**Positive and Negative Parenting in Conduct Disorder with High versus Low Levels of Callous-Unemotional Traits: Supplementary Materials**

**Correlations between CD Symptoms, CU Traits and Parenting**

Correlations between parenting (after regressing out variance associated with IQ, sex, pubertal status, SES and site of data collection), CD symptoms and CU traits are presented in **Table S1** (CD group) and **Table S2** (TD group). In the CD group, both CD symptoms and CU traits were negatively correlated with parental involvement and parental reinforcement, and positively correlated with poor supervision, inconsistent discipline and corporal punishment, although these correlations were not always significant. It is notable that CU traits were not significantly correlated with either measure of negative parenting, although they were positively correlated with both measures of positive parenting as well as poor supervision. In the TD group, CU traits were likewise negatively correlated with both measures of positive parenting, and positively correlated with poor supervision and inconsistent discipline. As expected, given the lack of variation in CD symptoms in this group, there were no significant correlations between CD symptoms and any measure of parenting.

**Table S1.** Pearson correlations between CD symptoms, CU traits and parenting for youths with CD

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1.  | 2.  | 3.  | 4.  | 5. | 6. | 7. |
| 1. CD symptoms  |  - | 0.18\* | -0.16\* | -0.04 | 0.28\* | 0.15\* | 0.09 |
| 2. CU traits |  | -  | -0.18\* | -0.11\* | 0.16\* | 0.07 | 0.06 |
| 3. Parental involvement |  |  | - | 0.65\* | -0.30\* | -0.13\* | -0.09 |
| 4. Positive reinforcement |  |  |  | - | -0.24\* | -0.08 | -0.13\* |
| 5. Poor supervision |  |  |  |  | - | 0.38\* | 0.19\* |
| 6. Inconsistent discipline |  |  |  |  |  | - | 0.25\* |
| 7. Corporal punishment  |  |  |  |  |  |  | - |

Notes: CD = conduct disorder, CU = callous-unemotional traits

**\*** *p* < .05, 2-tailed

**Table S2.** Pearson correlations between CD symptoms, CU traits and parenting for TD youths

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1.  | 2.  | 3.  | 4.  | 5. | 6. | 7. |
| 1. CD symptoms  |  - | 0.04 | -0.05 | -0.07 | 0.09 | 0.05 | 0.09 |
| 2. CU traits |  | -  | -0.22\* | -0.32\* | 0.11\* | 0.22\* | 0.06 |
| 3. Parental involvement |  |  | - | 0.58\* | -0.39\* | -0.17\* | -0.10\* |
| 4. Positive reinforcement |  |  |  | - | -0.31\* | -0.13\* | -0.16\* |
| 5. Poor supervision |  |  |  |  | - | 0.31\* | 0.20\* |
| 6. Inconsistent discipline |  |  |  |  |  | - | 0.25\* |
| 7. Corporal punishment  |  |  |  |  |  |  | - |

Notes: TD = typically developing, CD = conduct disorder, CU = callous-unemotional traits.

**\*** *p* < .05, 2-tailed

**Imputation of Missing Data**

Missing data for the PDS were imputed separately, before the decision was made to impute missing values for other measures. The procedure for the PDS imputation is thus described separately from the other measures. Missing PDS values were imputed based on the full FemNAT-CD sample. Missing data in a multi-item instrument are best handled by imputation at the item level (Eekhout, de Vet, Twisk, Brand, de Boer, & Heymans, 2014). Consequently, missing values of single items were imputed first and their scores were calculated based on the imputed items. The imputation was done in SAS® version 9.4 using the procedure PROC MI. Imputation by fully conditional specification (FCS) was used, which offers a flexible method to specify the multivariate imputation model for arbitrary missing patterns, including both categorical and continuous variables (Liu & De, 2015). As the items were measured at an ordinal level, the logistic regression method was specified in the FCS statement. For imputation diagnostics, distributions of the observed and imputed items and scores were checked. The imputation of the PDS items was done separately in males and in females due to the presence of sex specific items. Variables included in the imputation model were: sex specific items of the PDS, remaining PDS items, age at collection of PDS data and age at informed consent. These were used to impute age at PDS (if missing), CD/TD status and site. Imputation for the remaining measures was conducted separately, following the same procedure as above. The following variables were included in the imputation model: all items of the respective questionnaire, age, IQ, group (CD/TD), sex, site, comorbidities (PTSD, ADHD, ODD, depression, anxiety), and items of other questionnaires if correlated with at least one of the items with ≥ 0.4. For imputation diagnostics, distribution of the observed and imputed items and scores were checked. For the parent-report APQ (Essau et al., 2006a) missing items were only imputed if at least one item was present.

**Repetition of Angle-GMLVQ Classification Analyses after Regressing Out Variance Associated with SES**

Angle-GMLVQ classification analyses were repeated after regressing out variance associated with SES, where these data were available (157 CD/HCU, 159 CD/LCU and 421 TD). As shown in **Table S3**, performance was similar, albeit slightly poorer, to the main Angle-GMLVQ classifier results.

