**Supplemental Material**

**Confirmatory factor analysis and structural equation models general specifications**

Confirmatory factor analysis (CFA) was applied to test the bifactor structure psychopathology using the internalizing-externalizing framework of the Child and Behavioral Checklist (CBCL) (Achenbach & Rescorla, 2001), executive functions second order model and academic achievement latent variables. CFA was performed in Mplus version 8.6 (Muthén & Muthén, 2017) and implemented in R version 4.0.3 using the *MplusAutomation* package (Hallquist & Wiley, 2018), which was also used to extract factor scores generated in Mplus. All bifactor reliability indices were calculated using the *BifactorIndicesCalculator* package in R (Dueber, 2017). We used indicators from the two time points simultaneously and defined the subjects as a cluster in all models, except reading comprehension (as described in (Simioni et al., 2019). We applied delta parameterization and weighted least squares with diagonal weight matrix with standard errors and mean- and variance-adjusted chi-square test statistics (WLSMV) estimator for the CBCL and academic achievement models and robust maximum likelihood (MLR) for the executive functions model. We applied full-information maximum likelihood (FIML) to deal with missing data. Model fit parameters were Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), Tucker Lewis Index (TLI) and Standardized Root Mean Squared Error (SRMR). Values of RMSEA near or below 0.080 represent acceptable model fit, and values lower than 0.060 represent good-to-excellent model fit (Hu & Bentler, 1999; Kline, 2015). CFI and TLI values near or above 0.900 represent acceptable model fit, whereas values higher than 0.950 represent a good-to-excellent model fit. SRMR lower or equal than 0.100 indicate adequate fit, and lower than 0.060 in combination with previous indices indicate good fit (Hu & Bentler, 1999; Kline, 2015). Model-based reliability indices were also calculated for latent factors. We used omega (ω - analogous to the alpha coefficient, but appropriate for tests that have varying factor loadings), hierarchical omega (ωH - the proportion of total variance attributed to the general or specific factors), factor determinacy (FD - the correlation between the factor scores and the estimated factor) and H index (quantifies how well each latent factor is represented by the items loading on it). H index ≥ 7 represents a well-defined latent variable and FD ≥ 0.9 indicates that factor scores are appropriate to be used in analysis (Dueber, 2017; Hancock, G. R & Mueller, R. O., 2001; Rodriguez, Reise, & Haviland, 2016). In all regression analysis, individual's standardized factor scores of the below CFA-generated variables were saved and further regressed on age and gender.

Structural equation models (SEM) were applied for all models testing the associations between cerebellar standard deviation volumes with the selected outcomes (i.e., age- and sex-residualized measures of psychopathology, academic achievement, and executive functions). SEM was estimated using maximum likelihood estimator and FIML to handle missing data. Model fit parameters were the same as for CFA. SEM were carried out using the “sem” function in the *lavaan* package (Rosseel et al., 2018).

**CBCL factor model**

To access general (p-factor) and specific psychopathology (internalizing and externalizing dimensions), we used the 32 internalizing-related items and the 35 externalizing-related items according to the CBCL user guide (Achenbach & Rescorla, 2001), collected at the same day of the brain imaging acquisition, at baseline and follow-up. This model indicated a good global fit index (RMSEA=0.031 [90% CI = 0.029, 0.032]; CFI=0.929, TLI=0.925). Factor loadings and model-based indices are reported in supplementary table S1. Factor score trajectories from each individual are depicted in Figure S2 and described in Table S3.

The model presented factors with good H index and factor determinacy, which indicates the appropriateness of using the factor scores. Furthermore, feeling unloved, talking too much about suicide, mood changes, sulks and suspiciousness items collapsed in the specific factors and were indicators of the p-factor only. Conversely, shyness was exclusive indicator of the internalizing factor and indicators of substance use were heavily loaded by the externalizing factor. This pattern is common in several bifactor models using the CBCL and the p-factors are highly correlated, even when entire domains and sets of items are not used in the model (Hoffmann et al., 2021). Nevertheless, it is possible to observe, by the factor loadings pattern in Table S1, that the p-factor may represents mood dysregulation/dysphoria, internalizing may represent high shyness and externalizing represents specific substance use problems (Heinrich et al., 2021).

**Executive functions factor model**

Executive functions were modelled as previously published (Martel et al., 2017), including the longitudinal extension. As described in the main text, working memory was measured by both verbal (VWM) and visuospatial (VSWM) tasks. Inhibitory control was measured by the conflict control task (CCT) and the Go/No-go (GNG) task. Temporal processing was measured using the time anticipation (TA) task. Executive function was modelled as a second-order factor (common factor among the aforementioned first-order factors). This model indicated a good global fit index (RMSEA=0.007 [90% CI = 0.000, 0.028]; CFI=0.999, TLI=0.999). Factor loadings are reported in supplementary table S2. Factor score trajectories from each individual are depicted in Figure S2 and described in Table S3.

**Academic achievement (school subjects, reading and writing ability)**

Achievement on school subjects was evaluated using CBCL items regarding performance on school subjects at both time points. Parents rated their children as failing, below average, average or above average. This unidimensional model loaded highly on the items/subjects of Portuguese/literature (0.842), mathematics (0.666), history/social studies (0.878), geography (0.898), English/Spanish (0.758), biology (0.917), sciences (0.893) and computer class (0.686). School subject model indicated an adequate global fit index (RMSEA=0.095 [90% CI = 0.089, 0.100]; CFI=0.989, TLI=0.985) and model-based indices (ω=0.900, ωH=0.900, FD=0.978, H=0.957). Factor score trajectories from each individual are depicted in Figure S2 and described in Table S3.

