The Journal of Child Psychology and Psychiatry

Journal of Child Psychology and Psychiatry **:* (2024), pp **–** doi:10.1111/jcpp.13944

Mapping potential pathways from polygenic liability through brain structure to psychological problems across the transition to adolescence

Benjamin B. Lahey,[1†] E. Leighton Durham,[2†] Sarah J. Brislin,[3] Peter B. Barr,[4] Danielle M. Dick,[3] Tyler M. Moore,[5] Brandon L. Pierce,[1] Lin Tong,[1] Gabrielle E. Reimann,[2] Hee Jung Jeong,[2] Randolph M. Dupont,[2] and Antonia N. Kaczkurkin[2]

[1]University of Chicago, Chicago, IL, USA; [2]Vanderbilt University, Nashville, TN, USA; [3]Rutgers University, Piscataway, NJ, USA; [4]SUNY Downstate Health Sciences University, Brooklyn, NY, USA; [5]University of Pennsylvania, Philadelphia, PA, USA

Background: We used a polygenic score for externalizing behavior (extPGS) and structural MRI to examine potential pathways from genetic liability to conduct problems via the brain across the adolescent transition. **Methods:** Three annual assessments of child conduct problems, attention-deficit/hyperactivity problems, and internalizing problems were conducted across across 9–13 years of age among 4,475 children of European ancestry in the Adolescent Brain Cognitive Development[SM] Study (ABCD Study®). **Results:** The extPGS predicted conduct problems in each wave (R^2 = 2.0%–2.9%). Bifactor models revealed that the extPRS predicted variance specific to conduct problems (R^2 = 1.7%–2.1%), but also variance that conduct problems shared with other measured problems (R^2 = .8%–1.4%). Longitudinally, extPGS predicted levels of specific conduct problems (R^2 = 2.0%), but not their slope of change across age. The extPGS was associated with total gray matter volume (TGMV; R^2 = .4%) and lower TGMV predicted both specific conduct problems (R^2 = 1.7%–2.1%) and the variance common to all problems in each wave (R^2 = 1.6%–3.1%). A modest proportion of the polygenic liability specific to conduct problems in each wave was statistically mediated by TGMV. **Conclusions:** Across the adolescent transition, the extPGS predicted both variance specific to conduct problems and variance shared by all measured problems. The extPGS also was associated with TGMV, which robustly predicted conduct problems. Statistical mediation analyses suggested the hypothesis that polygenic variation influences individual differences in brain development that are related to the likelihood of conduct problems during the adolescent transition, justifying new research to test this causal hypothesis. **Keywords:** Polygenic score; brain structure; general factor of psychopathology; externalizing.

Introduction

Externalizing psychological problems[i] in youth refer to oppositional behavior, conduct problems, attention-deficit/hyperactivity disorder (ADHD), and substance misuse (Krueger et al., 2021; Lahey et al., 2004), which are associated with adverse consequences for the individual and society (Costello, Egger, & Angold, 2005; Lahey, 2021). We examine potential biological factors in externalizing problems using integrated genetic, neuroimaging, and behavioral methods. Although social factors undoubtedly play a role in externalizing problems, twin studies reveal that externalizing problems are substantially heritable and share their genetic influences with other problems (Allegrini et al., 2020; Beauchaine, Hinshaw, & Pang, 2010; Cosgrove et al., 2011; Du Rietz et al., 2021; Pettersson et al., 2019). To assess heritable influences at the individual level, polygenic scores have been developed that quantify aspects of genetic risk for psychological problems captured by the weighted sum of many single nucleotide polymorphisms (SNPs; Wray et al., 2021). A polygenic score for externalizing problems (extPGS) has been developed from data on approximately 1.5 million individuals (Karlsson Linner et al., 2021) using genomic structural equation modeling to identify genetic variance common to externalizing phenotypes, including ADHD, alcohol misuse, tobacco and cannabis use, and risky sexual behavior. In independent tests, the extPGS explained 10.5% of the variance in a broad externalizing factor, 1.65% of the variance in ADHD, 3.10% of the variance in conduct disorder, and 1.96% of the variance in oppositional defiant disorder (Karlsson Linner et al., 2021).

Previous studies have found inverse associations between concurrently measured externalizing problems and brain structure (Cao et al., 2022; Durham et al., 2021; Fairchild et al., 2019; Hoogman et al., 2017, 2019; Mackey et al., 2019; Mooney et al., 2021; Patel, Parker, Salum, Pausova, & Paus, 2022; Zhang et al., 2022). A meta-analysis by the ENIGMA ADHD Working Group identified widespread cortical and subcortical structural differences in those with ADHD compared with controls (Hoogman et al., 2017, 2019). Critically, gray matter volume and other metrics are substantially heritable (Albaugh et al., 2019; Mooney et al., 2021; Pol,

[†]Joint first authorship.

Conflict of interest statement: No conflicts were declared.

Peper, Boomsma, & Kahn, 2007; Zhao et al., 2019), raising the possibility that genetic variants that influence risk for externalizing problems may do so partly by influencing individual differences in brain structure related to risk for externalizing problems.

Note that previous cross-sectional analyses have attempted to map hypothesized mediated biological risk pathways from genetic polymorphisms to individual differences in brain structure and to psychological problems (Forbes et al., 2021; Hagler et al., 2019). These initial tests of mediation were severely limited by the use of variables measured at the same points in time, however, because concurrently measured variables bias estimates of mediation and do not rule out reverse causation (Maxwell, Cole, & Mitchell, 2011; Shrout, 2011). Therefore, the present analyses use *prospective* data to test statistical mediation from the extPGS through brain structure to psychological problems measured 1 and 2 years later.

Methods

Sample

Data were from Release 4.0 of the longitudinal Adolescent Brain Cognitive Development[SM] Study (ABCD Study®) conducted at 21 sites on 11,876 from 9- to 10-year-old participants at baseline.

Ethical considerations

Data were collected at each site under institutional review board approval from each participating university for sharing de-identified data. Waivers were obtained for analyses of the public use de-identified data from the Institutional Review Boards of the University of Chicago and Vanderbilt University.

Measures

The Child Behavior Checklist (CBCL; Achenbach, 2009) is a parent rating scale describing child psychological problems on a scale of 0 = *not true (as far as you know)*, 1 = *somewhat or sometimes true*, or 2 = *very true or often true*. We used a subset of CBCL items based on factor analyses that eliminated CBCL items that did not address psychological problems, such as biting fingernails, constipation, and wishing to be the opposite sex, as well as items with low endorsements in late childhood (Moore et al., 2020). Confirmatory correlated factors (CF) models defined conduct problems, ADHD, and internalizing factors (Table S1) and confirmatory bifactor models defined a general factor and specific conduct problems, ADHD, and internalizing factors (Table S2). The 66 items used in the present analyses and their factor loadings during each annual assessment are in these tables.

Polygenic score (PGS)

Saliva was collected from children as per the ABCD Study biospecimen protocol. Genomic DNA was genotyped at Rutgers University using the NIDA Affymetrics Smokescreen array (Baurley, Edlund, Pardamean, Conti, & Bergen, 2016). Following quality control (https://nda.nih.gov/study.html?id=634), genotyping data consisted of 10,627 samples and 517,724 genotyped SNPs. Samples were removed for problematic genotyping plates (n = 56); low call rates (n = 126); poorly called or monomorphic SNPs (<90%); and reported versus DNA sex mismatches (n = 10). Distribution of SNP and sample call rates were inspected; 155 samples with call rates <97% were excluded. Furthermore, 2,195 SNPs were excluded due to violation of the Hardy–Weinberg equilibrium with $p < 10^{-10}$. Quality control used PLINK (1.9; Purcell et al., 2007), after which 8,326 individuals and 462,767 SNPs remained for imputation. The Michigan imputation server pipeline v1.2.4 (https://imputationserver.sph.umich.edu/index.html#!) conducted genome-wide imputation using Minimac 4 with the Haplotype Reference Consortium panel r1.1 as the reference population. Subsequently, SNPs with minor allele frequency <1% or low imputation quality (R^2 < .8) were removed. Additionally, 12 samples were identified as outliers in a population structure plot (StataCorp, 2019) using principal component analysis in the R 4.1.0 prcomp package (RCoreTeam, 2021) and were removed from analysis, leaving 8,314 samples and 8,574,976 SNPs for PGS analysis. The extPGS was calculated using effect sizes from summary statistics for the general externalizing factor using PRS-CS (Karlsson Linner et al., 2021). PRS-CS uses a Bayesian regression and continuous shrinkage method to correct for the nonindependence among nearby SNPs. SNPs in the extPGS were limited to those from HapMap3 that overlapped between the original GWAS summary statistics and the LD reference panel (1,000 Genomes EUR reference panel).

Brain image acquisition, processing, and quality assurance

The ABCD Study structural imaging protocol was described previously (Casey et al., 2018). Protocols were harmonized for all scanner platforms. Scanning included 3D T1- and 3D T2-weighted images of brain structure, taking place over one or two scanning sessions. The 21 data collection sites across the United States used several models of 3 tesla (3 T) Siemens, General Electric, or Philips scanners; specific scanner models used for data collection were General Electric Discovery MR750, Siemens Prisma, Siemens Prisma Fit, Philips Achieva dStream, and Philips Ingenia. Imaging parameters were repetition time 2,400 to 2,500 ms; echo time 2 to 2.9 ms; field of view (FOV) 256 × 240 to 256; FOV phase of 93.75% to 100%; matrix 256 × 256; 176 to 225 slices; inversion delay 1,060 ms; flip angle of 8°; voxel resolution of 1 × 1 × 1 mm; total acquisition time from 5 min 38 s to 7 min 12 s.

