Supplementary Online Content


eMethods 1. Additional Methodological Details Related to the Meta-analysis Procedure

eTable 1. Computed CU traits Percent of Maximum Possible Scores for Youths With CP

eMethods 2. Computation of the CU Traits Percent of Maximum Possible Scores

eFigure 1. Meta-regression Results Showing Association Between Age Range and Quadratic Age With Gray Matter Differences Overlaid Onto the Main Meta-analysis Results (A) and Meta-regression Plots (B)
eFigure 2. Meta-regression Results Showing Association Between IQ Difference and Gray Matter Differences Overlaid Onto the Main Meta-analysis Results (A) and Meta-regression Plot (B)
eFigure 3. Meta-regression Results Showing Association Between Lifetime CD Symptom Severity Score and Gray Matter Differences Overlaid Onto the Main Meta-analysis Results (A) and Meta-regression Plot (B)
eTable 2. Age Range (Years) Across Studies and Mean Quadratic Age (Age Squared) in Youths With CP

eTable 3. Group Comparison on IQ Measures, IQ Group Difference Score, and Inclusion of IQ as a Covariate Across Studies

eTable 4. The Diagnostic Instruments Used to Assess CD Symptoms Across Studies and Mean Lifetime CD Symptom Severity Score

eMethods 3. Meta-regression With Age Range and Mean Age in Youths With CP

eMethods 4. Meta-regression With IQ Group Difference

eTable 5. Excluded Studies Based on Selection Criteria

eTable 6. Results of the Jack-knife Reliability Analyses of the Main Meta-analysis Findings

eTable 7. Results of the Jack-knife Reliability Analyses of the Subgroup Childhood-Onset CP vs TD Youths Meta-analysis

eFigure 4. Meta-regression Results Showing an Association Between the Proportions of Male:Female Youths With CP and GMV (A) and an Association Between CU Percent of Maximum Possible Scores and GMV (B)
eReferences

This supplementary material has been provided by the authors to give readers additional information about their work.
eMethods 1. Additional Methodological Details Related to the Meta-analysis Procedure

Conversion of the z-values to uncorrected p-values using Seed-based d-Mapping (formerly signed differential mapping; SDM) (http://www.sdmproject.com/) does not conform to traditional transformation conventions. In their paper describing anisotropic effect –size SDM, Radua et al., (2012)1 noted on p. 607: “Despite the fact that meta-analytic estimates are z values, assessment of their statistical significance is not straightforward.” This is primarily due to the fact that the p-values of SDM Z-scores are usually much different to the p-values associated with standard z-scores and are computed using randomisations. To ensure stability, we carried out 20 randomizations in total, as per the recommendation in the SDM tutorial and by Joaquim Radua (17/04/2015), developer of SDM. Pre-processing of studies consists of creating a map of variances (i.e. a measure of effect-size, with each study weighted depending on the size of its sample as well as within-study variability and between-study heterogeneity).

Where a statistical parametric map (or t-map) is available, these are converted to unbiased effect size and variance maps using Hedge’s d.2 For peak coordinates, the re-creation is based on converting the peak t-value to Hedge’s effect size and then applying a non-normalized Gaussian kernel to the voxels close to the peak (see1,3). When a statistical map is not available the effect size can only be extracted for those voxels containing a peak whilst it must be estimated for the neighbouring voxels. This estimation is conducted by assigning an effect size to each voxel depending on its distance to close peaks1,4. We used a full-width half-maximum (FWHM) of 20mm based on empirical evidence1,3 that is in line with other meta-analytical tools (e.g. ALE and MKDA)5. The meta-analytic use of null effect sizes when pre-processing voxels far from a peak voxel biases the z-values towards 0, making a normal distribution of z-values inappropriate. Previous tests based on randomizing the location of peaks is also not appropriate in combination with the use of raw statistical parametric maps1,4, as is the case in our meta-analysis (85% statistical maps included; 11 out of 13 studies). The AES-SDM method1 used here solves this issue by randomizing the location of the voxels within the standard SDM grey matter template, assuming that effect sizes (rather than only peaks) are randomly distributed throughout the brain.