Reading and writing items were fitted in a two-correlated factor model. These abilities were measured by the School Performance Test (“Teste de Desempenho Escolar” - TDE) (Stein, 1998). It is composed of right/wrong items of reading and writing tests. We used 61 read decoding items and 12 writing items (writing words based on oral dictation) to assess its respective ability, at baseline and follow-up. The reading and writing correlated model indicated an excellent global fit index (RMSEA=0.018 [90% CI = 0.018, 0.019]; CFI=0.992, TLI=0.992) and model-based indices for reading (ω=0.973, ωH=0.882, FD=0.987, H=0.973) and writing (ω=0.998, ωH=0.983, FD=1.356, H=0.998). Factor score trajectories from each individual are depicted in Figure S2 and described in Table S3.

**REFERENCES**

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA School-Age Forms & Profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth & Families: Library of Congress.

Dueber, D. (2017). Bifactor Indices Calculator: A Microsoft Excel-Based Tool to Calculate Various Indices Relevant to Bifactor CFA Models. *Educational, School, and Counseling Psychology Research Tools*. https://doi.org/10.13023/edp.tool.01

Hallquist, M. N., & Wiley, J. F. (2018). MplusAutomation: An R Package for Facilitating Large-Scale Latent Variable Analyses in Mplus. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(4), 621–638. https://doi.org/10.1080/10705511.2017.1402334

Hancock, G. R & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. In *Structural equation modeling: Present and future* (pp. 195–216). Lincolnwood, IL: Scientific Software International.

Heinrich, M., Geiser, C., Zagorscak, P., Burns, G. L., Bohn, J., Becker, S. P., … Knaevelsrud, C. (2021). On the Meaning of the “P Factor” in Symmetrical Bifactor Models of Psychopathology: Recommendations for Future Research From the Bifactor-(S-1) Perspective. *Assessment*, 10731911211060298. https://doi.org/10.1177/10731911211060298

Hoffmann, M. S., Moore, T. M., Axelrud, L. K., Tottenham, N., Zuo, X.-N., Rohde, L. A., … Salum, G. A. (2021). *Reliability and Validity of Bifactor Models of Dimensional Psychopathology in Youth from three Continents* (p. 2021.06.27.21259601). https://doi.org/10.1101/2021.06.27.21259601

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling, Fourth Edition*. Guilford Publications.

Martel, M. M., Pan, P. M., Hoffmann, M. S., Gadelha, A., do Rosário, M. C., Mari, J. J., … Salum, G. A. (2017). A general psychopathology factor (P factor) in children: Structural model analysis and external validation through familial risk and child global executive function. *Journal of Abnormal Psychology*, *126*(1), 137–148. https://doi.org/10.1037/abn0000205

Muthén, L. K., & Muthén, B. O. (2017). *Mplus User’s Guide* (Eighth Edition). Los Angeles: Muthén & Muthén. Retrieved from www.StatModel.com

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, *21*(2), 137–150. https://doi.org/10.1037/met0000045

Rosseel, Y., Oberski, D., Byrnes, J., Vanbrabant, L., Savalei, V., Merkle, E., … Jorgensen, T. D. (2018). lavaan: Latent Variable Analysis (Version 0.6-3). Retrieved from https://CRAN.R-project.org/package=lavaan

Simioni, A. R., Pine, D. S., Sato, J. R., Pan, P. M., Fonseca, R. P., Schafer, J., … Salum, G. A. (2019). A Cognitive Development Chart for School-age Children and Adolescents. *MedRxiv*, 19012963. https://doi.org/10.1101/19012963

