static Expressions: Total Eye Preference



Emotion

Supplementary Materials for Martin-Key et al. - *Facial Emotion Recognition and Eye Movement Behaviour in Conduct Disorder*

Data Analytic Strategy:

To create parsimonious models of the data, we followed the approach suggested by Bates, Kliegl, Vasishth, and Baayen (2015), in which the initial model includes as many random and fixed effects as are supportable by the data. This was followed by a stepwise backward elimination process to remove factors that did not significantly improve the model's fit to the data. The initial model included subject, ADHD, mood, anxiety and age as random (intercept) factors (note that PTSD, alcohol and substance abuse were present in one subject only, and thus these variables are subsumed by the subject factor). In addition, for analyses collapsed across emotional expressions, emotion was included as a random factor. By-subject and by-emotion variation in the repeated measured factors (fixed effects) were also included as random effects (slopes). CD, gender, CU traits, SES, IQ, stimulus type (static vs. dynamic) and emotional intensity were included as fixed effects, in addition to all two-way interactions between these variables.

Models were evaluated using the 'fitlme' function in Matlab. After fitting the most complex model, random factors that did not significantly improve to the model were iteratively removed, with higher-order factors (i.e., slopes) removed before lower-order factors. The contribution of each candidate factor was evaluated using a likelihood ratio test (Matlab's 'compare' function) to compare the more complex model with a reduced (nested) model excluding the candidate factor. The model was simplified following a non-significant decrease in goodness of fit ($p > .05$); we note that inclusion/exclusion based on coefficient t -tests is not advised in mixed-effects models (Baayen, Davidson & Bates, 2008; Bates et al., 2015). Fixed factors were then iteratively removed via the same process. However, because IQ and SES differed significantly between participant groups, we followed the conservative approach of including these (potentially confounding) factors in the final model, even when their effects were non-significant.

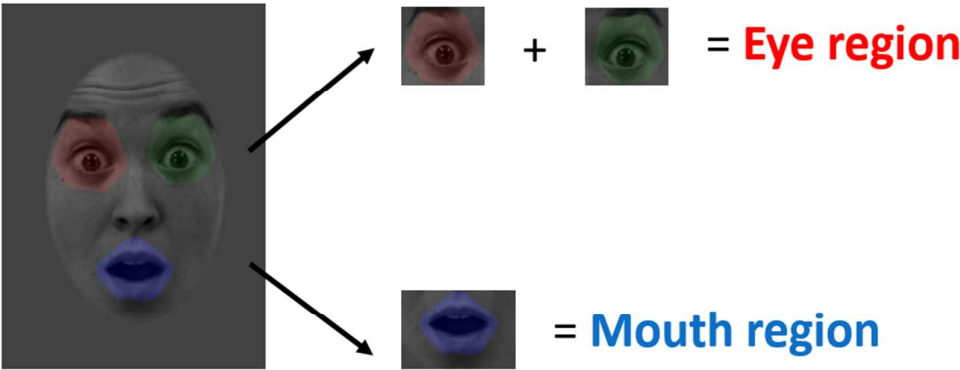


Figure S1. An example static stimulus. Bespoke software was used to manually select three regions of interest (ROIs) for each of our face stimuli: two around the eyes (comprising the eye region) and one around the mouth (mouth region). ROIs for dynamic stimuli were created on a frame-by-frame basis.