Linear meta-regression analysis

Based on Monte Carlo randomizations (N = 20) used to compute statistical significance, simple linear regressions were carried out, weighted by the square root of the sample size and restricted to predict SDM values within the observed range of the computed variable (i.e. from -1 to 1). The main output for each variable indicates the regression slope (e.g. the amount of grey matter change per unit increase in mean CU trait score), with clusters reported that show a significant trend across youths with CP along with a predicted significant difference with TD youths (e.g. a reported grey matter difference between youths with CP and TD youths in studies measuring CU traits).
**eTable 1. Computed CU traits Percent of Maximum Possible Scores for Youths With CP**

<table>
<thead>
<tr>
<th>Studies</th>
<th>CU measure and score</th>
<th>CP Youths CU traits (Percent of Maximum Possible) score&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>De Brito et al.,&lt;sup&gt;6&lt;/sup&gt; 2009</td>
<td>APSD CU scale 7.9</td>
<td>65.83</td>
</tr>
<tr>
<td>Fairchild et al.,&lt;sup&gt;7&lt;/sup&gt; 2011</td>
<td>YPI CU scale 0.74 &amp; ICU total 34.2</td>
<td>60.75</td>
</tr>
<tr>
<td>Fairchild et al.,&lt;sup&gt;8&lt;/sup&gt; 2013</td>
<td>YPI CU scale 0.62</td>
<td>62</td>
</tr>
<tr>
<td>Cope et al.,&lt;sup&gt;9&lt;/sup&gt; 2014</td>
<td>PCL-YV Factor 1 9.8 &amp; ICU total 30</td>
<td>51.46</td>
</tr>
<tr>
<td>Sebastian et al.,&lt;sup&gt;10&lt;/sup&gt; 2015</td>
<td>ICU total 52.5</td>
<td>72.92</td>
</tr>
</tbody>
</table>

Abbreviations: APSD = Antisocial Process Screening Device; CP = conduct problems; TD = typically developing; CU = callous-unemotional; PCL-YV = Psychopathy Checklist – Youth Version; YPI = Youth Psychopathy Inventory

<sup>a</sup> Calculation of Percent of Maximum Possible scores: *Percent of maximum possible = ((variable name – minimum_score)/(maximum_score – minimum_score)) * 100
eMethods 2. Computation of the CU Traits Percent of Maximum Possible Scores

The measures of CU scores differed across the five studies that assessed those traits (eTable 1). The CU Percent of Maximum Possible scores calculated for each study were based on the CU scale of the Youth Psychopathy Inventory (YPI) (studies\(^7\))\(^9\), of the Antisocial Process Screening Device\(^12\) (study\(^10\)), and the total score of the Inventory of Callous-Unemotional Traits (ICU)\(^13\) (study\(^10\)). For the study by Fairchild et al. (2011)\(^7\) we created an averaged Percent of Maximum Possible score using the total ICU score and the CU scale of the YPI. For the study by Cope et al. (2014)\(^9\) we created an averaged Percent of Maximum Possible score using the CU scale of the YPI Factor 1 score of the Psychopathy Checklist – Youth Version\(^14\). For each of those measures of psychopathic traits, we decided to focus on CU traits only because it is the dimension of psychopathic traits in youths that has recently been included as a specifier for CP in the DSM 5\(^15\). This is based on two decades of research showing that, in contrast to other dimensions of psychopathic traits in youths (e.g., impulsivity, narcissism), CU traits offer incremental validity in identifying a particularly severe group of youths with CP\(^16\). The results of the meta-regression using the CU Percent of Maximum Possible Scores are presented in eFigure 4B.
eFigure 1. Meta-regression Results Showing Association Between Age Range and Quadratic Age With Gray Matter Differences Overlaid Onto the Main Meta-analysis Results (A) and Meta-regression Plots (B)