Stein, L. M. (1998). *TDE Teste de Desempenho Escolar*. São Paulo: Casa do Psicólogo.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Table S1 - Child Behavior Checklist factor model (clustered within each subject) | | | | |
| Item | Content | Factor loadings | | |
| P-factor | Internalizing | Externalizing |
| 14 | Cries | 0.569 | 0.173 |  |
| 29 | Fears | 0.306 | 0.363 |  |
| 30 | Fears school | 0.308 | 0.375 |  |
| 31 | Fears do bad | 0.311 | 0.331 |  |
| 32 | Perfect | 0.060 | 0.209 |  |
| 33 | Unloved | 0.741 | 0.082 |  |
| 35 | Worthless | 0.701 | 0.221 |  |
| 45 | Nervous | 0.719 | 0.133 |  |
| 50 | Fearful | 0.557 | 0.321 |  |
| 52 | Feels too guilty | 0.576 | 0.304 |  |
| 71 | Self-conscious | 0.345 | 0.621 |  |
| 91 | Talks about suicide | 0.750 | -0.049 |  |
| 112 | Worries | 0.414 | 0.352 |  |
| 5 | Little they enjoy | 0.568 | 0.273 |  |
| 42 | Prefers alone | 0.219 | 0.596 |  |
| 65 | Won’t talk | 0.475 | 0.417 |  |
| 69 | Secretive | 0.258 | 0.612 |  |
| 75 | Shy | 0.062 | 0.698 |  |
| 102 | Lacks energy | 0.324 | 0.537 |  |
| 103 | Sad | 0.686 | 0.388 |  |
| 111 | Withdrawn | 0.372 | 0.625 |  |
| 47 | Nightmares | 0.474 | 0.215 |  |
| 49 | Constipate | 0.160 | 0.303 |  |
| 51 | Dizzy | 0.358 | 0.436 |  |
| 54 | Tired | 0.518 | 0.394 |  |
| 56A | Aches | 0.290 | 0.365 |  |
| 56B | Headaches | 0.389 | 0.364 |  |
| 56C | Nausea | 0.349 | 0.434 |  |
| 56D | Eye problems | 0.299 | 0.248 |  |
| 56E | Skin problems | 0.169 | 0.094 |  |
| 56F | Stomach | 0.437 | 0.273 |  |
| 56G | Vomit | 0.363 | 0.352 |  |
| 2 | Drinks | 0.314 |  | 0.583 |
| 26 | No guilt | 0.370 |  | 0.287 |
| 28 | breaks rules at home | 0.582 |  | 0.661 |
| 39 | Bad friends | 0.356 |  | 0.637 |
| 43 | Lies or cheats | 0.533 |  | 0.477 |
| 63 | Prefers older | 0.284 |  | 0.158 |
| 67 | Runs away | 0.500 |  | 0.416 |
| 72 | Sets fires | 0.499 |  | 0.310 |
| 73 | Sex problems | 0.447 |  | 0.089 |
| 81 | Steals from home | 0.444 |  | 0.515 |
| 82 | Steals outside home | 0.442 |  | 0.540 |
| 90 | Swears | 0.619 |  | 0.421 |
| 96 | Thinks about sex | 0.412 |  | 0.261 |
| 99 | Smoke | 0.236 |  | 0.873 |
| 101 | Truants | 0.389 |  | 0.549 |
| 105 | Use drugs | 0.228 |  | 0.900 |
| 106 | Vandalism | 0.426 |  | 0.588 |
| 3 | Argues | 0.541 |  | 0.308 |
| 16 | Mean | 0.551 |  | 0.394 |
| 19 | Demands a lot of attention | 0.639 |  | 0.074 |
| 20 | Destroys own things | 0.521 |  | 0.458 |
| 21 | Destroys other | 0.622 |  | 0.465 |
| 22 | Disobedient at home | 0.557 |  | 0.615 |
| 23 | Disobedient at school | 0.423 |  | 0.675 |
| 37 | Fights | 0.497 |  | 0.505 |
| 57 | Attacks | 0.541 |  | 0.395 |
| 68 | Screams | 0.682 |  | 0.265 |
| 86 | Stubborn | 0.757 |  | 0.251 |
| 87 | Mood changes | 0.858 |  | -0.007 |
| 88 | Sulks | 0.756 |  | 0.054 |
| 89 | Suspicious | 0.662 |  | -0.145 |
| 94 | Teases | 0.488 |  | 0.459 |
| 95 | Temper | 0.784 |  | 0.270 |
| 97 | Threatens | 0.603 |  | 0.452 |
| 104 | Loud | 0.510 |  | 0.329 |
| *Reliability* | |  |  |  |
|  | ω | 0.944 | 0.892 | 0.936 |
|  | ωH | 0.739 | 0.367 | 0.365 |
|  | FD | 0.966 | 0.868 | 0.941 |
|  | H | 0.975 | 0.929 | 0.968 |
| Note: ω, omega is a model-based reliability index analogous to the alpha coefficient, but appropriate for tests that have varying factor loadings; ωH, hierarchical omega is the proportion of total variance attributed to the general or specific factors; FD, factor determinacy is the correlation between the factor scores and the estimated factor; H index quantifies how well each latent factor is represented by the items loading on it. H index ≥ 7 represents a well-defined latent variable and FD ≥ 0.9 indicates that factor scores could be used (Dueber, 2017; Hancock, G. R & Mueller, R. O., 2001; Rodriguez, Reise, & Haviland, 2016). | | | | |