The ABCD Data Analysis and Informatics Center performed centralized processing and analysis of structural data using a collection of processing steps via the Multi-Modal Processing Stream, a software package developed and maintained at the Center for Multimodal Imaging and Genetics at the University of California, San Diego (UCSD). This pipeline consisted of: (a) preprocessing [correction for gradient nonlinearity distortions, intensity scaling and inhomogeneity correction, registration to an averaged reference brain in standard space, and manual quality control (QC)]; (b) brain segmentation (cortical surface reconstruction and subcortical segmentation performed based on automated, atlas-based, segmentation procedures in FreeSurfer v5.3); (c) derivation of morphometric measures [calculation of average volume in each cortical parcel of the standard FreeSurfer Desikan parcellation scheme (Desikan et al., 2006) and in each subcortical region (Fischl et al., 2002)]; and finally, (d) postprocessing QC (manual review for motion, intensity inhomogeneity, white matter underestimation, pial overestimation, and magnetic susceptibility artifact). Detailed descriptions of these methods have been published (Casey et al., 2018; Hagler et al., 2019).

Statistical analyses

Because the extPGS was derived from data on only persons of European ancestries, we included only genotyped children who

self-identified as non-Hispanic white, were identified as European ancestry based on principal components from the 1,000 Genomes Project, and who passed quality assurance for both genotyping and structural magnetic resonance imaging. Demographic characteristics of the analyzed subsample are in Table S3. The analyses used longitudinal data from three annual assessments of 4,475 children of European ancestry at baseline, 4,343 in the first follow-up, and 4,092 in the second follow-up. The extPGS, brain volumes, and child problem factors were standardized by *z*-transformation on children of European ancestry prior to analyses.

We examined possible predictive relationships from genetic risk to individual differences in brain to psychological problems using two complementary measurement models to define the dimensions of psychological problems (Forbes, Greene, et al., 2021; Moore et al., 2020). One of the most important properties of psychological problems is that they are positively correlated (Forbes et al., 2021; Krueger et al., 2021; Lahey, 2021; Watson et al., 2022). Therefore, CF models usefully define dimensions of psychological problems while preserving such correlations. In contrast, bifactor measurement models are valuable for a different purpose. Rather than defining clinical phenotypes, bifactor models are a statistical tool for testing the hierarchical hypothesis that *some risk factors are specific to each particular dimension of problems*, whereas *other risks are nonspecifically shared by all psychological problems* (Caspi et al., 2014; Lahey, Krueger, Rathouz, Waldman, & Zald, 2017; Neumann et al., 2022; Zald & Lahey, 2017). To accomplish this, bifactor models partition the variance common to all psychological problems from the variance that is unique to particular subsets of problems to define orthogonal general and specific factors. Thus, bifactor measurement models are well suited for the present analyses of genetic and neural correlates of psychological problems. That is, they allow valid tests to determine whether polygenic risk for externalizing problems is prospectively associated with variance that is specific only to the dimension of ADHD and/or to the dimensions of conduct problems, or is associated with the variance shared by all psychological problems, or both (Laceulle, Chung, Vollebergh, & Ormel, 2020; Lahey et al., 2017; Lahey, Van Hulle, Singh, Waldman, & Rathouz, 2011; Waszczuk et al., 2020). Because some researchers prefer only to use CF models (Watts, Poore, & Waldman, 2019), however, results based on both measurement models were used.

Associations between the extPGS and measured GMVs were tested using Statistical Analysis Software (SAS; https://www.sas.com) procedures for complex surveys. Generalized linear regression models took clustering within families and stratification by sites into account and incorporated poststratification weights (Heeringa & Berglund, 2018). Response variables were 68 cortical regions, which were derived from a surface-based atlas procedure previously developed (Desikan et al., 2006), and 19 automatically labeled subcortical regions (Fischl et al., 2002). Associations of the extPGS and brain volume measures with latent dimensions of psychological problems in each assessment wave were estimated in Mplus (Muthén & Muthén, 1998-2017), taking clustering and stratification into account. These adjustments were needed to calculate appropriate standard errors because children within the same families and data collection sites are likely to be more similar to one another than randomly selected children in the population. Poststratification weights (Heeringa & Berglund, 2018) based on the known demographics of the US population to adjust estimates of parameters to be more representative of the population.

Parent-reported sex, age in months, and the first 10 principal components for ancestry were included in all analyses as covariates and MRI model, and manufacturer was a covariate in analyses of brain volumes. Associations of the extPGS with total cortical gray matter volume, total subcortical gray matter volume, and total cortical volume were tested separately. Subsequent tests of associations between the extPGS and the 68 cortical regions and 19 subcortical regions were conducted with and without total intracranial volume (TICV) as a covariate to evaluate the possibility that any associations with volumes were global rather than regional.

As shown in Figure 1, all direct and indirect paths between extPRS and brain volumes measured at baseline, and problem behavior dimensions in each of the two annual follow-up assessments were tested using structural equation models in Mplus (Muthén & Muthén, 1998-2017). Direct and indirect paths from the extPGS to each latent problem dimension in each of the three annual assessments through TGMV were estimated in Mplus using 99% bootstrapped confidence intervals. Paths were estimated in separate analyses for problem factors defined using either CF or bifactor models.

Longitudinal growth models were estimated in Mplus (Muthén & Muthén, 1998-2017) using data from each annual assessment to test the associations of the extPGS with the intercept and slope of the general factor and the specific conduct problems, ADHD, and internalizing factors. Factor scores were extracted from bifactor models of parent problem ratings in Mplus separately for each wave then converted to age-based data. Longitudinal data were available on 2,321 (9 years), 4,068 (10 years), 3,629 (11 years), 1,866 (12 years), and 310 (13 years) participants.

As stated in each table of results, family-wise error rates (Benjamini & Hochberg, 1995; Peña, Habiger, & Wu, 2011) were adjusted for false discovery in six families of analyses of related tests of significance. Interactions between sex and the extPRS were tested in exploratory analyses but were not significant for any outcome in any model.

Results

Model psychometrics

CF and bifactor models each fit the CBCL data well in all waves (Tables S2 and S3). Table S4 shows that the general and specific factors defined in bifactor models in all three waves had acceptable H indices >0.70 (Hancock & Mueller, 2001), and each specific factor was reliable according to omega statistics. Explained common variance and omegaH indicated that the robust general factor explained more than half of the estimated variance in each specific factor. Factor determinacies showed that all factors were well defined (Rodriguez, Reise, & Haviland, 2016). Note that the factor loadings on the ADHD factors and the H index for ADHD in both models in every wave were lower than other factors, suggesting that the measurement of ADHD was less strong than other factors.

Polygenic score prediction of behavior

Wave-by-wave analyses. When correlated dimensions of psychological problems were defined in CF models, the extPGS concurrently and prospectively predicted conduct problems (R^2 = 2.0%–2.9%) and ADHD problems (R^2 = 1.4%–1.7%), but not internalizing problems in each of the three annual assessments (Table 1, upper rows). Because the variance in each factor defined in CF models reflects both the variance specific to that factor and the variance that

it shares with the other CF, estimates of association with extPRS were repeated for problem dimensions defined in bifactor models. These analyses reveal a more complex picture. When problem dimensions were defined using bifactor models, extPGS positively predicted both the general (R^2 = .8%–1.4%) and specific conduct problems factors (R^2 = 1.7%–2.1%), and inversely predicted the specific internalizing factor (R^2 = .2%–.6%), in each wave. The extPGS only inconsistently predicted the specific ADHD factor defined in bifactor modeling across waves (Table 1, lower rows).

Longitudinal analyses. Longitudinal growth models revealed that the extPGS predicted the levels (intercepts) of the general factor and the specific conduct problems, ADHD, and internalizing factors defined in bifactor models across 9 through 13 years of age, but not the rates of change (slopes) of any problem factor with increasing age (Figure 2; Table S5). Coefficients for the significant associations of extPGS with each factor were positive, except for the inverse associations with specific internalizing problems.

Polygenic score prediction of brain structure

Significant associations were found in separate analyses between the extPGS and total cortical GMV [β = −.06 (SE = .02), p < .0001], total subcortical GMV [β = −.04 (SE = .02), p < .0115], and total GMV in the entire brain [β = −.06 (SE = .02), p < .0001]. The extPGS accounted for 0.36% of variance in TGMV. The upper rows of Table 2 present the results of separate complementary analyses at the regional level. The extPGS was significantly inversely associated with GMV in 37 cortical and subcortical regions after FDR correction. When TICV was added as a covariate, however, the extPGS was significantly associated with only one regional GMV after FDR correction (lower rows of Table 2), suggesting that the association of the extPGS with gray matter volume is more global than regional.