Feature relevance scores are displayed in **Figure S1**. Again, the pattern of feature relevance scores in terms of positive and negative parenting was similar to the main classifier results. The only substantial change in feature relevance was for the HCU-TD model, where the relevance of corporal punishment and poor supervision was higher than in the main models, while the relevance of other features decreased somewhat. In summary, accounting for SES did not substantially affect the pattern of results observed.



**Figure S1.** Feature relevance for a) HCU-TD model, b) LCU-TD model and c) HCU-LCU models after accounting for SES. Bars show percentage of re-samplings in which feature relevance was in the top 20% of relevance scores across all re-samplings with MCER ≤ 0.40

**Table S3.** Angle-GMLVQ model performance after accounting for SES (mean (95% confidence intervals of the mean))

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Mixed-TD  | HCU-TD  | LCU-TD  | F (*p)*, partial η2 | HCU-LCU  |
| Accuracy |  0.69(0.69, 0.69) | 0.73(0.72, 0.73) | 0.66(0.66, 0.67) | 504.10 (*<.001*), .25 | 0.55(0.55, 0.55) |
| PPV | 0.70(0.70, 0.71) | 0.58(0.58, 0.58) | 0.50(0.50, 0.50) | 3593.00 (*<.001*), .71 | 0.55(0.55, 0.55) |
| NPV | 0.68(0.68, 0.68) | 0.82(0.82, 0.82) | 0.78(0.78, 0.78) | 3410.06 (*<.001*), .70 | 0.55(0.55, 0.56) |
| TPR | 0.66(0.65, 0.66) | 0.67(0.67, 0.70) | 0.60(0.59, 0.60) | 235.01 (*<.001*), .14 | 0.56(0.55, 0.56) |
| TNR | 0.72(0.72, 0.73) | 0.75(0.75, 0.75) | 0.70(0.69, 0.70) | 208.10 (*<.001*), .12 | 0.54(0.54, 0.55) |
| MCER | 0.31(0.31, 0.31) | 0.29(0.28, 0.29) | 0.35(0.35, 0.35) | 521.72 (*<.001*), .26 | 0.45(0.45, 0.45) |

Notes: Mixed-TD = model classifying youths with conduct disorder with mixed levels of callous unemotional traits and typically developing youths, HCU-TD = model classifying youths with conduct disorder with high levels of callous unemotional traits and typically developing youths, LCU-TD = model classifying youths with conduct disorder and low levels of callous-unemotional traits and typically developing youths. PPV = positive predictive value, NPV = negative predictive value, TPR = true positive rate, TNR = true negative rate, MCER = macro-averaged classification error rate. Groups with different superscript indices differ significantly in post-hoc comparisons (*p* < .05, Bonferroni corrected). Note that the HCU-LCU model (column 6) was not included in statistical tests as comparisons between this and other models were not relevant to hypotheses

**SVM Analyses**

Classification analyses were repeated with SVM classifiers to ensure that classifier performance was broadly similar across different classifier types. SVM classifiers were trained and tested using the standard MATLAB function ‘fitcsvm’. To optimise parameters, we trained and tested six classifiers for each model; a linear SVM and SVMs with second, third, fourth, fifth and sixth-order non-linear polynomial transformation kernels. We then selected the SVM with the lowest MCER for comparison with the corresponding Angle-GMLVQ classifier.

Of all the SVM classifiers tested, the linear SVM performed best for each model. MCER values for the linear SVMS were 0.28 for the Mixed-TD model, 0.25 for the HCU-TD model, 0.32 for the LCU-TD model and 0.42 for the HCU-LCU model. For the Mixed-TD, HCU-TD and LCU-TD models the linear SVM classifier outperformed the corresponding Angle-GMLVQ model by a very small but statistically significant margin (Mixed-TD: t (1998) = 2.90, *p* =.004, partial η2 = .004. HCU-TD: t (1998) = 5.04, *p* < .001, partial η2 = .01. LCU-TD: t (1998) = 3.95, *p* < .001, partial η2 = 0.01). For the HCU-LCU model, the two classifiers did not differ significantly (t (1998) = 1.30, *p* =.19, partial η2 = .001). These differences in performance indicate that the absolute magnitudes of differences in parenting scores between participants were generally very slightly more informative than the pattern of differences alone.

**Sex Differences**

Parenting scores per group are reported separately for females (**Table S4**) and males (**Table S5**). Main effects of sex and sex by group interactions were investigated for each parenting measure separately. There were no main effects of sex for parental involvement, positive reinforcement, poor supervision or inconsistent discipline. (Parental involvement: F (1, 750) = 0.29, *p* = .59. Positive reinforcement: F (1, 750) = 1.24, *p* = .27. Poor supervision: F (1, 750) = 2.45, *p* = .12. Inconsistent discipline: F (1, 750) = 0.04, *p* = .85). However, there was a main effect of sex on corporal punishment, with males experiencing significantly more corporal punishment than females (F (1, 750) = 2.31, *p* = .10).