|  |  |  |  |
| --- | --- | --- | --- |
| Table S2 - Executive functions factor model (clustered within each subject) | | | |
| Indicator | Factor loadings (standardized) | | |
| Working memory | Inhibitory control | Temporal processing |
| VWM | 0.668 |  |  |
| VSWM | 0.719 |  |  |
| CCT |  | 0.755 |  |
| GNG |  | -0.450 |  |
| TA |  |  | 0.945 |
| Executive function | 0.681 | 0.563 | 0.540 |
| Note: Standardized factor loadings are interpreted as change in the indicator in the indicator standard deviation unit for a standard deviation unit change in the factor variable ("std.all" standardization). As described in the main text, working memory was measured by both verbal (VWM) and visuospatial (VSWM) tasks. Inhibitory control was measured by the conflict control task (CCT) and the Go/No-go (GNG) task. Temporal processing was measured using the time anticipation (TA) task. | | | |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table S3 - Descriptive data from typically development and test samples at baseline and follow-up | | | | | | | | | |
|  | Typically development | |  | Test | |  | Overall | |  |
|  | Baseline | Follow-up | Statistic | Baseline | Follow-up | Statistic | Baseline | Follow-up | Statistic |
| (N=216) | (N=103) | (N=461) | (N=344) | (N=677) | (N=447) |
| Age (years) |  |  |  |  |  |  |  |  |  |
| Mean (SD) | 10.6 (1.84) | 14.1 (1.77) | t(207) = -16.50, p = 1.1E-39 | 10.7 (1.96) | 14.4 (1.95) | t(742) = -26.60, p = 7.2E-110 | 10.6 (1.92) | 14.3 (1.91) | t(959) = -31.50, p = 5.11E-150 |
| Median [Min, Max] | 10.0 [7.00, 15.0] | 14.0 [11.0, 19.0] | 11.0 [7.00, 15.0] | 14.0 [11.0, 19.0] | 10.0 [7.00, 15.0] | 14.0 [11.0, 19.0] |
| Sex |  |  |  |  |  |  |  |  |  |
| Boys | 123 (56.9%) | 63 (61.2%) | χ2 (1,319) = 0.511, p = 0.474 | 258 (56.0%) | 195 (56.7%) | χ2 (1,805) = 0.04, p = 0.838 | 381 (56.3%) | 258 (57.7%) | χ2 (1,1124) = 0.228, p = 0.633 |
| Girls | 93 (43.1%) | 40 (38.8%) | 203 (44.0%) | 149 (43.3%) | 296 (43.7%) | 189 (42.3%) |
| Socioeconomic status |  |  |  |  |  |  |  |  |  |
| D/E | 0 (0%) | 0 (0%) | χ2 (1,319) = 4.81, p = 0.028 | 76 (16.5%) | 44 (12.8%) | χ2 (2,805) = 4.00, p = 0.135 | 76 (11.2%) | 44 (9.8%) | χ2 (2,1124) = 0.545, p = 0.761 |
| C | 180 (83.3%) | 75 (72.8%) | 329 (71.4%) | 267 (77.6%) | 509 (75.2%) | 342 (76.5%) |
| A/B | 36 (16.7%) | 28 (27.2%) | 56 (12.1%) | 33 (9.6%) | 92 (13.6%) | 61 (13.6%) |
| Trauma exposure |  |  |  |  |  |  |  |  |  |
| Low level | 213 (98.6%) | 98 (95.1%) |  | 307 (66.6%) | 188 (54.7%) | χ2 (1,760) = 3.08, p = 0.079 | 520 (78.0%) | 286 (70.8%) | χ2 (1,1071) = 6.944, p = 0.008 |
| High level | 0 (0%) | 0 (0%) |  | 147 (31.9%) | 118 (34.3%) | 147 (22.0%) | 118 (29.2%) |
| Any psychiatric disorder |  |  |  |  |  |  |  |  |  |
| Self (yes) | 0 (0%) | 0 (0%) |  | 208 (45.1%) | 167 (48.5%) | χ2 (1,805) = 0.93, p = 0.335 | 208 (30.7%) | 167 (37.4%) | χ2 (1,1124) = 5.333, p = 0.021 |
| Parental (yes) | 0 (0%) |  |  | 221 (47.9%) |  |  | 221 (32.6%) |  |  |
| School achievement (factor score) | |  |  |  |  |  |  |  |  |
| Mean (SD) | 0.139 (0.805) | 0.258 (0.810) | t(201) = -1.22, p = 0.222 | -0.189 (1.07) | -0.118 (0.996) | t(753) = -0.96, p = 0.336 | -0.084 (1.01) | -0.030 (0.969) | t(969) = -0.89, p = 0.375 |
| Median [Min, Max] | -0.0260 [-2.39, 1.94] | 0.0410 [-2.02, 1.94] | -0.0340 [-2.79, 1.94] | -0.094 [-2.79, 1.94] | -0.026 [-2.79, 1.94] | -0.034 [-2.79, 1.94] |
| Reading ability (factor score) |  |  |  |  |  |  |  |  |  |
| Mean (SD) | -0.424 (0.857) | 0.147 (0.563) | t(278) = -7.03, p = 1.6E-11 | -0.635 (0.915) | -0.096 (0.648) | t(776) = -9.61, p = 9.7E-21 | -0.567 (0.901) | -0.040 (0.637) | t(1082) = -11.32, p = 3.62E-28 |
| Median [Min, Max] | -0.414 [-2.78, 0.885] | 0.145 [-0.985, 0.885] | -0.581 [-2.80, 0.885] | -0.096 [-2.00, 0.885] | -0.544 [-2.80, 0.885] | -0.050 [-2.00, 0.885] |