Prediction of psychological problems by total gray matter volume

Table 3 shows that TGMV was inversely associated concurrently and prospectively with conduct problems (R^2 = 3.1%–4.7%), ADHD (R^2 = 1.7%–3.0%), and internalizing problems (R^2 = <1.0%–1.0%) defined in CF models in all annual waves, adjusting for multiple testing. Using bifactor models, TGMV was associated inversely with both the general factor (R^2 = 1.6%–3.1%) and the specific conduct problems factor (R^2 = 1.7%–2.1%) in all waves, but not with specific ADHD. In addition, specific internalizing problems were modestly associated in the opposite direction with TGMV in each wave.

Tests of statistical mediation

Because the extPGS was found to be associated with both TGMV and conduct problems, and TGMV was found to be associated with conduct problems, formal tests were conducted of indirect statistical paths from the extPGS through TGMV to each problem dimension in each wave (see Figure 1). In all three annual waves, the direct and indirect paths from extPGS to both conduct problems and ADHD

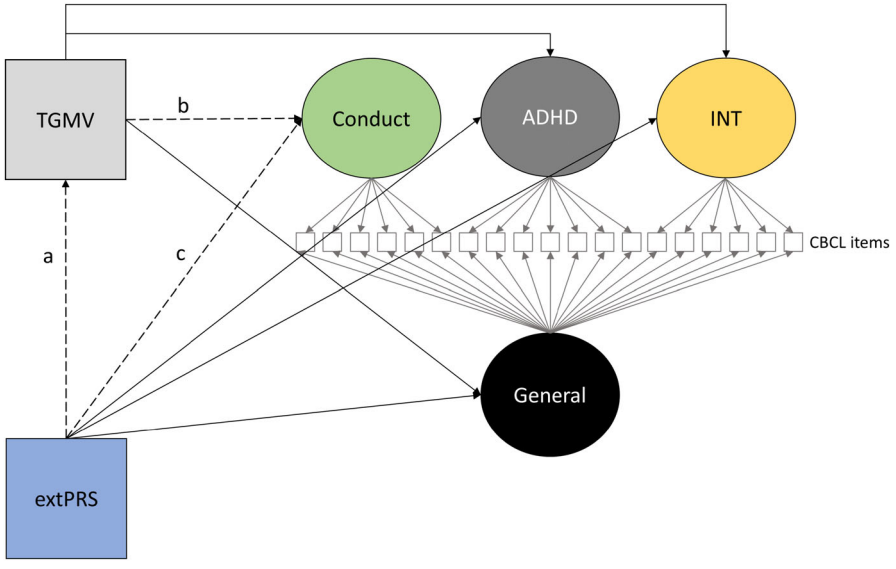

Figure 1 Diagram of all paths tested in the formal analyses of statistical mediation from extPRS and total gray matter volume (TGMV) measured at baseline to problem behavior dimensions based on latent factors from the bifactor models estimated in each of the two annual follow-up assessments. The example of the prospective paths tested for latent specific conduct problems in each wave is illustrated using dashed lines for the direct path from extPGS to specific conduct problems (path c) and for the mediated path between extPGS and specific conduct problems through TGMV (path ab)

Table 1 Results of tests of associations controlling age, sex, and the first 10 principal components for ancestry between a standardized polygenic score for externalizing problems (extPGS) and dimensions of psychological problems defined alternatively by CF and bifactor modeling of parent ratings in the baseline assessment (N = 4,452) and first (N = 4,311) and second (N = 4,054) annual follow-up assessments among European ancestry children

Correlated factors models of psychological problem ratings								
Baseline assessment								
Internalizing			Conduct problems			ADHD		
β (SE)	p	R^2	β (SE)	p	R^2	β (SE)	p	R^2
.040 (.020)	.052	.2	**.142 (.019)**	**<.001**	**2.0**	**.118 (.019)**	**<.001**	**1.4**
Follow-up Wave 1 assessment								
.028 (.020)	.164	.2	**.163 (.20)**	**<.001**	**2.7**	**.132 (.20)**	**<.001**	**1.7**
Follow-up Wave 2 assessment								
.035 (.021)	.097	.1	**.170 (.020)**	**<.001**	**2.9**	**.130 (.020)**	**<.001**	**1.7**
Bifactor models of psychological problem ratings								
Baseline assessment								
General factor			Specific internalizing			Specific conduct problems			Specific ADHD		
β (SE)	p	R^2	β (SE)	p	R^2	β (SE)	p	R^2	β (SE)	p	R^2
.092 (.022)	**<.001**	**.8**	**−.047 (.020)**	**.022**	**.2**	**.145 (.025)**	**<.001**	**2.1**	**.065 (.028)**	**.021**	**.4**
Follow-up Wave 1 assessment											
.113 (.02)	**<.001**	**1.3**	**−.069 (.020)**	**.001**	**.5**	**.132 (.026)**	**<.001**	**1.7**	**.066 (.028)**	**.019**	**.4**
Follow-up Wave 2 assessment											
.119 (.022)	**<.001**	**1.4**	**−.080 (.022)**	**<.001**	**.6**	**.146 (.025)**	**<.001**	**2.1**	.043 (.030)	.155	.2

Factor loadings in the measurement models used in each predictive regression were fixed based on unstandardized factor loadings in a bifactor model without covariates. Tabled p's are unadjusted; coefficients in bold are significant at $p < .05$ after false discovery rate correction for 21 tests of significance.

defined in CF models were 'significant'. That is, the 99% confidence intervals for both the direct and the indirect paths from extPGS to conduct problems and to ADHD did not include zero in any wave (top rows of Table 4). When bifactor modeling was used (lower rows of Table 4), the 99% confidence intervals for both the direct and the indirect paths from extPGS through TGMV to specific conduct problems similarly did not include zero in any wave. The proportion of the association between extPRS and a problem dimension mediated by the indirect path can be calculated as the indirect effect divided by the total effect (sum of the direct and indirect paths). Thus, the proportions of the total observed association between extPGS and specific conduct problems that were mediated by TGMV in the three annual assessments were 7.3%, 9.8%, and 6.6%, respectively.

Sensitivity tests

To check for potential inflation of R^2 values due to overfitting, we performed 10,000 random split-half (50%/50%) cross-validation estimations of the direct effects within the mediation model, in which the model was estimated in one half and used in the second half to predict the bifactor-defined problem scores in the first follow-up. Correlations between true and predicted values of problem dimensions were compared between training and testing sets. As illustrated in the histograms in Figure S1, the generalizability of prediction between training and testing sets was nearly perfect, indicating little or no inflation. Prediction of the specific ADHD factor did produce some anomalous bimodal results toward the lower end of the distribution, but this was limited to the direct effect because the R^2 values for the indirect effect were near-zero and there was not a significant predictive effect to cross-validate.

Because polygenic scores capture only part of the genetic influences on a phenotype, we also evaluated potential bias in the estimated mediated path due to genetic confounding using Gsens (Pingault et al., 2022). Two complementary external estimates of h^2 based on SNP heritability (Choi et al., 2022), and twin studies (Hicks, Krueger, Iacono, McGue, & Patrick, 2004) were used in these tests of bias. As shown in Table S6, there was significant genetic confounding of the mediated relationship between extPGS and specific conduct problems in all three annual assessments based on both estimates of h^2. Nonetheless, all estimates of the mediated associations remained statistically significant (p range .005 to 6.56e−97) after adjustment for genetic confounding.

Discussion

The present findings confirm and extend evidence from earlier twin studies of genetic influences on child externalizing problems (Rhee & Waldman, 2002) using longitudinal analyses and