Age-related meta-regression results. A. Main meta-analysis results (CP < TD) shown in dark blue, association between age-range across groups and gray matter differences shown in light blue, association between quadratic age (mean age in the CP youths squared) and gray matter differences shown in red/pink (left fusiform gyrus effect shown circled). All results thresholded at p < 0.05. Slices are shown in the axial plane with MNI-coordinates included. B. Association between the age-range and gray matter differences in the left amygdala (k = 140 voxels; Intercept = -0.4) (shown in light blue). Association between quadratic age and gray matter differences in right amygdala (k = 59 voxels; Intercept = -0.41) and left fusiform gyrus (k = 12 voxels; Intercept = -0.4). Effect sizes (SDM-estimates) used to create the meta-regression plots were extracted from the peak of maximum slope significance. Note that this may overestimate the effect that would be observed in a wider anatomical region. The meta-regression SDM-estimate value is derived from the proportion of studies that reported gray matter changes near the voxel so it is expected that some values are at 0 or near +/- 1. Each included study is represented as a numbered dot, with the dot size reflecting total sample size. Large dots indicate total samples > 80 participants; medium dots > 40 participants; and small dots > 20 participants. Study key: 1 = Sterzer et al.,2007; 2 = Huebner et al.,2008; 3 = De Brito et al.,2009; 4 = Dalwani et al.,2011; 5 = Fairchild et al.,2011; 6 = Fahim et al.,2011; 7 = Stevens et al.,2012; 8 = Fairchild et al.,2013; 9 = Olvera et al.,2014; 10 = Cope et al.,2014; 11 = Hummer et al.,2014; 12 = Michalska et al.,2015; 13 = Sebastian et al., 2015.
eFigure 2. Meta-regression Results Showing Association Between IQ Difference and Gray Matter Differences Overlaid Onto the Main Meta-analysis Results (A) and Meta-regression Plot (B)

IQ-difference meta-regression results. A. Main meta-analysis results (CP < TD) are shown in dark blue, association between IQ-difference scores (TD youths (N = 257) minus CP youths (N = 329)) and gray matter differences in right fusiform gyrus is shown in red (circled). All results thresholded at p < 0.05. Slices are shown in the sagittal and axial planes with MNI-coordinates included. B. Association between IQ-difference score and gray matter differences in right fusiform gyrus (k = 88 voxels: Intercept = -1.29). Effect sizes (SDM-estimates) used to create the meta-regression plot were extracted from the peak of maximum slope significance. Note that this may overestimate the effect that would be observed in a wider anatomical region. The meta-regression SDM-estimate value is derived from the proportion of studies that reported gray matter changes near the voxel so it is expected that some values are at 0 or near +/- 1. Each included study is represented as a numbered dot, with the dot size reflecting total sample size. Large dots indicate total samples > 80 participants; medium dots > 40 participants; and small dots > 20 participants. Study key: 1 = Sterzer et al.,2007; 2 = Huebner et al.,2008; 3 = De Brito et al.,2009; 4 = Dalwani et al.,2011; 5 = Fairchild et al.,2011; 6 = Fahim et al.,2011; 7 = Stevens et al.,2012; 8 = Fairchild et al.,2013; 9 = Olvera et al.,2014; 10 = Cope et al.,2014; 11 = Hummer et al.,2014; 12 = Michalska et al.,2015; 13 = Sebastian et al., 2015.
eFigure 3. Meta-regression Results Showing Association Between Lifetime CD Symptom Severity Score and Gray Matter Differences Overlaid Onto the Main Meta-analysis Results (A) and Meta-regression Plot (B)

Lifetime CD-symptom severity meta-regression results. A. Main meta-analysis results (CP < TD) are shown in dark blue, association between mean lifetime CD-symptom severity score in youths with CP (N = 211) and gray matter differences in right superior temporal gyrus shown in green (circled). All results thresholded at p < 0.05. Slices are shown in the sagittal and axial planes with MNI-coordinates included. B. Association between mean lifetime CD-symptom severity score and gray matter differences in right superior temporal gyrus (k = 51 voxels: Intercept = -0.85). Effect sizes (SDM-estimates) used to create the meta-regression plot were extracted from the peak of maximum slope significance. Note that this may overestimate the effect that would be observed in a wider anatomical region. The meta-regression SDM-estimate value is derived from the proportion of studies that reported gray matter changes near the voxel so it is expected that some values are at 0 or near +/- 1. Each included study is represented as a numbered dot, with the dot size reflecting total sample size. Large dots indicate total samples > 80 participants; medium dots > 40 participants; and small dots > 20 participants. Study key: 1 = Sterzer et al.,2007; 2 = Huebner et al.,2008; 3 = De Brito et al.,2009; 4 = Dalwani et al.,2011; 5 = Fairchild et al.,2011; 6 = Fahim et al.,2011; 7 = Stevens et al.,2012; 8 = Fairchild et al.,2013; 9 = Olvera et al.,2014; 10 = Cope et al.,2014; 11 = Hummer et al.,2014; 12 = Michalska et al.,2015; 13 = Sebastian et al., 2015.
**eTable 2.** Age Range (Years) Across Studies and Mean Quadratic Age (Age Squared) in Youths With CP

