NeuroImage: Clinical 36 (2022) 103195

Contents lists available at ScienceDirect

NeuroImage: Clinical

journal homepage: www.elsevier.com/locate/ynicl

The relationship between adolescents' externalizing and internalizing symptoms and brain development over a period of three years

Irina Jarvers [a,*], Stephanie Kandsperger [a], Daniel Schleicher [a], Ayaka Ando [b], Franz Resch [b], Julian Koenig [c], Michael Kaess [b,d], Romuald Brunner [a]

[a] Department for Child and Adolescent Psychiatry and Psychotherapy, University of Regensburg, Germany
[b] Clinic for Child and Adolescent Psychiatry, Centre for Psychosocial Medicine, University Hospital Heidelberg, Germany
[c] University of Cologne, Faculty of Medicine and University Hospital Cologne, Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, Cologne, Germany
[d] University Hospital of Child and Adolescent Psychiatry and Psychotherapy, University of Bern, Switzerland

ARTICLE INFO

Keywords:
Brain development
Psychopathology
Externalizing
Internalizing
Adolescence

ABSTRACT

Background: Adolescence is a crucial period for both brain maturation and the emergence of mental health disorders. Associations between brain morphology and internalizing/externalizing symptomatology have been identified in clinical or at-risk samples, but age-related developmental differences were rarely considered. The current study investigated the longitudinal relationship between internalizing/externalizing symptoms and brain development in the absence of psychiatric disorders during early and late adolescence.
Methods: 98 healthy adolescents within two cohorts (younger: 9 years, older: 12 years) participated in annual assessments for three years; a clinical assessment measuring their externalizing and internalizing symptoms (SDQ) and an MRI assessment measuring their brain volume and white matter microstructure, including fractional anisotropy (FA), mean diffusivity (MD) and average path length.
Results: Linear mixed effect models and cross-lagged panel models showed that larger subcortical gray matter volume predicted more externalizing symptoms in older adolescents whereas decreases of subcortical gray matter volume predicted more externalizing symptoms for younger adolescents. Additionally, longer average white matter path length predicted more externalizing symptoms for older adolescents, while decreases in cerebral white matter volume were predictive of more externalizing symptoms for younger adolescents. There were no predictive effects for internalizing symptoms, FA or MD.
Conclusions: Delays in subcortical brain maturation, in both early and late adolescence, are associated with increases in externalizing behavior which indicates a higher risk for psychopathology and warrants further investigations.

1. Introduction

Adolescence is a critical developmental period for brain maturation (Blakemore, 2012) and the emergence of mental health problems (Paus et al., 2008). In brain development, white matter (WM) has been shown to increase gradually with age, indicating growing axonal myelination (Barnea-Goraly et al., 2005). This is reflected in measures of WM integrity such as fractional anisotropy (FA), which assess the overall directionality of water diffusion and increases with age, and mean diffusivity (MD), which assess the molecular diffusion rate and decrease with age (Schmithorst and Yuan, 2010). Whereas WM increases over development, gray matter (GM) only increases during the first decade of childhood, before declining in the second (Mills et al., 2016). Sex differences are prominent in brain development with boys generally having larger brain volumes and limbic regions independent of age (Ruigrok et al., 2014). Overall, a fundamental reorganization of the brain is assumed to take place in adolescence with subcortical areas developing earlier and frontal areas developing the latest (Konrad et al., 2013).

Internalizing and externalizing symptomatology predicts the development of mental health disorders (Goodwin et al., 2004; Reef et al.,

* Corresponding author at: Clinic of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University of Regensburg, Universitätsstraße 84, Regensburg, Bavaria 93053, Germany.
E-mail address: irina.jarvers@ukr.de (I. Jarvers).

https://doi.org/10.1016/j.nicl.2022.103195
Received 12 April 2022; Received in revised form 8 September 2022; Accepted 10 September 2022
Available online 14 September 2022

2010) and has been shown to be stable over a period of 6 years (Verhulst and Van der Ende, 1992). Internalizing symptoms refer to negative behaviors that are directed towards the self and increase with age, whereas externalizing symptoms refer to negative behaviors towards the environment that usually decrease after preschool (Fanti and Henrich, 2010).