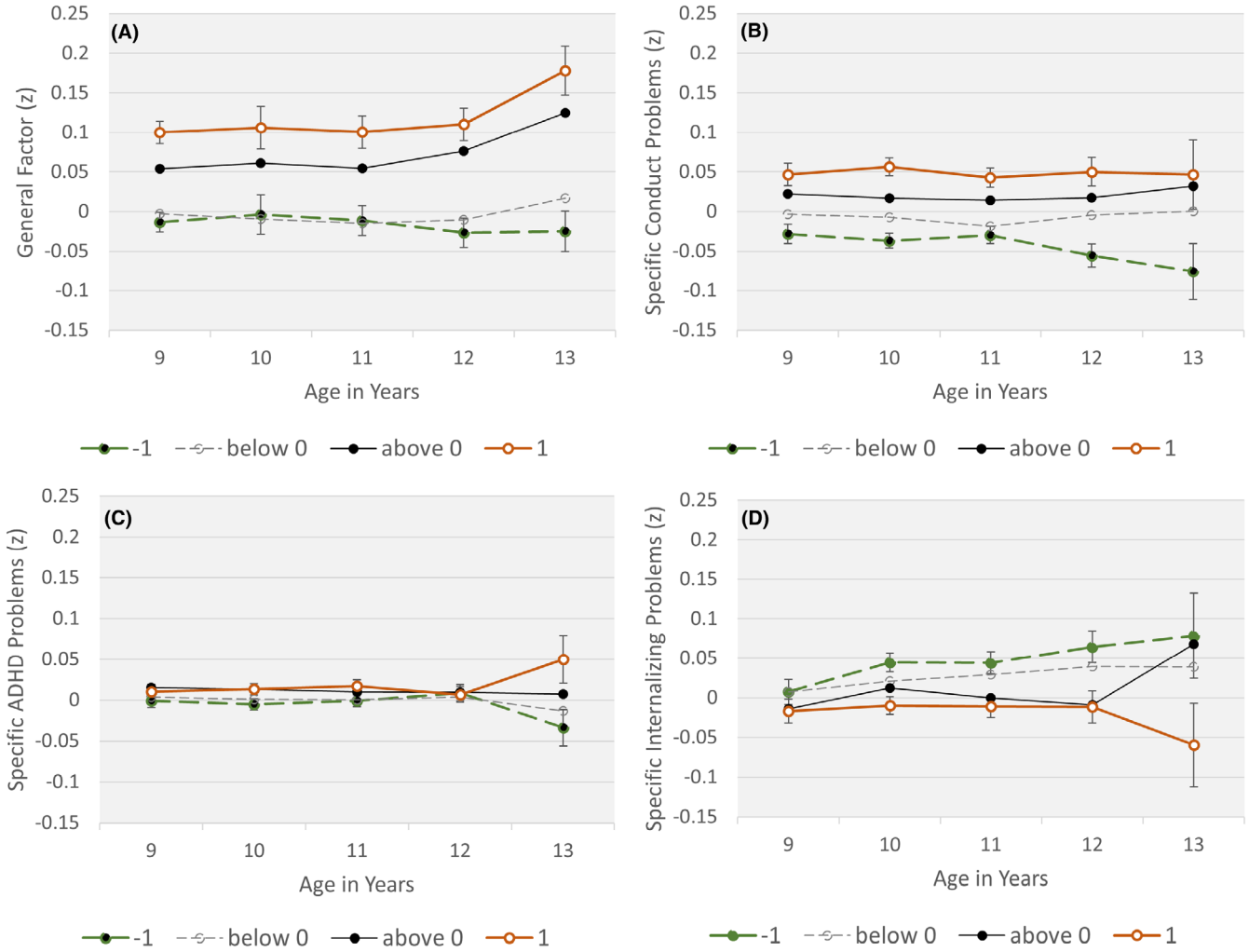

Figure 2 Growth modeling revealed significant associations between the externalizing polygenic score (extPGS) and the intercepts (levels), but not the slopes of change across age, for each general and specific latent factor defined in bifactor measurement models (inversely for internalizing). These findings are illustrated using standardized scores for repeated measures within children on the (a) general factor, (b) specific conduct problems, (c) attention-deficit disorder (ADHD), and (d) internalizing factors for parent-rated psychological problems across 9 through 13 years of age in four groups defined by the child's externalizing extPGS (−1: ≤ 1 *SD* below mean; below 0: > −1 *SD* and <0; above 0: ≥ 0 and <1 *SD*; +1: ≥1 *SD* above mean). Standard error bars are presented for the −1 and +1 *SD* polygenic score groups

polygenic score methods, which are based on different assumptions than cross-sectional twin studies. In particular, the present findings support the utility of the extPGS as a measure of part of the genetic influences on conduct problems quantifiable from common SNPs (Wray et al., 2021) by independently replicating its association with childhood conduct problems in a new sample (Karlsson Linner et al., 2021). Furthermore, the present longitudinal analyses showed that prospective associations of the extPGS with conduct problems endured over 3 years in the same youth.

The present findings of a significant association between the extPGS and TGMV also confirm twin studies (Pol et al., 2007; Zhao et al., 2019), and previous studies using other polygenic scores (Alemany et al., 2019; Alnaes et al., 2019), that brain volume is substantially heritable. Crucially, our findings expand on previous cross-sectional findings that smaller TGMV is associated with greater child behavior problems (Kaczkurkin et al., 2020; Snyder, Hankin, Sandman, Head, & Davis, 2017) by revealing robust *prospective* associations of TGMV with both the variance specific to conduct problems and variance shared by all problem dimensions when they are measured at different points in time during the adolescent transition.

Both the extPGS and TGMV were associated with conduct problems and ADHD when these dimensions were defined in CF models (Tables 1 and 3). This is important, but it does not reveal if these variables were separately associated with each dimension for different reasons or were associated with the variance *shared* by the two dimensions. Results based on bifactor models resolved these alternative explanations. The extPGS was associated with the general factor of psychological problems over time (Figure 2; Table 1; Table S5), indicating

Table 2 Association of the externalizing PGS with regional gray matter volumes controlling age, sex, scanner, and the first 10 principal components for ancestry

	Left			Right		
Cortical regions	β	SE	*p*	β	SE	*p*
Analyses not controlling total intracranial volume						
Frontal cortex						
Frontal pole	**−.050**	**.017**	**.0028**	−.028	.016	.0800
Rostral middle frontal gyrus	−.035	.017	.0344	−.028	.016	.0825
Medial orbitofrontal	**−.052**	**.015**	**.0005**	**−.050**	**.015**	**.0011**
Lateral orbitofrontal	**−.043**	**.017**	**.0099**	**−.058**	**.016**	**.0002**
Pars orbitalis	−.012	.016	.4663	**−.055**	**.016**	**.0008**
Pars opercularis	−.010	.017	.5542	−.021	.017	.1977
Pars triangularis	−.028	.016	.0894	−.017	.017	.3127
Superior frontal gyrus	**−.055**	**.016**	**.0007**	−.051	**.016**	**.0013**
Caudal middle frontal	**−.048**	**.016**	**.0031**	−.026	.016	.1128
Paracentral gyrus	**−.044**	**.017**	**.0094**	−.051	**.016**	**.0019**
Precentral gyrus	**−.065**	**.017**	**<.0001**	−.064	**.016**	**<.0001**
Cingulate						
Rostral anterior cingulate	**−.051**	**.016**	**.0019**	−.028	.017	.0936
Caudal anterior cingulate	−.022	.017	.1994	−.051	**.017**	**.0025**
Posterior cingulate	**−.040**	**.017**	**.0177**	−.029	.016	.0805
Isthmus cingulate	−.016	.016	.2994	−.028	.016	.0785
Temporal lobe						
Banks of superior temporal sulcus	−.016	.017	.3216	−.013	.017	.4593
Superior temporal gyrus	**−.071**	**.016**	**<.0001**	−.054	**.016**	**.0008**
Middle temporal gyrus	−.024	.016	.1342	−.047	**.016**	**.0037**
Inferior temporal gyrus	−.022	.016	.1669	−.049	**.016**	**.0023**
Fusiform gyrus	**−.043**	**.017**	**.0103**	−.041	**.016**	**.0089**
Entorhinal	**−.040**	**.017**	**.0171**	−.023	.016	.1516
Transverse temporal gyrus	**−.047**	**.017**	**.0046**	−.035	.017	.0361
Parahippocampal gyrus	−.027	.017	.1144	−.014	.017	.4122
Temporal pole	−.035	.017	.0335	−.023	.016	.1695
Parietal lobe						
Superior parietal	**−.039**	**.016**	**.0148**	−.047	**.016**	**.0038**
Inferior parietal lobule	−.019	.017	.2800	−.047	**.016**	**.0039**
Postcentral gyrus	−.011	.016	.5073	−.010	.017	.5527
Supramarginal	−.032	.017	.0515	−.032	.017	.0554
Precuneus	−.010	.016	.5485	−.014	.016	.4104
Occipital lobe						
Lateral occipital	**−.057**	**.016**	**.0004**	−.039	**.016**	**.0120**
Lingual gyrus	**−.041**	**.016**	**.0117**	−.043	**.016**	**.0071**
Pericalcarine	**−.040**	**.017**	**.0197**	−.046	**.017**	**.0074**
Cuneus	−.019	.017	.2592	−.024	.017	.1637
Insular cortex	**−.041**	**.015**	**.0072**	−.044	**.015**	**.0039**
Subcortical regions						
Cerebellum cortex	**−.045**	**.016**	**.0038**	−.044	**.015**	**.0042**
Thalamus proper	−.018	.016	.2569	−.011	.016	.4907
Caudate	−.004	.017	.8039	.003	.017	.8432
Putamen	.021	.017	.2022	.010	.016	.5165
Pallidum	.004	.015	.7743	.016	.015	.3107
Hippocampus	−.032	.016	.0405	−.025	.016	.1192
Amygdala	−.022	.015	.1440	−.017	.016	.2744
Accumbens area	−.004	.016	.7830	−.034	.017	.0454
Ventral diencephalon	−.034	.015	.0256	−.029	.015	.0613
Brain stem	−.020	.015	.1797	–	–	–
Analyses controlling total intracranial volume						
Frontal Lobe						
Frontal pole	−.037	.016	.0216	−.016	.016	.3033
Rostral middle frontal gyrus	−.014	.014	.3321	−.007	.014	.6063
Medial orbitofrontal	−.031	.012	.0128	−.027	.012	.0307
Lateral orbitofrontal	−.017	.013	.1728	−.034	.012	.0051
Pars orbitalis	.006	.015	.6778	−.038	.015	.0095
Pars opercularis	.005	.016	.7714	−.007	.015	.6532
Pars triangularis	−.014	.015	.3432	−.003	.016	.8387
Superior frontal gyrus	−.031	.013	.0152	−.027	.013	.0333
Caudal middle frontal	−.030	.015	.0405	−.007	.014	.6131
Paracentral gyrus	−.026	.015	.0938	−.033	.015	.0250
Precentral gyrus	−.043	.014	.0019	−.044	.014	.0017