<table>
<thead>
<tr>
<th>Studies</th>
<th>Age range (years)</th>
<th>Mean Age CP Youths (quadratic term)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sterzer et al., 17 2007</td>
<td>6</td>
<td>163.8</td>
</tr>
<tr>
<td>Huebner et al., 18 2008</td>
<td>5</td>
<td>210.3</td>
</tr>
<tr>
<td>De Brito et al., 6 2009</td>
<td>3.3</td>
<td>134.6</td>
</tr>
<tr>
<td>Daiwani et al., 19 2011</td>
<td>4</td>
<td>275.6</td>
</tr>
<tr>
<td>Fahim et al., 20 2011</td>
<td>0</td>
<td>70.6</td>
</tr>
<tr>
<td>Fairchild et al., 7 2011</td>
<td>5</td>
<td>316.8</td>
</tr>
<tr>
<td>Stevens &amp; Haney-Caron, 21 2012</td>
<td>1</td>
<td>256.0</td>
</tr>
<tr>
<td>Fairchild et al., 8 2013</td>
<td>6</td>
<td>295.8</td>
</tr>
<tr>
<td>Olvera et al., 22 2014</td>
<td>4</td>
<td>249.6</td>
</tr>
<tr>
<td>Cope et al., 9 2014</td>
<td>4.1</td>
<td>302.8</td>
</tr>
<tr>
<td>Hummer et al., 23 2014</td>
<td>4</td>
<td>234.1</td>
</tr>
<tr>
<td>Michalska et al., 24 2015</td>
<td>2</td>
<td>102.0</td>
</tr>
<tr>
<td>Sebastian et al., 10 2015</td>
<td>6</td>
<td>204.5</td>
</tr>
</tbody>
</table>

Abbreviations: CP = conduct problems.
**eTable 3.** Group Comparison on IQ Measures, IQ Group Difference Score, and Inclusion of IQ as a Covariate Across Studies

<table>
<thead>
<tr>
<th>Studies</th>
<th>Group comparison on IQ p value</th>
<th>IQ group difference</th>
<th>IQ included in the main analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sterzer et al., 17 2007</td>
<td>.18</td>
<td>6.6</td>
<td>Yes</td>
</tr>
<tr>
<td>Huebner et al., 18 2008</td>
<td>.36</td>
<td>2.2</td>
<td>No</td>
</tr>
<tr>
<td>De Brito et al., 6 2009</td>
<td>&lt;.001</td>
<td>11.5</td>
<td>Yes</td>
</tr>
<tr>
<td>Dalwani et al., 19 2011</td>
<td>.001</td>
<td>7.1</td>
<td>Yes</td>
</tr>
<tr>
<td>Fahim et al., 20 2011</td>
<td>.24 (Block design); .85 (Vocabulary)</td>
<td>NA</td>
<td>No</td>
</tr>
<tr>
<td>Fairchild et al., 7 2011</td>
<td>.20</td>
<td>2.1</td>
<td>No</td>
</tr>
<tr>
<td>Stevens &amp; Haney-Caron, 22 2012</td>
<td>.14</td>
<td>6.1</td>
<td>No</td>
</tr>
<tr>
<td>Fairchild et al., 8 2013</td>
<td>.03</td>
<td>6.0</td>
<td>No</td>
</tr>
<tr>
<td>Olvera et al., 22 2014</td>
<td>.09</td>
<td>6.7</td>
<td>No</td>
</tr>
<tr>
<td>Cope et al., 9 2014</td>
<td>&lt;.001</td>
<td>17.6</td>
<td>No</td>
</tr>
<tr>
<td>Hummer et al., 23 2014</td>
<td>.17</td>
<td>4.2</td>
<td>No</td>
</tr>
<tr>
<td>Michalska et al., 24 2015</td>
<td>IQ data not collected</td>
<td>NA</td>
<td>No</td>
</tr>
<tr>
<td>Sebastian et al., 10 2015</td>
<td>.06</td>
<td>7.3</td>
<td>No</td>
</tr>
</tbody>
</table>

Notes: IQ = intelligence quotient; NA = not available.
eTable 4. The Diagnostic Instruments Used to Assess CD Symptoms Across Studies and Mean Lifetime CD Symptom Severity Score