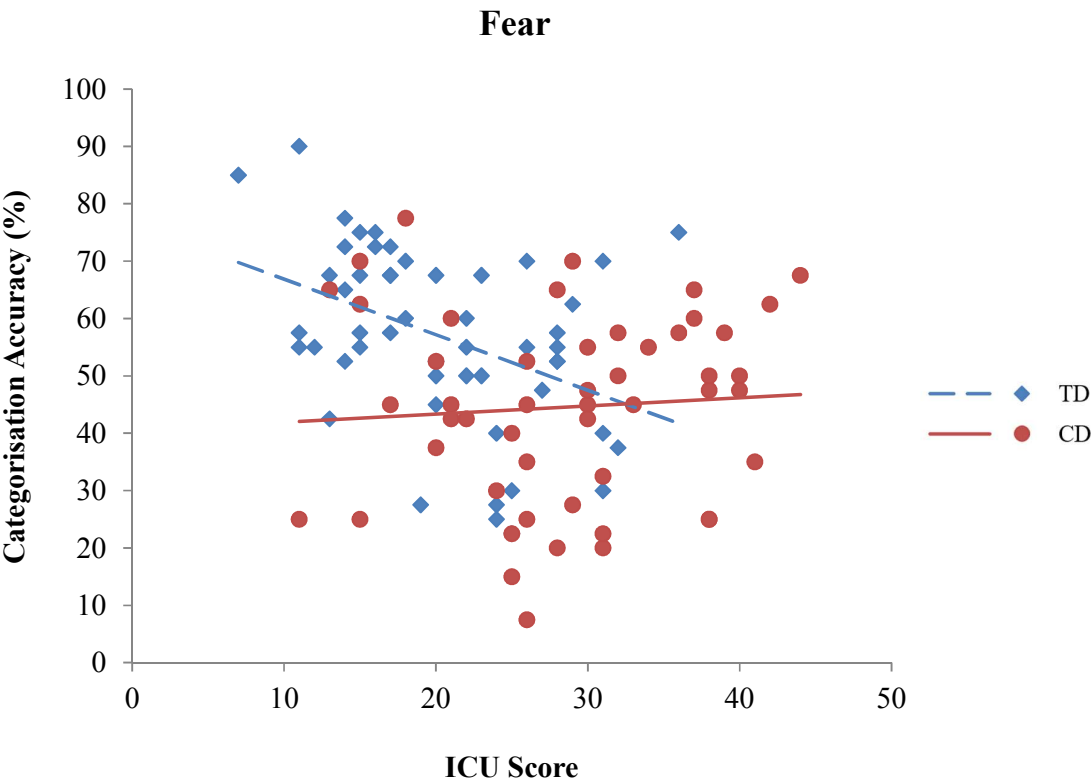


Figure S2. Fear categorisation accuracy as a function of Inventory of Callous-Unemotional traits (ICU) score, averaged across stimulus type (static and dynamic). Each symbol represents a single participant and lines show the linear best fits to the data from participants with CD and typically-developing controls. *Note:* CD, conduct disorder; TD, typically-developing.

Table S1. Full models for facial emotion categorisation and eye-movement behaviour, across all emotions and for individual emotions.

		CD	Gender	CU	IQ	SES	Stimulus type	Emotion intensity	CD*CU	CD* Intensity	IQ*SES	Gender* Stimulus type	IQ* Emotion intensity	CU* Emotion intensity
Categorisation Accuracy	All	-5.51***	5.67***	-	.07	-3.37	-	.68***	-	-	-	-	-	-
	Anger	-7.18*	6.83*	-	.24*	-3.52	8.40**	1.01***	-	-	-	-	-	-
	Fear	-29.08**	-	-1.02**	.29*	3.54	-	.70***	1.00*	-	-	-	-	-
	Disgust	-	13.22***	-	-1.40**	-99.90**	-10.31***	.42***	-	-.14**	.99**	-	-	-
	Happiness	-	-	-	.01	-2.76	7.52***	.52***	-	-	-	-	-	-
	Sadness	-	6.46**	-	.03	-7.98**	17.80***	.71***	-	-	-	-	-	-
	Surprise	-	3.47*	-	.07	-5.72**	8.61***	.77***	-	-	-	-	-	-
	Neutral	-	7.00*	-	.08	-1.45	-	-	-	-	-	-	-	-
Initial Eye Preference	All	-	21.96*	-	-.07	-13.57	-	-	-	-	-	-	-	-
	Anger	-	29.56**	-	-.13	-20.30	13.57**	-	-	-	-	-	-	-
	Fear	-	34.41**	-	-.43	-24.11	-	-	-	-	-	-	-	-
	Disgust	-	-	-	.09	-10.96	-11.77*	-.18**	-	-	-	-	-	-
	Happiness	-	-	-	-.18	-19.11	-	-	-	-	-	-	-	-
	Sadness	-11.97**	45.00**	-	-.13	-6.48	33.48***	-.33***	-	-	-	-19.42*	-	-
	Surprise	-	63.57***	-1.32**	.64***	2.77	58.65***	-	-	-	-	-30.46***	-.004**	.01*