(continues)

Table 2 (continued)

Cortical regions	Left			Right		
	β	SE	*p*	β	SE	*p*
Cingulate						
Rostral anterior cingulate	−.030	.014	.0336	−.010	.015	.4927
Caudal anterior cingulate	−.006	.016	.6954	−.037	.016	.0198
Posterior cingulate	−.020	.015	.1730	−.008	.014	.5629
Isthmus cingulate	.001	.014	.9258	−.011	.015	.4605
Temporal lobe						
Banks of superior temporal sulcus	−.002	.015	.9116	.004	.015	.7902
Superior temporal gyrus	**−.048**	**.013**	**.0003**	−.031	.013	.0178
Middle temporal gyrus	−.003	.013	.8366	−.024	.013	.0624
Inferior temporal gyrus	−.001	.013	.9388	−.028	.013	.0368
Fusiform gyrus	−.021	.014	.1317	−.018	.013	.1448
Entorhinal	−.029	.016	.0699	−.013	.016	.3922
Transverse temporal gyrus	−.032	.015	.0397	−.018	.015	.2315
Parahippocampal gyrus	−.013	.016	.4287	.003	.015	.8579
Temporal pole	−.022	.016	.1573	−.008	.015	.5905
Parietal Lobe						
Superior parietal	−.020	.014	.1691	−.028	.014	.0527
Inferior parietal lobule	.001	.015	.9465	−.028	.014	.0478
Postcentral gyrus	.010	.014	.4683	.012	.014	.4049
Supramarginal	−.013	.014	.3654	−.012	.014	.4117
Precuneus	.012	.014	.3851	.009	.013	.5024
Occipital lobe						
Lateral occipital	−.040	.014	.0051	−.021	.014	.1315
Lingual gyrus	−.025	.015	.0990	−.025	.014	.0763
Pericalcarine	−.027	.016	.0971	−.033	.016	.0455
Cuneus	−.007	.016	.6667	−.011	.016	.5171
Insular cortex	−.019	.012	.1292	−.022	.013	.0831
Subcortical regions						
Cerebellum cortex	−.024	.013	.0592	−.024	.013	.0626
Thalamus proper	.009	.012	.4457	.016	.011	.1476
Caudate	.017	.014	.2227	.025	.014	.0652
Putamen	.040	.015	.0070	.030	.014	.0303
Pallidum	.023	.014	.0955	.036	.013	.0055
Hippocampus	−.009	.013	.4873	−.003	.013	.8347
Amygdala	−.004	.013	.7621	.004	.013	.7647
Accumbens area	.015	.014	.2726	−.013	.015	.3606
Ventral diencephalon	−.007	.011	.5088	−.002	.011	.8633
Brain stem	.004	.012	.7541	–	–	–

Tabled *p*'s are unadjusted; coefficients in bold are significant at $p < .05$ after false discovery rate adjustment for 87 tests of significance in family A analyses (not controlling total intracranial volume) and 87 tests of significance in family B analyses (controlling total intracranial volume).

that the extPGS predicted the variance shared by all measured problems. Nonetheless, the extPGS also was associated with the orthogonal specific conduct problems factor over time. Similar results were found for associations with TGMV (Table 3). Thus, both extPGS and TGMV were associated with conduct problems through both the variance specific to conduct problems and through the variance that conduct problems shared with other dimensions. In contrast, associations of extPGS and TGMV with ADHD were primarily through the variance shared with all problem dimensions. Alternatively, this finding could reflect the weaker measurement of ADHD than the other dimensions.

Notably, a comparison of the results of analyses based on CF and bifactor measurement models revealed potentially important findings on internalizing problems. Consistent with previous hypotheses that low levels of anxiety are associated with conduct problems (Frick, Lilienfeld, Ellis, Loney, & Silverthorn, 1999), the extPGS polygenic score was associated *inversely* with variance uniquely associated with specific internalizing problems, but only in bifactor models that partitioned the variance that is not shared by internalizing with other problems.

Tests of mediated pathways

The present analyses were the first to combine polygenic scores, brain structure, and longitudinal data on child psychological problems in the same models to examine potential temporally ordered risk pathways from genome to brain to behavior. Formal tests of statistical mediation found evidence that the association of the extPGS with conduct problems was significantly mediated by TGMV in all waves

Table 3 Results of tests of associations between standardized total gray matter volume and dimensions of psychological problems defined alternatively by CF and bifactor modeling of parent ratings of European ancestry children in the baseline assessment, and first and second annual follow-up assessments, controlling age, sex, scanner, and the first 10 principle components for ancestry among European ancestry children with nonmissing externalizing polygenic scores

Correlated factors models of psychological problem ratings								
Internalizing			Conduct problems			ADHD		
β (SE)	R^2	*p*	β (SE)	R^2	*p*	β (SE)	R^2	*p*
Baseline assessment (N = 4,063)								
−.089 (.024)	**1.0%**	**<.0001**	**−.216 (.023)**	**4.7%**	**<.0001**	**−.175 (.023)**	**3.0%**	**<.0001**
Follow-up Wave 1 assessment (N = 3,940)								
−.051 (.024)	**<1.0%**	**.036**	**−.176 (.025)**	**3.1%**	**<.0001**	**−.131 (.024)**	**1.7%**	**<.0001**
Follow-up Wave 2 assessment (N = 3,705)								
−.056 (.025)	**<1.0%**	**.025**	**−.196 (.024)**	**3.8%**	**<.0001**	**−.136 (.024)**	**1.8%**	**<.0001**

Correlated factors models of psychological problem ratings											
General psychopathology			Specific internalizing			Specific conduct problems			Specific ADHD		
β (SE)	R^2	*p*	β (SE)	R^2	*p*	β (SE)	R^2	*p*	β (SE)	R^2	*p*
Baseline assessment (N = 4,063)											
−.177 (.025)	**3.1%**	**<.0001**	**.062 (.027)**	**<1.0%**	**.021**	**−.135 (.030)**	**1.8%**	**<.0001**	−.041 (.032)		.210
Follow-up Wave 1 assessment (N = 3,940)											
−.128 (.025)	**1.6%**	**<.0001**	**.058 (.024)**	**<1.0%**	**.016**	**−.145 (.032)**	**2.1%**	**<.0001**	−.026 (.033)		.431
Follow-up Wave 2 assessment (N = 3,705)											
−.154 (.027)	**2.4%**	**<.0001**	**.078 (.28)**	**1.0%**	**.005**	**−.129 (.032)**	**1.7%**	**<.0001**	.003 (.035)		.942

Factor loadings in the measurement models used in each predictive regression were fixed based on unstandardized factor loadings in a bifactor model without covariates. Tabled *p*'s are unadjusted; coefficients in bold are significant at $p < .05$ after false discovery rate correction for 21 tests of significance.

when conduct problems were defined in either CF or bifactor models. In contrast, no significant mediation of the association between the extPGS and the general factor was found (Table 4). Rather, most of the concurrent and prospective associations between the extPGS and psychological problems were found to be direct, rather than being mediated by TGMV. That is, these findings suggest that TGMV statistically mediates the association of extPGS with only the variance that is specific to conduct problems and not the variance that conduct problems share with other problem dimensions. The genetic sensitivity analyses suggested that the observed mediation by TGMV of the association between extPGS and specific conduct problems in all assessment waves was not an artifact of bias. Confidence in this finding is substantial because the twin-based estimate used was quite high (h^2 = 0.81), providing a conservative test of genetic confounding.

Limitations and future directions

Because the current analyses were based only on children of European ancestry, the findings are not generalizable to other populations. Individuals of non-European ancestry were excluded because the extPGS was derived from samples of individuals of only European descent (Karlsson Linner et al., 2021). Polygenic scores generally underperform when there is mismatch in ancestries between the discovery and target samples, which could result in misrepresentations of the role of polygenic risk in other populations (Duncan et al., 2019; Martin et al., 2019). It is imperative, however, that future molecular genetic studies address the role of genetic risk for psychological problems in non-European populations using appropriate methods.

Across the age span of 9–13 years when many children are entering puberty, the extPGS was associated with levels of problem behaviors over time, but not the rate of change in these problems with increasing age (Figure 2). Because greater changes in psychological problems with age are expected later in adolescence, however, it is imperative to address this issue again when additional waves of the ABCD Study data are available.