<table>
<thead>
<tr>
<th>Studies</th>
<th>CD symptoms assessed</th>
<th>Diagnostic instrument</th>
<th>Lifetime CD symptoms M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sterzer et al., 17 2007</td>
<td>Yes</td>
<td>DCL-SSV</td>
<td>NA</td>
</tr>
<tr>
<td>Huebner et al., 18 2008</td>
<td>Yes</td>
<td>K-SADS-PL</td>
<td>NA</td>
</tr>
<tr>
<td>De Brito et al., 6 2009</td>
<td>No</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Dalwani et al., 19 2011</td>
<td>Yes</td>
<td>DISC</td>
<td>6.44 (2.83)</td>
</tr>
<tr>
<td>Fahim et al., 20 2011</td>
<td>No</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Fairchild et al., 7 2011</td>
<td>Yes</td>
<td>K-SADS-PL</td>
<td>8.30 (2.16)</td>
</tr>
<tr>
<td>Stevens &amp; Haney-Caron, 21 2012</td>
<td>Yes</td>
<td>K-SADS-PL</td>
<td>6.36 (2.38)</td>
</tr>
<tr>
<td>Fairchild et al., 8 2013</td>
<td>Yes</td>
<td>K-SADS-PL</td>
<td>7.59 (2.26)</td>
</tr>
<tr>
<td>Olvera et al., 22 2014</td>
<td>Yes</td>
<td>K-SADS-PL</td>
<td>8.17 (2.70)</td>
</tr>
<tr>
<td>Cope et al., 9 2014</td>
<td>Yes</td>
<td>K-SADS-PL</td>
<td>11.10 (1.71)</td>
</tr>
<tr>
<td>Hummer et al., 23 2014</td>
<td>Yes</td>
<td>K-SADS-PL</td>
<td>4.13 (3.38)</td>
</tr>
<tr>
<td>Michalska et al., 24 2015</td>
<td>Yes</td>
<td>DISC</td>
<td>NA</td>
</tr>
<tr>
<td>Sebastian et al., 10 2015</td>
<td>No</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

Notes: DCL-SSV = Diagnose-Checklisten Störungen des Sozialverhaltens; K-SADS-PL = Kiddie-Sads-Present and Lifetime Version; DISC = Diagnostic Interview Schedule; NA = not available
The large age range of the samples included in some of the studies (e.g., 8,10,17) is an important point to consider, which we6 and others23 have previously identified. Given evidence from other disorders (e.g., 25) that mean age of patients is associated with the magnitude of the grey matter difference observed between the clinical and the control group, we ran a meta-regression examining this aspect in the youths with CP. Despite the cross-sectional nature of the data used here, this type of analysis could provide preliminary information as to whether grey matter reductions observed in youths with CP is associated with age and, if so, in which region(s). Because the association between age and grey matter in the amygdala/insula and the superior frontal gyrus follows quadratic and cubic functions, respectively (see Tables 2 in 26,27), we used quadratic-age (age-squared) and age-cubic for this analysis. This approach is consistent with a number of previous meta-analyses (28,29). The results of that analysis are presented in eFigure 1 (analyses using mean quadratic age (age-squared) and age-cubic produced the same results, so the age-cubic results are not reported here).
**eMethods 4. Meta-regression With IQ Group Difference**

Out of the 13 studies, two studies compared groups that differed significantly in terms of IQ, but those studies did not include IQ as a covariate of no interest in their main analyses (studies 8,9). However, Fairchild et al. (2013)8 showed that including IQ as a covariate did not alter their main findings. A third study did not measure IQ in their sample so we could not examine its influence on their results (study 24). Of the remaining 10 studies, two compared groups that did differ on IQ (studies 6,19), but included IQ as a covariate in their main analysis, while the remaining eight studies compared groups that did not differ on IQ (studies 7,10,17,18,20-23). Thus, it appears that for all studies, but three, it is unlikely that IQ might have substantially influenced our main results.