Total Eye Preference	Neutral	-	31.38**	-	-.04	-14.32	-	-	-	-	-	-	-	-
	All	-	-	-	-.04	-7.31	4.78***	-	-	-	-	-	-	-
	Anger	-	8.91*	-	-.04	-5.86	4.88***	-	-	-	-	-	-	-
	Fear	-8.84*	9.39*	-	-.13	-3.16	6.55***	.05**	-	-	-	-	-	-
	Disgust	-	-	-	-.02	-4.96	4.68***	-	-	-	-	-	-	-
	Happiness	-	8.60*	-	-.001	-5.99	4.32***	-	-	-	-	-	-	-
	Sadness	-	12.31**	-	-.13	-6.23	3.92**	-.05*	-	-	-	-	-	-
	Surprise	-	9.13*	-	-.04	-5.14	6.84***	.15***	-	-	-	-	-	-
	Neutral	-	9.41*	-	.01	7.09	-	-	-	-	-	-	-	-

Key: *B*, unstandardised coefficient; CD, Conduct Disorder; CU, callous-unemotional traits; Initial eye preference, the percentage of trials in which the participant moved from the fixation cross to fixate on the eyes first, minus the percentage of trials in which they moved to fixate the mouth first; Total eye preference, the percentage of trial time spent fixating the eye region minus the percentage of time spent fixating the mouth. **p* < .05, ***p* < .01, ****p* < .001 as determined by likelihood ratio tests.

Table S2. *Improved models of categorisation accuracy when including measures of eye movement behaviour as predictors.*

Categorisation Accuracy		Predictors (<i>B</i>)												
		CD	Gender	CU	CD*CU	IQ	SES	Initial eye preference	Total eye preference	Stimulus type	Emotional intensity	Total eye preference* Stimulus type	Initial eye preference* Emotional intensity	Total eye preference* Emotional intensity
	All	-5.51 ^{***}	5.65 ^{***}	-	-	0.07	-3.35	-.09 ^{***}	.22 ^{***}	-	.68 ^{***}	-	.001 ^{***}	-.003 ^{***}
	Anger	-6.68 [*]	5.59 [*]	-	-	.26 [*]	-2.87	-	-.69 [*]	5.59 [*]	.97 ^{***}	.27 [*]	-	.005 [*]
	Fear	-29.63 ^{**}	-	-1.03 ^{**}	1.01 [*]	.28 [*]	3.22	-.10 [*]	-	-	.69 ^{***}	-	.001 ^{**}	-
	Happiness	-	-	-	-	0.01	-2.31	-	.31 ^{**}	7.53 ^{***}	.52 ^{***}	-	-	-.003 [*]
	Sadness	-	7.20 ^{**}	-	-	0.03	-8.44 ^{**}	-.33 ^{***}	-	15.28 ^{***}	.49 ^{***}	-	.003 ^{***}	-

Key: *B*, unstandardised coefficient; CD, Conduct Disorder; CU, callous-unemotional traits; Initial eye preference, the percentage of trials in which the participant moved from the fixation cross to fixate on the eyes first, minus the percentage of trials in which they fixated the mouth first; Total eye preference, the percentage of overall trial time spent fixating the eye region minus the percentage of time spent fixating the mouth. * $p < .05$, ** $p < .01$, *** $p < .001$.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