Although the present findings are revealing, they cannot support the inference that variations in the SNPs that define the extPGS contribute causally to variations in brain, which causally influence risk for conduct problems. This is because differences in SNPs and brain structure between children from different families are confounded by many potential environmental differences between families, such as maltreatment and economic deprivation (D'Onofrio, Sjolander, Lahey, Lichtenstein, & Oberg, 2020; Lahey & D'Onofrio, 2010). To infer causation, it will be necessary to move from case–control and other designs that compare different children from different families to within-family designs that can break

Table 4 Standardized regression coefficients with 99% bootstrapped confidence intervals for the results of tests of direct association paths from the externalizing polygenic score (PGS) to each dimension of psychological problems and indirect paths from the externalizing PGS to problem dimensions through total gray matter volume in each of three annual assessments, with problem dimensions alternatively defined in correlated factors models (upper panels) or bifactor models (lower panels)

	Indirect path			Direct path		
Problem dimensions	β	99% CI	R^2 (%)	β	99% CI	R^2 (%)
Correlated factors models of outcomes						
Baseline assessment						
Internalizing	**.005**	**0.001, 0.011**	**<.1**	.034	−0.019, 0.091	
Conduct problems	**.012**	**0.003, 0.022**	**<.1**	**.135**	**0.089, 0.186**	**1.2**
ADHD	**.010**	**0.002, 0.018**	**<.1**	**.109**	**0.059, 0.162**	**1.2**
Follow-up Wave 1 assessment						
Internalizing	.003	−0.001, 0.009		.024	−0.031, 0.083	
Conduct problems	**.010**	**0.003, 0.020**	**<.1**	**.158**	**0.105, 0.209**	
ADHD	**.007**	**0.002, 0.015**	**<.1**	**.126**	**0.078, 0.177**	
Follow-up Wave 2 assessment						
Internalizing	.003	−0.001, 0.008		.030	−0.023, 0.082	
Conduct problems	**.011**	**0.003, 0.020**	**<.1**	**.157**	**0.104, 0.211**	**2.5**
ADHD	**.007**	**0.002, 0.014**		**.111**	**0.059, 0.155**	**1.2**
Bifactor models of outcomes						
Baseline assessment						
General factor	**.009**	**0.001, 0.018**	**<.1**	**.087**	**0.027, 0.147**	**1.0**
Specific internalizing	−.005	−0.015, 0.000		−.047	−0.105, 0.027	
Specific conduct problems	**.011**	**0.003, 0.114**	**<.1**	**.140**	**0.077, 0.214**	**1.0**
Specific ADHD	.004	−0.005, 0.118		.053	−0.024, 0.134	
Follow-up Wave 1 assessment						
General factor	.005	−0.064, 0.014		**.110**	**0.021 0.166**	**1.2**
Specific internalizing	−.005	−0.075, 0.063		−.068	−0.130, 0.083	
Specific conduct problems	**.014**	**0.002, 0.226**	**<.1**	**.128**	**0.014, 0.208**	**1.6**
Specific ADHD	.007	−0.005, 0.213		.059	−0.050, 0.137	
Follow-up Wave 2 assessment						
General factor	.009	−0.011, 0.022		**.108**	**0.048, 0.165**	**1.2**
Specific internalizing	**−.008**	**−0.023, −0.001**	**<.1**	**−.070**	**−0.136, −0.005**	**<.1**
Specific conduct problems	**.009**	**0.002, 0.047**	**<.1**	**.137**	**0.071, 0.214**	**1.9**
Specific ADHD	−.001	−0.025, 0.066		.026	−0.058, 0.115	

the confounding of genes and brain with unmeasured environmental variables. Nonetheless, it is encouraging that Karlsson Linner et al (Karlsson Linner et al., 2021) conducted tests of the associations of the extPGS with externalizing phenotypes in very large samples of adults of European ancestry that contain many siblings and found few differences in the magnitudes of associations of the extPGS with adult externalizing behavior when estimated in comparisons conducted within families versus between families. Additionally, a polygenic score for educational attainment was found to significantly predict impairing alcohol, nicotine, and marijuana use in a large sample of young adults using less confounded sibling comparisons (Salvatore et al., 2020). There are insufficient numbers of siblings in the ABCD Study sample to conduct adequately powered within-family tests, but the results of the present analyses encourage future investments in research designs that can map causal paths from genetic variation through brain to significant psychological problems. This would advance understanding of the biological nature of psychological problems and could ultimately lead to improved strategies of prevention and intervention.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article:

Figure S1. Illustration of correlations between true and predicted values of general factor and specific problem dimensions between random split-half training and testing sets from the ABCD Study.

Table S1. Summary of demographic characteristics of the 4,475 children with nonmissing data on brain volumes, polygenic risk scores, and parent rating of child psychological problems at baseline.

Table S2. Standardized factor loadings for the confirmatory correlated factors model with three factors based on CBCL items measured in three annual assessments in children of European descent.

Table S3. Standardized factor loadings for the confirmatory bifactor model with three specific factors based on CBCL items measured in three annual assessments in children of European descent.

Table S4. Psychometric indices for each latent factor defined in bifactor models of parent-rated psychological problems in the baseline and first and second annual follow-up assessments.

Table S5. Results of longitudinal growth models testing associations of the externalizing polygenic score and

parent-classified child sex with extracted general and specific factors of psychological problems defined by bifactor models of parent ratings of their children across 9–13 years of age, controlling for the first 10 principal components for ancestry.

Table S6. Results of sensitivity analyses conducted in Gsens62 to evaluate bias in the estimates of mediated associations of the polygenic score with specific conduct problems through total gray matter volume assessed in the three annual assessment owing to imperfect measurement of each variable.

Acknowledgements

This research was supported by grants UG3DA045251 (BBL) from the National Institute on Drug Abuse, R01MH098098 (BBL), R01MH117014 (TMM), R00MH117274 (ANK), and T32MH18921 (ELD is a trainee on this grant) from the National Institute of Mental Health, UL1TR000430 (BBL) and UL1TR000445 (BBL) from the National Center for Advancing Translational Sciences, a Young Investigator Grant from the Brain & Behavior Research Foundation (ANK), a Sloan Research Fellowship (ANK), a David H. and Beverly A. Barlow grant from the American Psychological Foundation (ANK), and the Lifespan Brain Institute of the University of Pennsylvania and the Children's Hospital of Philadelphia (TMM). This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. 1937963 (GER). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors have declared that they have no competing or potential conflicts of interest.

Author contributions

Benjamin Lahey contributed to the conceptualization, methodology, statistical analysis, writing original draft, review and editing, and funding acquisition. E. Leighton Durham contributed to the writing of the original draft, review, editing, and visualization. Sarah Brislin contributed to the methodology, sensitivity analyses, and resources. Peter Barr contributed to the methodology, sensitivity analyses, resources, data curation, review, and editing. Danielle Dick contributed to the methodology, resources, data curation, review, and editing. Tyler Moore and Brandon Pierce contributed to the methodology, review, and editing. Lin Tong contributed to the methodology, resources, and data curation. Gabrielle E. Reimann and Hee Jung Jeong contributed to the review and editing. Randolph Dupont contributed to the data curation. Antonia N. Kaczkurkin contributed to the conceptualization, methodology, statistical analysis, writing of original draft, review, editing, funding acquisition, and project administration.

Data availability statement

Data used in the preparation of this article are obtained from the Adolescent Brain Cognitive DevelopmentSM (ABCD) Study (https://abcdstudy.org), held in the NIMH Data Archive (NDA). This is a multisite, longitudinal study designed to recruit more than 10,000 children age 9–10 and follow them over 10 years into early adulthood. The ABCD Study® is supported by the National Institutes of Health and additional federal partners under award numbers U01DA041048, U01DA050989, U01DA051016, U01DA041022, U01DA051018, U01DA051037, U01DA050987, U01DA041174, U01DA041106, U01DA041117, U01DA041028, U01DA041134, U01DA050988, U01DA051039, U01DA041156, U01DA041025, U01DA041120, U01DA051038, U01DA041148, U01DA041093, U01DA041089, U24DA041123, and U24DA041147. A full list of supporters is available at https://abcdstudy.org/federal-partners.html. A listing of participating sites and a complete listing of the study investigators can be found at https://abcdstudy.org/consortium_members/. ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in the analysis or writing of this report. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators. The ABCD data repository grows and changes over time. The ABCD data used in this report came from RRID: SCR_015769, DOI 10.15154/1523041 (data release 4.0) and NDA study DOI 10.15154/1527791. DOIs can be found at https://nda.nih.gov/abcd/study-information. The Externalizing Consortium has been supported by the National Institute on Alcohol Abuse and Alcoholism (R01AA015416 – administrative supplement to DMD), and the National Institute on Drug Abuse (R01DA050721 to DMD). Additional funding for investigator effort has been provided by K02AA018755, U10AA008401, and P50AA022537, as well as a European Research Council Consolidator Grant (647,648 EdGe to Koellinger). The content is solely the responsibility of the authors and does not necessarily represent the official views of the above funding bodies. The Externalizing Consortium would like to thank the following groups for making the research possible: 23andMe, Add Health, Vanderbilt University Medical Center's BioVU, Collaborative Study on the Genetics of Alcoholism (COGA), the Psychiatric Genomics Consortium's Substance Use Disorders working group, UK10K Consortium, UK Biobank, and Philadelphia Neurodevelopmental Cohort. More details about the Externalizing Consortium can be found at Externalizing.org.

Endnote

i. The term 'psychological problems' is used as a denotative synonym for 'psychopathology', but is preferred because psychopathology literally means 'sick mind', which fosters stigma (Lahey, 2021).