| Writing ability (factor score) |  |  |  |  |  |  |  |  |  |
| Mean (SD) | -0.392 (0.846) | 0.183 (0.577) | t(271) = -7.05, p = 1.5-E11 | -0.576 (0.885) | -0.070 (0.640) | t(777) = -9.26, p = 2.0E-19 | -0.517 (0.876) | -0.011 (0.634) | t(1076) = -11.05, p = 5.6E-27 |
| Median [Min, Max] | -0.302 [-2.63, 0.923] | 0.188 [-1.03, 0.923] | -0.517 [-2.66, 0.923] | -0.090 [-2.11, 0.923] | -0.445 [-2.66, 0.923] | -0.005 [-2.11, 0.923] |
| Reading comprehension (factor score) | |  |  |  |  |  |  |  |  |
| Mean (SD) |  | 0.138 (0.756) |  |  | -0.041 (0.802) |  |  | 0.001 (0.794) |  |
| Median [Min, Max] |  | 0.124 [-1.29, 2.35] |  |  | 0.000 [-1.29, 2.35] |  |  | 0.000 [-1.29, 2.35] |  |
| P-factor (factor score) |  |  |  |  |  |  |  |  |  |
| Mean (SD) | -0.480 (0.665) | -0.507 (0.769) | t(175) = 0.31, p = 0.759 | 0.224 (0.902) | 0.234 (0.890) | t(738) = -0.14, p = 0.887 | -0.001 (0.896) | 0.063 (0.918) | t(929) = -1.15, p = 0.250 |
| Median [Min, Max] | -0.470 [-2.12, 0.988] | -0.383 [-2.12, 0.859] | 0.134 [-1.99, 2.60] | 0.229 [-1.78, 2.31] | -0.058 [-2.12, 2.60] | 0.073 [-2.12, 2.31] |
| Internalizing-specific (factor score) | |  |  |  |  |  |  |  |  |
| Mean (SD) | -0.262 (0.692) | 0.162 (0.708) | t(194) = -5.01, p = 1.2E-6 | -0.031 (0.784) | 0.241 (0.823) | t(713) = -4.72, p = 2.9E-6 | -0.105 (0.763) | 0.223 (0.798) | t(916) = -6.83, p = 1.54E-11 |
| Median [Min, Max] | -0.301 [-1.74, 1.68] | 0.184 [-1.32, 1.90] | -0.059 [-2.01, 2.20] | 0.189 [-1.67, 2.65] | -0.148 [-2.01, 2.20] | 0.189 [-1.67, 2.65] |
| Externalizing-specific (factor score) | |  |  |  |  |  |  |  |  |
| Mean (SD) | -0.121 (0.623) | -0.076 (0.731) | t(173) = -0.54, p = 0.587 | 0.026 (0.784) | 0.220 (0.886) | t(680) = -3.22, p = 0.001 | -0.021 (0.739) | 0.152 (0.861) | t(843) = -3.48, p = 0.001 |
| Median [Min, Max] | -0.219 [-1.43, 1.87] | -0.208 [-1.45, 2.22] | -0.018 [-2.47, 1.91] | 0.219 [-2.14, 3.46] | -0.093 [-2.47, 1.91] | 0.137 [-2.14, 3.46] |
| Executive function (factor score) | |  |  |  |  |  |  |  |  |
| Mean (SD) | 0.093 (0.633) | 0.157 (0.602) | t(210) = -0.87, p = 0.383 | 0.015 (0.694) | -0.076 (0.793) | t(674) = 1.69, p = 0.092 | 0.040 (0.675) | -0.022 (0.759) | t(869) = 1.39, p = 0.166 |
| Median [Min, Max] | 0.120 [-1.69, 1.83] | 0.244 [-1.42, 1.57] | 0.050 [-2.34, 1.73] | 0.042 [-3.49, 1.61] | 0.092 [-2.34, 1.83] | 0.105 [-3.49, 1.61] |
| Working memory (factor score) | |  |  |  |  |  |  |  |  |
| Mean (SD) | 0.129 (0.996) | 0.259 (1.07) | t(189) = -1.04, p = 0.299 | -0.005 (1.13) | -0.053 (1.14) | t(730) = 0.60, p = 0.547 | 0.038 (1.09) | 0.019 (1.13) | t(925) = 0.28, p = 0.781 |
| Median [Min, Max] | 0.123 [-2.47, 3.24] | 0.281 [-2.84, 3.60] | -0.015 [-3.96, 3.83] | 0.128 [-3.88, 2.89] | 0.043 [-3.96, 3.83] | 0.150 [-3.88, 3.60] |
| Inhibitory control (factor score) | |  |  |  |  |  |  |  |  |
| Mean (SD) | 0.102 (0.956) | 0.139 (0.885) | t(215) = -0.33, p = 0.740 | -0.033 (0.950) | -0.088 (1.00) | t(711) = 0.78, p = 0.435 | 0.010 (0.953) | -0.035 (0.978) | t(930) = 0.77, p = 0.443 |
| Median [Min, Max] | 0.179 [-2.26, 2.31] | 0.382 [-1.97, 1.61] | 0.012 [-3.65, 2.04] | 0.032 [-3.87, 1.81] | 0.079 [-3.65, 2.31] | 0.151 [-3.87, 1.81] |
| Temporal processing (factor score) | |  |  |  |  |  |  |  |  |
| Mean (SD) | 0.099 (1.08) | 0.147 (0.919) | t(233) = -0.40, p = 0.690 | 0.104 (1.08) | -0.154 (1.33) | t(640) = 2.95, p = 0.003 | 0.103 (1.08) | -0.085 (1.25) | t(848) = 2.59, p = 0.010 |
| Median [Min, Max] | 0.253 [-3.48, 2.24] | 0.255 [-3.42, 1.85] | 0.227 [-3.83, 2.18] | 0.109 [-5.17, 2.20] | 0.236 [-3.83, 2.24] | 0.126 [-5.17, 2.20] |
| Note: The test sample was composed by subjects not included in the typically developing sample. | | | | | | | | | |