There was no significant sex by group interaction for parental involvement (F (2, 750) = 5.61, *p* = .02; see **Figure S2**). For positive reinforcement, there was a significant interaction (F (2, 750) = 3.05, *p* = .048), but pairwise comparisons did not reveal significant differences between males and females within any group (CD/HCU: *p* = .15, CD/LCU: *p* = .07, TD: *p* = .43; see **Figure S3**). The same was true for poor supervision (F (2, 750) = 4.92, *p* = .01. CD/HCU: *p* = .06, CD/LCU: *p* = .09, TD: *p* = .06; see **Figure S4**). For inconsistent discipline, TD males scored significantly higher than TD females (F (2, 750) = 3.98, *p* = .02. CD/HCU: *p* = .14, CD/LCU: *p* = .68, TD: *p* = .01; see **Figure S5**). Finally, for corporal punishment, males in the CD/LCU group scored significantly higher than females (F (2, 750) = 5.00, *p* = .01. CD/HCU: *p* = .27, CD/LCU: *p* = .02, TD: *p* = .37; see **Figure S6**).In summary, there was a tendency for males to experience higher levels of negative parenting than females, but these differences were small and not consistent across measures and groups.

**Table S4.** Group differences in parenting for females only (mean (95% confidence intervals of the mean))

|  |  |  |  |
| --- | --- | --- | --- |
| APQ subscales | CD/HCU (*n* = 89) | CD/LCU (*n* = 86) | TD (*n* = 261) |
| *Positive parenting* |
| Parental involvement | 24.10 (22.93, 25.27) | 27.16 (25.63, 28.70) | 31.28 (30.68, 31.87) |
| Positive reinforcement | 17.27 (16.28, 18.26) | 18.34 (17.21, 19.46) | 21.07 (20.61, 21.52) |
| Poor supervision | 31.37 (29.85, 32.89) | 28.90 (27.11, 30.68) | 21.93 (21.15, 22.71) |
| *Negative parenting* |
| Inconsistent discipline | 19.91 (19.09, 20.73) | 18.97 (18.03, 19.90) | 15.46 (15.04, 15.87) |
| Corporal punishment | 5.58 (4.98, 6.19) | 5.65 (5.07, 6.23) | 3.89 (3.70, 4.07) |

Notes: CD/HCU = conduct disorder with high levels of callous-unemotional traits, CD/LCU = conduct disorder with low levels of callous-unemotional traits, TD = typically developing. APQ = Alabama Parenting Questionnaire

**Table S5.** Group differences in parenting for males only (mean (95% confidence intervals of the mean))

|  |  |  |  |
| --- | --- | --- | --- |
| APQ subscales | CD/HCU (*n* = 89) | CD/LCU (*n* = 86) | TD (*n* = 261) |
| *Positive parenting* |
| Parental involvement | 24.93 (23.24, 26.63) | 28.08 (26.47, 29.69) | 24.93 (23.24, 26.63) |
| Positive reinforcement | 17.81 (16.67, 18.95) | 19.60 (18.51, 20.69) | 20.44 (19.80, 21.07) |
| Poor supervision | 29.27 (27.51, 31.02) | 27.03 (25.23, 28.82) | 23.23 (22.25, 24.21) |
| *Negative parenting* |
| Inconsistent discipline | 19.07 (18.23, 19.91) | 18.73 (17.94, 19.52) | 16.37 (15.87, 16.87) |
| Corporal punishment | 5.23 (4.66, 5.79) | 4.65 (4.21, 5.09) | 4.07 (3.82, 4.32) |

Notes: CD/HCU = conduct disorder with high levels of callous-unemotional traits, CD/LCU = conduct disorder with low levels of callous-unemotional traits, TD = typically developing. APQ = Alabama Parenting Questionnaire



**Figure S2.** Sex differences per group for parental involvement



**Figure S3.** Sex differences per group for positive reinforcement



**Figure S4.** Sex differences per group for poor supervision



**Figure S5.** Sex differences per group for inconsistent discipline



**Figure S6.** Sex differences per group for corporal punishment

**Supplementary References**

Eekhout, I., de Vet, H. C., Twisk, J. W., Brand, J. P., de Boer, M. R., & Heymans, M. W. (2014). Missing data in a multi-item instrument were best handled by multiple imputation at the item score level. *Journal of Clinical Epidemiology*, *67*(3), 335-342. doi: https://doi.org/10.1016/j.jclinepi.2013.09.009

Essau, C. A., Sasagawa, S., & Frick, P. J. (2006a). Psychometric properties of the Alabama parenting questionnaire. *Journal of Child and Family Studies,* *15*(5), 595-614. doi: https://doi.org/10.1007/s10826-006-9036-y

Liu, Y., & De, A. (2015). Multiple imputation by fully conditional specification for dealing with missing data in a large epidemiologic study. *International Journal of Statistics in Medical Research*, *4*(3), 287. doi: 10.6000/1929-6029.2015.04.03.7