However, across the 13 studies included in our meta-analysis the mean IQ for TD youths (M = 104; SD = 4.27) and those with CP (M = 97; SD = 3.71) differed significantly (p < 0.001). Therefore, to investigate the extent to which IQ might have influenced our main results, similar to our analysis examining the influence of ADHD and age-related effects described above, we conducted an additional meta-regression to examine the influence of IQ (as a difference score) on the reported grey matter differences. The results of that analysis are presented in eFigure 2.
**eTable 5. Excluded Studies Based on Selection Criteria**

<table>
<thead>
<tr>
<th>No.</th>
<th>Studies</th>
<th>Title</th>
<th>Reason for Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Li et al., 2005</td>
<td>Adolescents with disruptive behavior disorder investigated using an optimized MR diffusion tensor imaging protocol</td>
<td>Used diffusion tensor imaging (DTI)</td>
</tr>
<tr>
<td>2</td>
<td>Finger et al., 2012</td>
<td>Impaired functional but preserved structural connectivity in limbic white matter tracts in youth with conduct disorder or Oppositional Defiant Disorder plus psychopathic traits</td>
<td>Used diffusion tensor imaging (DTI)</td>
</tr>
<tr>
<td>3</td>
<td>Passamonti et al., 2012</td>
<td>Abnormal anatomical connectivity between the amygdala and orbitofrontal cortex in conduct disorder</td>
<td>Used diffusion tensor imaging (DTI)</td>
</tr>
<tr>
<td>4</td>
<td>Sarkar et al., 2013</td>
<td>Frontotemporal white-matter microstructural abnormalities in adolescents with conduct disorder: A diffusion tensor imaging study</td>
<td>Used diffusion tensor imaging (DTI)</td>
</tr>
<tr>
<td>5</td>
<td>Haney-Caron et al., 2014</td>
<td>DTI-measured white matter abnormalities in adolescents with Conduct Disorder</td>
<td>Used diffusion tensor imaging (DTI)</td>
</tr>
<tr>
<td>6</td>
<td>Zhang, Gao, et al., 2014</td>
<td>Sex differences of uncinate fasciculus structural connectivity in individuals with conduct disorder</td>
<td>Used diffusion tensor imaging (DTI)</td>
</tr>
<tr>
<td>7</td>
<td>Zhang, Zhu, et al., 2014</td>
<td>Increased structural connectivity in corpus callosum in adolescent males with conduct disorder</td>
<td>Used diffusion tensor imaging (DTI)</td>
</tr>
<tr>
<td>8</td>
<td>Hyatt et al., 2012</td>
<td>Cortical thickness and folding deficits in conduct-disordered adolescents</td>
<td>Used surface-based rather than voxel-based morphometry</td>
</tr>
<tr>
<td>9</td>
<td>Wallace et al., 2014</td>
<td>Cortical and subcortical abnormalities in youths with conduct disorder and elevated callous-unemotional traits</td>
<td>Used surface-based rather than voxel-based morphometry</td>
</tr>
<tr>
<td>10</td>
<td>Sarkar et al., 2014</td>
<td>Reduced cortical surface area in adolescents with conduct disorder</td>
<td>Used surface-based rather than voxel-based morphometry</td>
</tr>
<tr>
<td>11</td>
<td>Ermer et al., 2012</td>
<td>Aberrant paralimbic gray matter in criminal psychopathy</td>
<td>Did not include a comparison with TD youths/Report ROI-based findings only</td>
</tr>
<tr>
<td>12</td>
<td>Ermer et al., 2013</td>
<td>Aberrant paralimbic gray matter in incarcerated male adolescents with psychopathic traits</td>
<td>Did not include a comparison with TD youths/Report ROI-based findings only</td>
</tr>
<tr>
<td>13</td>
<td>White et al., 2013</td>
<td>The relationship between large cavum septum pellucidum and antisocial behavior, callous-unemotional traits and psychopathy in adolescents</td>
<td>Used manual tracing methods</td>
</tr>
<tr>
<td>14</td>
<td>De Brito et al., 2011</td>
<td>Small, but not perfectly formed: decreased white matter concentration in boys with psychopathic tendencies</td>
<td>Reported structural changes in white matter concentration</td>
</tr>
<tr>
<td>15</td>
<td>Benegal et al., 2007</td>
<td>Gray matter volume abnormalities and externalizing symptoms in subjects at high risk for alcohol dependence</td>
<td>Focused on patients at risk for alcohol dependence</td>
</tr>
</tbody>
</table>
**eTable 6.** Results of the Jack-knife Reliability Analyses of the Main Meta-analysis Findings