|  |  |  |  |
| --- | --- | --- | --- |
| Table S4 - Regression model predicting response to follow-up (used to calculate inverse probability weight) | | | |
| Predictors at baseline | Odds Ratios | CI | p |
| Age (years) | 0.85 | [0.77, 0.93] | **0.001** |
| Trauma exposure | 0.99 | [0.68, 1.47] | 0.979 |
| Any current psychiatric condition (subject) | 0.75 | [0.52, 1.09] | 0.131 |
| Any current psychiatric condition (parental) | 1.28 | [0.89, 1.83] | 0.184 |
| Writing ability (factor score) | 1.14 | [0.91, 1.42] | 0.250 |
| P-factor (factor score) | 0.91 | [0.75, 1.11] | 0.345 |
| Internalizing-specific (factor score) | 1.10 | [0.88, 1.37] | 0.393 |
| Externalizing-specific (factor score) | 1.04 | [0.83, 1.30] | 0.732 |
| Temporal processing | 1.00 | [0.86, 1.16] | 0.975 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Table S5 - Model fit statistics of cerebellum development over three years | | | | | | | |
| Model | df | AIC | BIC | Log-likelihood | Test | Likelihood ratio | p-value |
| *Left hemisphere* | |  |  |  |  |  |  |
| Linear | 6 | 6004.684 | 6027.199 | -2996.342 |  |  |  |
| Quadratic | 8 | 5966.716 | 5996.686 | -2975.358 | Linear vs. Quadratic | 41.967 | <.0001 |
| Cubic | 10 | 5945.540 | 5982.938 | -2962.770 | Quadratic vs. Cubic | 25.176 | <.0001 |
| *Right hemisphere* | |  |  |  |  |  |  |
| Linear | 6 | 6005.017 | 6027.533 | -2996.509 |  |  |  |
| Quadratic | 8 | 5968.608 | 5998.577 | -2976.304 | Linear vs. Quadratic | 40.410 | <.0001 |
| Cubic | 10 | 5944.287 | 5981.685 | -2962.144 | Quadratic vs. Cubic | 28.320 | <.0001 |
| Note: Models are generalized least square models with main effects for age, sex and age by sex interactions, as well as interactions for polynomial terms (quadratic and cubic), using random age slopes for each participant. df, degrees of freedom; AIC, Alkaike information criteria ; BIC, Bayesian information criteria; | | | | | | | |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Table S6 - Results of the cubic model for cerebellum development over three years | | | | |
| Model | Estimate (mm3) | SE | t-value | p-value |
| *Left hemisphere* | |  |  |  |
| Age | 2111.290 | 183.681 | 11.494 | 0.000 |
| Age2 | -125.020 | 110.132 | -1.135 | 0.257 |
| Age3 | -100.220 | 67.796 | -1.478 | 0.140 |
| Sex (female) | -3116.240 | 590.976 | -5.273 | 0.000 |
| Age\*Sex | -671.360 | 303.498 | -2.212 | 0.028 |
| Age2\*Sex | -250.460 | 173.797 | -1.441 | 0.151 |
| Age3\*Sex | -35.350 | 132.003 | -0.268 | 0.789 |
| *Right hemisphere* | |  |  |  |
| Age | 2321.360 | 174.120 | 13.332 | 0.000 |
| Age2 | -39.210 | 104.484 | -0.375 | 0.708 |
| Age3 | -148.960 | 64.387 | -2.313 | 0.021 |
| Sex (female) | -3069.810 | 599.909 | -5.117 | 0.000 |
| Age\*Sex | -880.460 | 287.481 | -3.063 | 0.002 |
| Age2\*Sex | -315.790 | 164.850 | -1.916 | 0.056 |
| Age3\*Sex | 10.890 | 125.158 | 0.087 | 0.931 |
| Note: Cubic model fit are described in Table S4. SE, stanard error. | | | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Table S7** - Mixed, ICV-adjusted, static and dynamic regression model fits of cerebellar volume deviation predicting educational, psychopathology and executive function outcomes | | | | | | |
| Model | Hemisphere | Fit indices | | | | |
| RMSEA | RMSEA 90% CI | CFI | TLI | SRMR |
| Mixed | Left | 0.140 | [0.107, 0.175] | 0.996 | 0.903 | 0.026 |
| Right | 0.139 | [0.107, 0.174] | 0.996 | 0.904 | 0.026 |
| Mixed (ICV-adjusted | Left | 0.141 | [0.109, 0.177] | 0.996 | 0.901 | 0.026 |
| Right | 0.141 | [0.108, 0.176] | 0.996 | 0.902 | 0.026 |
| Static model (ICV-adjusted) | Left | 0.019 | [0.000, 0.083] | 1.000 | 0.994 | 0.013 |
| Right | 0.020 | [0.000, 0.084] | 1.000 | 0.993 | 0.013 |
| Dynamic model (ICV-adjusted) | Left | 0.086 | [0.024, 0.154] | 0.998 | 0.956 | 0.020 |
| Right | 0.086 | [0.022, 0.154] | 0.998 | 0.957 | 0.020 |
| Mediation model (ICV-adjusted) | Left | 0.062 | [0.011, 0.114] | 0.992 | 0.903 | 0.024 |
| Right | 0.061 | [0.011, 0.114] | 0.992 | 0.905 | 0.024 |