Effect size measures

Within the manuscript, we have reported simple effect sizes, as quantified using Cohen’s *d* (Cohen, 1988). This measure can be used, for example, to understand the difference in categorisation performance between those with and without a CD diagnosis. However, in the context of multi-predictor models, we may wish to understand the contribution of CD status in predicting categorisation performance, given that other variables such as IQ might account for some of the differences between these two groups. In other words, we might wish to quantify the unique contribution of CD status, after accounting for variability in emotion categorisation that is associated with variables such as IQ, SES, and gender. We can use Cohen’s *f*-squared (f^2) to quantify the effect size that is associated with a single predictor (or subset of predictors) within the context of these other variables (see Selya, Rose, Dierker, Hedeker, & Mermelstein et al., 2012). Note that the interpretation of the magnitude of Cohen’s f^2 differs from that of Cohen’s *d*. According to Cohen (1988), $f^2 \geq .02$, $f^2 \geq .15$, and $f^2 \geq .35$ represent small, medium, and large effect sizes, respectively. Table S3 shown below provides this ‘contextual’ effect size measure for key predictors (CD, gender, CU traits, and their interactions) within models of emotion categorisation, eye movement behavior, and finally, the effect size associated with eye movement parameters when added to the best-fitting models of emotion categorisation.

Table S3. *Effect sizes (expressed as Cohen's f -squared) associated with key predictor variables, in the context of multi-predictor models.*

Dependent Variable	Emotion	Predictor(s)	Cohen's f^2
Emotion categorisation	All emotions	CD	.01
		Gender	.02
	Anger	CD	.03
		Gender	.03
	Fear	CU, CD*CU	.03
	Disgust	CD*Intensity	.08
		Gender	.18
	Sad	Gender	.03
	Surprise	Gender	.01
	Neutral	Gender	.12
Initial eye preference	All emotions	Gender	.07
	Anger	Gender	.24
	Fear	Gender	.29
	Sadness	CD	.02
		Gender, Gender*Stimulus Type	.07
	Surprise	CU, CU*Intensity	.02
		Gender, Gender*Stimulus Type	.05
Total eye preference	Anger	Gender	.25
	Fear	CD	.17
		Gender	.26

Emotion categorisation (with eye preference scores considered as predictors)	Happiness	Gender	.20
	Sadness	Gender	.38
	Surprise	Gender	.24
	Neutral	Gender	.26
	All emotions	Initial eye preference, Initial*Intensity, Total eye preference, Total*Intensity	.01
	Anger	Total eye preference, Total*Intensity, Total*Stimulus type	.04
	Fear	Initial eye preference, Initial*Intensity	.02
	Happiness	Total eye preference, Total*Intensity	.03
	Sadness	Initial eye preference, Initial*Intensity	.04

Key: CD, Conduct Disorder; Cohen's f^2 , effect size measure in the context of multi-predictor models (where $f^2 \geq .02$, $f^2 \geq .15$, and $f^2 \geq .35$ represent small, medium, and large effect sizes, respectively); CU, callous-unemotional traits; Initial eye preference, the percentage of trials in which the participant moved from the fixation cross to fixate on the eyes first, minus the percentage of trials in which they fixated the mouth first; Total eye preference, the percentage of overall trial time spent fixating the eye region minus the percentage of time spent fixating the mouth.

References:

Baayen, R., Davidson, D., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(Special Issue: Emerging Data Analysis), 390-412. doi:10.1016/j.jml.2007.12.005

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). *Parsimonious Mixed Models*. ArXiv e-prints.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, N.J.: L. Erlbaum Associates.

Selya, A., Rose, J., Dierker, L., Hedeker, D., & Mermelstein, R. (2012). A practical guide to calculating Cohen's f^2 , a measure of local effect size, from PROC MIXED. *Frontiers in Psychology*, 3(4), doi:10.3389/fpsyg.2012.00111