<table>
<thead>
<tr>
<th>Clusters&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Left Amygdala (-32, 2, -20)</th>
<th>Left Insula (-42, 8, -8)</th>
<th>Right Insula (36, 20, -16)</th>
<th>Right IFG (48, 20, 2)</th>
<th>Left SFG (-6, 54, 28)</th>
<th>Left fusiform gyrus (-34, -78, 16)</th>
<th>Left postcentral gyrus (-56, -14, 36)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sterzer et al.,&lt;sup&gt;17&lt;/sup&gt; 2007</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Huebner et al.,&lt;sup&gt;18&lt;/sup&gt; 2008</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>De Brito et al.,&lt;sup&gt;9&lt;/sup&gt; 2009</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Dalwani et al.,&lt;sup&gt;19&lt;/sup&gt; 2011</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Fahim et al.,&lt;sup&gt;20&lt;/sup&gt; 2011</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Fairchild et al.,&lt;sup&gt;7&lt;/sup&gt; 2011</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Stevens &amp; Haney-Caron,&lt;sup&gt;21&lt;/sup&gt; 2012</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Fairchild et al.,&lt;sup&gt;8&lt;/sup&gt; 2013</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Olvera et al.,&lt;sup&gt;22&lt;/sup&gt; 2014</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Cope, Ermer, Gaudet et al.,&lt;sup&gt;9&lt;/sup&gt; 2014</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Hummer et al.,&lt;sup&gt;23&lt;/sup&gt; 2014</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Michalska et al.,&lt;sup&gt;24&lt;/sup&gt; 2015</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Sebastian et al.,&lt;sup&gt;10&lt;/sup&gt; 2015</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

Notes: Yes, brain region remains significantly decreased following exclusion as part of the jackknife sensitivity analysis; No, brain region is no longer significantly decreased in those analyses when the dataset is removed. Abbreviations: IFG = inferior frontal gyrus; SFG = superior frontal gyrus.

<sup>a</sup> Results thresholded at $p < 0.005$, requiring a peak $Z > 1$ and a cluster extent of 10 voxels.
### Table 7. Results of the Jack-knife Reliability Analyses of the Subgroup Childhood-Onset CP vs TD Youths

Meta-analysis

<table>
<thead>
<tr>
<th>Clusters</th>
<th>Left Amygdala (-26, 0, -12)</th>
<th>Left Insula (-40, 12, -12)</th>
<th>Right Insula (42, 18, 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Studies</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sterzer et al., 2007</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Huebner et al., 2008</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>De Brito et al., 2009</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Fahim et al., 2011</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Fairchild et al., 2011</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Michalska et al., 2015</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Notes: Yes, brain region remains significantly decreased following exclusion as part of the jackknife sensitivity analysis; No, brain region is no longer significantly decreased in those analyses when the dataset is removed.
eFigure 4. Meta-regression Results Showing an Association Between the Proportions of Male:Female Youths With CP and GMV (A) and an Association Between CU Percent of Maximum Possible Scores and GMV (B)

Meta-regression results. Slices are shown in the sagittal and axial planes with MNI coordinates of the selected slices representing the peak in the x,y,z direction. A. Association between the proportion of male compared to female youths with CP (N = 394) and gray matter differences in the left amygdala (k = 165 voxels: Intercept = -0.31) and the right inferior temporal gyrus (k = 115 voxels: Intercept = 0.75). B. Association between the normalized CU trait (Percent of Maximum Possible) score in youths with CP (N = 188) and gray matter differences in the left putamen (k = 14 voxels: Intercept = -0.91). Effect sizes (SDM-estimates) used to create the meta-regression plots were extracted from the peak of maximum slope significance. Note that this may overestimate the effect that would be observed in a wider anatomical region. The meta-regression SDM-estimate value is derived from the proportion of studies that reported gray matter changes near the voxel so it is expected that some values are at 0 or near +/-1. Each included study is represented as a numbered dot, with the dot size reflecting total sample size. Large dots indicate total samples > 80 participants; medium dots > 40 participants; and small dots > 20 participants. Study key: 1 = Sterzer et al., 2007; 2 = Huebner et al., 2008; 3 = De Brito et al., 2009; 4 = Dalwani et al., 2011; 5 = Fairchild et al., 2011; 6 = Fahim et al., 2011; 7 = Stevens et al., 2012; 8 = Fairchild et al., 2013; 9 = Olvera et al., 2014; 10 = Cope et al., 2014; 11 = Hummer et al., 2014; 12 = Michalska et al., 2015; 13 = Sebastian et al., 2015. Note, only 5 studies included a measure of CU traits in youths with CP.
References