| Note: All estimates were from structural equation models including all outcomes predicted by each hemisphere. All predictors and outcomes were standardized by age and sex; ICV, total intracranial volume. Model fit parameters were Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), Tucker Lewis Index (TLI) and Standardized Root Mean Squared Error (SRMR). | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table S9 - Results of inhibitory control mediation structural models of the association between baseline cerebellar volume deviations and future educational and psychopatholgy outcomes | | | | | | | | | | |
| Path | Predictor (baseline) |  | Outcome (follow-up) | β | 95% CI | | p-value | | SE | Z Statistic |
|  | LB | UB |
| *Left cerebellar hemisphere* | | | | | | | | | | |
| c | Cerebellar volume deviation | → | P factor | 0.01 | -0.12 | 0.13 | 0.917 |  | 0.06 | 0.10 |
| a | Cerebellar volume deviation | → | Inhibitory control | 0.17 | 0.05 | 0.31 | 0.006 | \*\* | 0.07 | 2.76 |
| b | Inhibitory control (follow-up) | → | P factor | -0.13 | -0.21 | -0.02 | 0.016 | \* | 0.05 | -2.40 |
| a\*b | Indirect effect |  |  | -0.02 | -0.04 | 0.00 | 0.074 |  | 0.01 | -1.79 |
|  |  |  |  |  |  |  |  |  |  |  |
| c | Cerebellar volume deviation | → | Internalizing specific factor | 0.09 | -0.04 | 0.19 | 0.184 |  | 0.06 | 1.33 |
| a | Cerebellar volume deviation | → | Inhibitory control | 0.17 | 0.05 | 0.31 | 0.006 | \*\* | 0.07 | 2.76 |
| b | Inhibitory control (follow-up) | → | Internalizing specific factor | -0.04 | -0.13 | 0.06 | 0.478 |  | 0.05 | -0.71 |
| a\*b | Indirect effect |  |  | -0.01 | -0.03 | 0.01 | 0.501 |  | 0.01 | -0.67 |
|  |  |  |  |  |  |  |  |  |  |  |
| c | Cerebellar volume deviation | → | Externalizing specific factor | 0.00 | -0.11 | 0.11 | 0.995 |  | 0.06 | 0.01 |
| a | Cerebellar volume deviation | → | Inhibitory control | 0.17 | 0.05 | 0.31 | 0.006 | \*\* | 0.07 | 2.76 |
| b | Inhibitory control (follow-up) | → | Externalizing specific factor | -0.10 | -0.20 | 0.03 | 0.134 |  | 0.06 | -1.50 |
| a\*b | Indirect effect |  |  | -0.02 | -0.04 | 0.01 | 0.195 |  | 0.01 | -1.30 |
|  |  |  |  |  |  |  |  |  |  |  |
| c | Cerebellar volume deviation | → | Achievement on school subjects | 0.02 | -0.11 | 0.14 | 0.814 |  | 0.06 | 0.24 |
| a | Cerebellar volume deviation | → | Inhibitory control | 0.17 | 0.05 | 0.31 | 0.006 | \*\* | 0.07 | 2.76 |
| b | Inhibitory control (follow-up) | → | Achievement on school subjects | 0.22 | 0.07 | 0.34 | 0.002 | \*\* | 0.07 | 3.05 |
| a\*b | Indirect effect |  |  | 0.04 | 0.00 | 0.08 | 0.061 |  | 0.02 | 1.87 |
|  |  |  |  |  |  |  |  |  |  |  |
| c | Cerebellar volume deviation | → | Reading ability | -0.01 | -0.09 | 0.07 | 0.829 |  | 0.04 | -0.22 |
| a | Cerebellar volume deviation | → | Inhibitory control | 0.17 | 0.05 | 0.31 | 0.006 | \*\* | 0.07 | 2.76 |
| b | Inhibitory control (follow-up) | → | Reading ability | 0.22 | 0.05 | 0.23 | 0.002 | \*\* | 0.04 | 3.15 |
| a\*b | Indirect effect |  |  | 0.04 | 0.00 | 0.05 | 0.044 | \* | 0.01 | 2.02 |
|  |  |  |  |  |  |  |  |  |  |  |
| c | Cerebellar volume deviation | → | Writing ability | -0.04 | -0.11 | 0.05 | 0.490 |  | 0.04 | -0.69 |
| a | Cerebellar volume deviation | → | Inhibitory control | 0.17 | 0.05 | 0.31 | 0.006 | \*\* | 0.07 | 2.76 |
| b | Inhibitory control (follow-up) | → | Writing ability | 0.21 | 0.05 | 0.22 | 0.002 | \*\* | 0.04 | 3.13 |
| a\*b | Indirect effect |  |  | 0.04 | 0.00 | 0.05 | 0.043 | \* | 0.01 | 2.02 |
|  |  |  |  |  |  |  |  |  |  |  |
| c | Cerebellar volume deviation | → | Reading comprehension | -0.03 | -0.12 | 0.06 | 0.535 |  | 0.04 | -0.62 |
| a | Cerebellar volume deviation | → | Inhibitory control | 0.17 | 0.05 | 0.31 | 0.006 | \*\* | 0.07 | 2.76 |
| b | Inhibitory control (follow-up) | → | Reading comprehension | 0.32 | 0.16 | 0.34 | 0.000 | \*\* | 0.05 | 5.54 |
| a\*b | Indirect effect |  |  | 0.05 | 0.01 | 0.08 | 0.013 | \* | 0.02 | 2.47 |
|  |  |  |  |  |  |  |  |  |  |  |
| *Right cerebellar hemisphere* | | | | | | | | | | |
| c | Cerebellar volume deviation | → | P factor | 0.01 | -0.11 | 0.12 | 0.930 |  | 0.06 | 0.09 |
| a | Cerebellar volume deviation | → | Inhibitory control | 0.16 | 0.04 | 0.29 | 0.008 | \*\* | 0.06 | 2.67 |
| b | Inhibitory control (follow-up) | → | P factor | -0.13 | -0.21 | -0.02 | 0.017 | \* | 0.05 | -2.39 |
| a\*b | Indirect effect |  |  | -0.02 | -0.04 | 0.00 | 0.079 |  | 0.01 | -1.76 |