Correspondence

Antonia Kaczkurkin, Department of Psychology, College of Arts and Science, Vanderbilt University, PMB 407817, 2301 Vanderbilt Place, Nashville, TN 37240-

7817, USA; Email: antonia.kaczkurkin@vanderbilt.eduBenjamin B. Lahey, Department of Public Health Sciences (MC2000), 5841 S. Maryland Avenue, Chicago, IL 60637, USA; Email: blahey@uchicago.edu

Key points

- A polygenic score for externalizing behavior developed on data from other samples predicted from 2.0% to 2.9% of the variance in conduct problems in each of three annual assessments of a large longitudinal cohort of 9- to 10-year-old children at baseline.
- Longitudinal analyses revealed that the externalizing polygenic score predicted levels of problem behaviors over 3 years during late childhood and early adolescence, but not their slopes of change.
- The externalizing polygenic score also predicted total brain volume in the baseline assessment and total brain volume robustly predicted from 1.7% to 2.1% of the variance specific to conduct problems and from 1.6% to 3.1% of the variance common to all problems across the annual assessments.
- A modest but significant proportion of polygenic liability for specific conduct problems in each wave was statistically mediated by total brain volume, indicating the plausibility of identifying pathways from genetic risk through variations in brain structure to serious conduct problems.

References

Achenbach, T.M. (2009). *The Achenbach system of empirically based Assessemnt (ASEBA): Development, findings, theory, and applications*. Burlington, VT: University of Vermont Research Center for Children, Youth and Families.

Albaugh, M.D., Hudziak, J.J., Ing, A., Chaarani, B., Barker, E., Jia, T., … Potter, A. (2019). White matter microstructure is associated with hyperactive/inattentive symptomatology and polygenic risk for attention-deficit/hyperactivity disorder in a population-based sample of adolescents. *Neuropsychopharmacology*, *44*, 1597–1603.

Alemany, S., Jansen, P.R., Muetzel, R.L., Marques, N., el Marroun, H., Jaddoe, V.W.V., … White, T. (2019). Common polygenic variations for psychiatric disorders and cognition in relation to brain morphology in the general pediatric population. *Journal of the American Academy of Child and Adolescent Psychiatry*, *58*, 600–607.

Allegrini, A.G., Cheesman, R., Rimfeld, K., Selzam, S., Pingault, J.B., Eley, T.C., & Plomin, R. (2020). The p factor: Genetic analyses support a general dimension of psychopathology in childhood and adolescence. *Journal of Child Psychology and Psychiatry*, *61*, 30–39.

Alnaes, D., Kaufmann, T., van der Meer, D., Córdova-Palomera, A., Rokicki, J., Moberget, T., … Cervenka, S. (2019). Brain heterogeneity in schizophrenia and its association with polygenic risk. *JAMA Psychiatry*, *76*, 739–748.

Baurley, J.W., Edlund, C.K., Pardamean, C.I., Conti, D.V., & Bergen, A.W. (2016). Smokescreen: A targeted genotyping array for addiction research. *BMC Genomics*, *17*, 1–12.

Beauchaine, T.P., Hinshaw, S.P., & Pang, K.L. (2010). Comorbidity of attention-deficit/hyperactivity disorder and early-onset conduct disorder: Biological, environmental, and developmental mechanisms. *Clinical Psychology: Science and Practice*, *17*, 327–336.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B*, *57*, 289–300.

Cao, Z.P., Cupertino, R.B., Ottino-Gonzalez, J., Murphy, A., Pancholi, D., Juliano, A., … ENIGMA Addiction Working Group. (2022). Cortical profiles of numerous psychiatric disorders and normal development share a common pattern. *Molecular Psychiatry*, *2022*, 698–709.

Casey, B.J., Cannonier, T., Conley, M.I., Cohen, A.O., Barch, D.M., Heitzeg, M.M., … ABCD Imaging Acquisition Workgroup. (2018). The adolescent brain cognitive development (ABCD) study: Imaging acquisition across 21 sites. *Developmental Cognitive Neuroscience*, *32*, 43–54.

Caspi, A., Houts, R.M., Belsky, D.W., Goldman-Mellor, S.J., Harrington, H.L., Israel, S., … Moffitt, T.E. (2014). The p factor: One general psychopathology factor in the structure of psychiatric disorders? *Clinical Psychological Science*, *2*, 119–137.

Choi, K.W., Wilson, M., Ge, T., Kandola, A., Patel, C.J., Lee, S.H., & Smoller, J.W. (2022). Integrative analysis of genomic and exposomic influences on youth mental health. *Journal of Child Psychology and Psychiatry*, *63*, 1196–1205.

Cosgrove, V.E., Rhee, S.H., Gelhorn, H.L., Boeldt, D., Corley, R.C., Ehringer, M.A., … Hewitt, J.K. (2011). Structure and etiology of co-occurring internalizing and externalizing disorders in adolescents. *Journal of Abnormal Child Psychology*, *39*, 109–123.

Costello, E.J., Egger, H., & Angold, A. (2005). 10-year research update review: The epidemiology of child and adolescent psychiatric disorders: I. Methods and public health burden. *Journal of the American Academy of Child and Adolescent Psychiatry*, *44*, 972–986.

Desikan, R.S., Segonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., … Killiany, R.J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*, 968–980.

D'Onofrio, B.M., Sjolander, A., Lahey, B.B., Lichtenstein, P., & Oberg, A.S. (2020). Accounting for confounding in observational studies. In T. Widiger & T.D. Cannon (Eds.), *Annual review of clinical psychology*, Vol 16 (pp. 25–48). San Mateo, CA: Annual Reviews.

Du Rietz, E., Pettersson, E., Brikell, I., Ghirardi, L., Chen, Q., Hartman, C., … Kuja-Halkola, R. (2021). Overlap between attention-deficit hyperactivity disorder and neurodevelopmental, externalising and internalising disorders: Separating unique from general psychopathology effects. *The British Journal of Psychiatry*, *218*, 35–42.

Duncan, L., Shen, H., Gelaye, B., Meijsen, J., Ressler, K., Feldman, M., … Domingue, B. (2019). Analysis of polygenic risk score usage and performance in diverse human populations. *Nature Communications*, *10*, 103328. https://doi.org/10.1038/s41467-019-11112-0

Durham, E.L., Jeong, H.J., Moore, T.M., Dupont, R.M., Cardenas-Iniguez, C., Cui, Z., … Kaczkurkin, A.N. (2021). Association of Gray Matter Volumes with general and specific dimensions of psychopathology in children. *Neuropsychopharmacology*, *11*, 1333–1339.

Fairchild, G., Hawes, D.J., Frick, P.J., Copeland, W.E., Odgers, C.L., Franke, B., … de Brito, S.A. (2019). Conduct disorder. *Nature Reviews Disease Primers*, *5*, 43. https://doi.org/10.1038/s41572-019-0095-y

Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., … Dale, A.M. (2002). Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron*, *33*, 341–355.

Forbes, M.K., Greene, A.L., Levin-Aspenson, H.F., Watts, A.L., Hallquist, M., Lahey, B.B., … Krueger, R.F. (2021). Three recommendations based on a comparison of the reliability and validity of the predominant models used in research on the empirical structure of psychopathology. *Journal of Abnormal Psychology*, *130*, 297–317.

Forbes, M.K., Sunderland, M., Rapee, R.M., Batterham, P.J., Calear, A.L., Carragher, N., … Krueger, R.F. (2021). A detailed hierarchical model of psychopathology: From individual symptoms up to the general factor of psychopathology. *Clinical Psychological Science*, *9*, 139–168.

Frick, P.J., Lilienfeld, S.O., Ellis, M., Loney, B., & Silverthorn, P. (1999). The association between anxiety and psychopathy dimensions in children. *Journal of Abnormal Child Psychology*, *27*, 383–392.

Hagler, D.J., Hatton, S., Cornejo, M.D., Makowski, C., Fair, D.A., Dick, A.S., … Hyde, L.W. (2019). Image processing and analysis methods for the adolescent brain cognitive development study. *NeuroImage*, *202*, 1–17. https://doi.org/10.1016/j.neuroimage.2019.116091

Hancock, G.R., & Mueller, R.O. (2001). Rethinking construct reliability within latent variable systems. In R. Cudeck & D. Sörbom (Eds.), *Structural equation modeling: Present and future—A festschrift in honor of Karl Jöreskog* (pp. 195–216). Elgin, IL: Scientific Software International.

Heeringa, S.G., & Berglund, P.A. (2018). *A guide for population-based analysis of the adolescent brain cognitive development (ABCD) study baseline data*. Ann Arbor, MI: Institute for Social Research.

Hicks, B.M., Krueger, R.F., Iacono, W.G., McGue, M., & Patrick, C.J. (2004). Family transmission and heritability of externalizing disorders—A twin-family study. *Archives of General Psychiatry*, *61*, 922–928.

Hoogman, M., Bralten, J., Hibar, D.P., Mennes, M., Zwiers, M.P., Schweren, L.S.J., … Franke, B. (2017). Subcortical brain volume differences in participants with attention deficit hyperactivity disorder in children and adults: A cross-sectional mega-analysis. *Lancet Psychiatry*, *4*, 310–319.

Hoogman, M., Muetzel, R., Guimaraes, J.P., Shumskaya, E., Mennes, M., Zwiers, M.P., … Franke, B. (2019). Brain imaging of the cortex in ADHD: A coordinated analysis of large-scale clinical and population-based samples. *American Journal of Psychiatry*, *176*, 531–542.