|  |  |  |  |  |  |  |  |  |  |  |
| c | Cerebellar volume deviation | → | Internalizing specific factor | 0.12 | 0.00 | 0.21 | 0.058 |  | 0.05 | 1.90 |
| a | Cerebellar volume deviation | → | Inhibitory control | 0.16 | 0.04 | 0.29 | 0.008 | \*\* | 0.06 | 2.67 |
| b | Inhibitory control (follow-up) | → | Internalizing specific factor | -0.05 | -0.14 | 0.06 | 0.430 |  | 0.05 | -0.79 |
| a\*b | Indirect effect |  |  | -0.01 | -0.02 | 0.01 | 0.461 |  | 0.01 | -0.74 |
|  |  |  |  |  |  |  |  |  |  |  |
| c | Cerebellar volume deviation | → | Externalizing specific factor | -0.01 | -0.12 | 0.10 | 0.840 |  | 0.05 | -0.20 |
| a | Cerebellar volume deviation | → | Inhibitory control | 0.16 | 0.04 | 0.29 | 0.008 | \*\* | 0.06 | 2.67 |
| b | Inhibitory control (follow-up) | → | Externalizing specific factor | -0.10 | -0.20 | 0.03 | 0.139 |  | 0.06 | -1.48 |
| a\*b | Indirect effect |  |  | -0.02 | -0.04 | 0.01 | 0.194 |  | 0.01 | -1.30 |
|  |  |  |  |  |  |  |  |  |  |  |
| c | Cerebellar volume deviation | → | Achievement on school subjects | 0.03 | -0.09 | 0.15 | 0.604 |  | 0.06 | 0.52 |
| a | Cerebellar volume deviation | → | Inhibitory control | 0.16 | 0.04 | 0.29 | 0.008 | \*\* | 0.06 | 2.67 |
| b | Inhibitory control (follow-up) | → | Achievement on school subjects | 0.22 | 0.07 | 0.33 | 0.003 | \*\* | 0.07 | 2.99 |
| a\*b | Indirect effect |  |  | 0.03 | 0.00 | 0.07 | 0.060 |  | 0.02 | 1.88 |
|  |  |  |  |  |  |  |  |  |  |  |
| c | Cerebellar volume deviation | → | Reading ability | -0.02 | -0.09 | 0.06 | 0.750 |  | 0.04 | -0.32 |
| a | Cerebellar volume deviation | → | Inhibitory control | 0.16 | 0.04 | 0.29 | 0.008 | \*\* | 0.06 | 2.67 |
| b | Inhibitory control (follow-up) | → | Reading ability | 0.22 | 0.05 | 0.23 | 0.002 | \*\* | 0.04 | 3.17 |
| a\*b | Indirect effect |  |  | 0.03 | 0.00 | 0.05 | 0.045 | \* | 0.01 | 2.01 |
|  |  |  |  |  |  |  |  |  |  |  |
| c | Cerebellar volume deviation | → | Writing ability | -0.05 | -0.11 | 0.04 | 0.356 |  | 0.04 | -0.92 |
| a | Cerebellar volume deviation | → | Inhibitory control | 0.16 | 0.04 | 0.29 | 0.008 | \*\* | 0.06 | 2.67 |
| b | Inhibitory control (follow-up) | → | Writing ability | 0.22 | 0.05 | 0.22 | 0.002 | \*\* | 0.04 | 3.16 |
| a\*b | Indirect effect |  |  | 0.03 | 0.00 | 0.04 | 0.044 | \* | 0.01 | 2.01 |
|  |  |  |  |  |  |  |  |  |  |  |
| c | Cerebellar volume deviation | → | Reading comprehension | -0.04 | -0.12 | 0.05 | 0.423 |  | 0.04 | -0.80 |
| a | Cerebellar volume deviation | → | Inhibitory control | 0.16 | 0.04 | 0.29 | 0.008 | \*\* | 0.06 | 2.67 |
| b | Inhibitory control (follow-up) | → | Reading comprehension | 0.32 | 0.16 | 0.34 | 0.000 | \*\* | 0.05 | 5.56 |
| a\*b | Indirect effect |  |  | 0.05 | 0.01 | 0.08 | 0.015 | \* | 0.02 | 2.43 |
|  |  |  |  |  |  |  |  |  |  |  |
| Note: All estimates were derived from one structural equation model by hemisphere. All variables in this model were standardized by age and sex. Inhibitory control was assessed at follow-up. Cerebellar volume deviation are deviance of cerebellar hemisphere growth standardized by age and gender and were adjusted for intracranial volume. c, direct path between predictor and outcome; a, path between predictor and mediator; b, path between mediator and outcome; LB, lower bound; UP, upper bound; \*, p < 0.05; \*\*, p < 0.01; \*\*\*, p < 0.001. | | | | | | | | | | |

Figure S1: Mean baseline differences of outcome variables between typical development and test samples.

(image attached on a separate file)

Figure S2: Trajectories of academic achievement outcomes, psychopathology and executive function.

(image attached on a separate file)

Figure S3: Scatterplots representing the association of left cerebellar volume devitation with psychopathology, executive function and school achievement outcomes. Fitted curves are derived from mixed regression models adjusted for the intracranial volume.

(image attached on a separate file)

Figure S4: Scatterplots representing the association of right cerebellar volume devitation with psychopathology, executive function and school achievement outcomes. Fitted curves are derived from mixed regression models adjusted for the intracranial volume.

(image attached on a separate file)