Kaczkurkin, A.N., Moore, T.M., Sotiras, A., Xia, C.H., Shinohara, R.T., & Satterthwaite, T.D. (2020). Approaches to defining common and dissociable neurobiological deficits associated with psychopathology in youth. *Biological Psychiatry*, *88*, 51–62.

Karlsson Linner, R., Mallard, T.T., Barr, P.B., Sanchez-Roige, S., Madole, J.W., Driver, M.N., et al. (2021). Multivariate analysis of 1.5 million people identifies genetic associations with traits related to self-regulation and addiction. *Nature Neuroscience*, *24*, 1367–1376.

Krueger, R.F., Hobbs, K.A., Conway, C.C., Dick, D.M., Dretsch, M.N., Eaton, N.R., … HiTOP Utility Workgroup. (2021). Validity and utility of hierarchical taxonomy of psychopathology (HiTOP): II. Externalizing superspectrum. *World Psychiatry*, *20*, 171–193.

Laceulle, O.M., Chung, J.M., Vollebergh, W.A.M., & Ormel, J. (2020). The wide-ranging life outcome correlates of a general psychopathology factor in adolescent psychopathology. *Personality and Mental Health*, *14*, 9–29.

Lahey, B.B. (2021). *Dimensions of psychological problems: Replacing diagnostic categories with a more science-based and less stigmatizing alternative*. New York, NY: Oxford University Press.

Lahey, B.B., Applegate, B., Waldman, I.D., Loft, J.D., Hankin, B.L., & Rick, J. (2004). The structure of child and adolescent psychopathology: Generating new hypotheses. *Journal of Abnormal Psychology*, *113*, 358–385.

Lahey, B.B., & D'Onofrio, B.M. (2010). All in the family: Comparing siblings to test causal hypotheses regarding environmental influences on behavior. *Current Directions in Psychological Science*, *19*, 319–323.

Lahey, B.B., Krueger, R.F., Rathouz, P.J., Waldman, I.D., & Zald, D.H. (2017). A hierarchical causal taxonomy of psychopathology across the life span. *Psychological Bulletin*, *143*, 142–186.

Lahey, B.B., Van Hulle, C.A., Singh, A.L., Waldman, I.D., & Rathouz, P.J. (2011). Higher-order genetic and environmental structure of prevalent forms of child and adolescent psychopathology. *Archives of General Psychiatry*, *68*, 181–189.

Mackey, S., Allgaier, N., Chaarani, B., Spechler, P., Orr, C., Bunn, J., … the ENIGMA Addiction Working Group. (2019). Mega-analysis of gray matter volume in substance dependence: General and substance-specific regional effects. *American Journal of Psychiatry*, *176*, 119–128.

Martin, A.R., Kanai, M., Kamatani, Y., Okada, Y., Neale, B.M., & Daly, M.J. (2019). Clinical use of current polygenic risk scores may exacerbate health disparities. *Nature Genetics*, *51*, 584–591.

Maxwell, S.E., Cole, D.A., & Mitchell, M.A. (2011). Bias in cross-sectional analyses of longitudinal mediation: Partial and complete mediation under an autoregressive model. *Multivariate Behavioral Research*, *46*, 816–841.

Mooney, M.A., Bhatt, P., Hermosillo, R.J.M., Ryabinin, P., Nikolas, M., Faraone, S.V., … Nigg, J.T. (2021). Smaller total brain volume but not subcortical structure volume related to common genetic risk for ADHD. *Psychological Medicine*, *51*, 1279–1288.

Moore, T.M., Kaczkurkin, A.N., Durham, E.L., Jeong, H.J., McDowell, M.G., Dupont, R.M., … Lahey, B.B. (2020). Criterion validity and relationships between alternative hierarchical dimensional models of general and specific psychopathology. *Journal of Abnormal Psychology*, *129*, 677–688.

Muthén, L.K., & Muthén, B.O. (1998-2017). *Mplus user's guide* (8th edn). Los Angeles, CA: Muthén & Muthén.

Neumann, A., Nolte, I.M., Pappa, I., Ahluwalia, T.S., Pettersson, E., Rodriguez, A., … Tiemeier, H. (2022). A genome-wide association study of total child psychiatric problems scores. *PLoS One*, *17*, e0273116.

Patel, Y., Parker, N., Salum, G.A., Pausova, Z., & Paus, T. (2022). General psychopathology, cognition, and the cerebral cortex in 10-year-old children: Insights from the adolescent brain cognitive development study. *Frontiers in Human Neuroscience*, *15*, 15781554. https://doi.org/10.3389/fnhum.2021.781554

Peña, E.A., Habiger, J.D., & Wu, W. (2011). Power-enhanced multiple decision functions controlling family-wise error and false discovery rates. *Annals of Statistics*, *39*, 556.

Pettersson, E., Lichtenstein, P., Larsson, H., Song, J., Attention Deficit/Hyperactivity Disorder Working Group of the iPSYCH-Broad-PGC Consortium, Autism Spectrum Disorder Working Group of the iPSYCH-Broad-PGC Consortium, Bipolar Disorder Working Group of the PGC, Eating Disorder Working Group of the PGC, Major Depressive Disorder Working Group of the PGC, Obsessive Compulsive Disorders and Tourette Syndrome Working Group of the PGC, Schizophrenia CLOZUK, Substance Use Disorder Working Group of the PGC, Agrawal, A., … Polderman, T.J.C. (2019). Genetic influences on eight psychiatric disorders based on

family data of 4 408 646 full and half-siblings, and genetic data of 333 748 cases and controls. *Psychological Medicine*, *49*, 1166–1173.

Pingault, J.B., Allegrini, A.G., Odigie, T., Frach, L., Baldwin, J.R., Rijsdijk, F., & Dudbridge, F. (2022). Research review: How to interpret associations between polygenic scores, environmental risks, and phenotypes. *Journal of Child Psychology and Psychiatry*, *63*, 1125–1139.

Pol, H.E.H., Peper, J.S., Boomsma, D.I., & Kahn, R.S. (2007). Genetic contributions to human brain morphology and intelligence. *European Neuropsychopharmacology*, *17*, S217.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., … Sham, P.C. (2007). PLINK: A toolset for whole-genome association and population-based linkage analysis. *American Journal of Human Genetics*, *81*, 559–575.

RCoreTeam. (2021). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rhee, S.H., & Waldman, I.D. (2002). Genetic and environmental influences on antisocial behavior: A meta-analysis of twin and adoption studies. *Psychological Bulletin*, *128*, 490–529.

Rodriguez, A., Reise, S.P., & Haviland, M.G. (2016). Evaluating Bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, *21*, 137–150.

Salvatore, J.E., Barr, P.B., Stephenson, M., Aliev, F., Kuo, S.I.C., Su, J., … Dick, D.M. (2020). Sibling comparisons elucidate the associations between educational attainment polygenic scores and alcohol, nicotine and cannabis. *Addiction*, *115*, 337–346.

Shrout, P.E. (2011). Commentary: Mediation analysis, causal process, and cross-sectional data. *Multivariate Behavioral Research*, *46*, 852–860.

Snyder, H.R., Hankin, B.L., Sandman, C.A., Head, K., & Davis, E.P. (2017). Distinct patterns of reduced prefrontal and limbic gray matter volume in childhood general and internalizing psychopathology. *Clinical Psychological Science*, *5*, 1001–1013.

StataCorp. (2019). *Stata statistical software: Release 16*. College Station, TX: StataCorp LLC.

Waszczuk, M.A., Eaton, N.R., Krueger, R.F., Shackman, A.J., Waldman, I.D., Zald, D.H., … Kotov, R. (2020). Redefining phenotypes to advance psychiatric genetics: Implications from hierarchical taxonomy of psychopathology. *Journal of Abnormal Psychology*, *129*, 143–161.

Watson, D., Levin-Aspenson, H.F., Waszczuk, M.A., Conway, C.C., Dalgleish, T., Dretsch, M.N., … HiTOP Utility Workgroup. (2022). Validity and utility of hierarchical taxonomy of psychopathology (HiTOP): III. Emotional dysfunction superspectrum. *World Psychiatry*, *21*, 26–54.

Watts, A.L., Poore, H.E., & Waldman, I.D. (2019). Riskier tests of the validity of the Bifactor model of psychopathology. *Clinical Psychological Science*, *7*, 1285–1303.

Wray, N.R., Lin, T., Austin, J., McGrath, J.J., Hickie, I.B., Murray, G.K., & Visscher, P.M. (2021). From basic science to clinical application of polygenic risk scores a primer. *JAMA Psychiatry*, *78*, 101–109.

Zald, D.H., & Lahey, B.B. (2017). Implications of the hierarchical structure of psychopathology for psychiatric neuroimaging. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *2*, 310–317.

Zhang, Y., Xu, B., Kim, H.H., Muetzel, R., Delaney, S.W., & Tiemeier, H. (2022). Differences in cortical morphology and child internalizing or externalizing problems: Accounting for the co-occurrence. *Journal of Child Psychology and Psychiatry Advances*, *2*, e12114.

Zhao, B.X., Ibrahim, J.G., Li, Y., Li, T., Wang, Y., Shan, Y., … Zhu, H. (2019). Heritability of regional brain volumes in large-scale neuroimaging and genetic studies. *Cerebral Cortex*, *29*, 2904–2914.

Accepted for publication: 8 November 2023