As both symptomatology and brain maturation emerge in adolescence, their relationship deserves examination. Especially in clinical samples, associations have been identified with higher levels of internalizing/externalizing symptomatology relating to smaller brain volumes and reduced cortical thickness (Kaczkurkin et al., 2020); for WM, internalizing disorders were shown to result in reduced FA and increased MD (Lichenstein et al., 2016). In healthy samples, a relationship between externalizing symptoms and a thinner PFC was implicated in children aged 12 (Ameis et al., 2014). A study with 254 children between 6 and 10 years examining internalizing/externalizing symptoms found GM volume reductions in prefrontal areas and limbic areas for internalizing symptoms (Snyder et al., 2017). Finally, Andre, Geeraert and Lebel (2019) found a relationship between reduced FA and internalizing and externalizing symptoms in children between 6 and 16 years. However, cross-sectional studies are unable to determine whether internalizing/externalizing symptoms impact brain maturation or vice versa.

Longitudinal studies are pivotal to provide insights into developmental trajectories, but findings are mixed, with increases (Dennis et al., 2019; Whittle et al., 2020) or decreases (Bos et al., 2018) in cortical metrics (i.e. volume, thickness) associated with internalizing/externalizing symptoms. Brain areas commonly found to be related to internalizing and externalizing symptoms are frontal cortical areas (Ducharme et al., 2014; Whittle et al., 2020) and limbic areas (Bos et al., 2018; Dennis et al., 2019; Muetzel et al., 2018). Ducharme et al. (2014) found a thinner prefrontal cortex (delayed GM growth) in children with more anxious and depressive symptoms and a thicker prefrontal cortex (delayed GM neuronal specification i.e. filtering of synaptic connections) in adolescents with more anxious and depressive symptoms. Whittle et al. (2020) also showed that higher cortical thickness in medial occipitofrontal cortex and the left postcentral gyrus at the age of 8 predicted externalizing symptoms two years later. Similar relationships were identified for cortical thickness in the bilateral orbitofrontal cortex and internalizing symptoms. Changes in limbic volume were observed in a study by Muetzel et al. (2018), who examined the bi-directional relationship between psychopathology, brain volume and FA in children between 6 and 8 with a 1-year follow-up. Externalizing/internalizing symptoms predicted smaller increases in subcortical GM volume and global FA. Also Bos et al. (2018) showed that decreases in hippocampal volume predicted aggression in adolescents and young adults with a mean age of 15 years. In contrast, Dennis et al. (2019) found that higher levels of irritability in 48 9- to 14-year-olds were related to GM increases in the hippocampus, the insula and frontal structures which suggested delayed maturation. In clinical samples, associations between brain matter reductions and externalizing symptoms are commonly identified at later stages in adolescence and thus when gray matter maturation (in limbic areas) has already been completed (i.e., Huebner et al., 2008). Partially, the observed differences may be due to the broad age ranges examined, across both early and late adolescence. Furthermore, some studies only examine the influence of symptoms on brain structure (Whittle et al., 2020), whereas others examine both causal directions (Muetzel et al., 2018). Additional work is required to examine the directionality of the relationship in early and late adolescence.

Previous studies with healthy samples included individuals with clinical or at-risk scores as these are, up to a certain percentage, representative of a community sample (Andre et al., 2019; Merikangas et al., 2009). However, psychiatric conditions among participants are rarely explicitly screened for and excluded. Participants can score high on a screening measure without fulfilling criteria for a diagnosis, but also score low while fulfilling them. Simpson's paradox may be an important reason to examine participants without diagnoses as a separate group in order to ensure that relationships observed are not only due to grouping (Kievit et al., 2013). Here it is important to explicitly screen participants for psychiatric disorders instead of automatically assuming the presence or absence of a disorder based on scores on a screening measure.

The purpose of the current study was to investigate the longitudinal relationship between internalizing/externalizing symptoms and brain structure in the absence of psychiatric disorders during early and late adolescence. As the relationship may be bidirectional and age specific, three analytic techniques were used, a) associations between internalizing/externalizing symptoms and brain structure across time via linear mixed effect models, b) simultaneous, bidirectional associations between internalizing/externalizing symptoms and brain structure over three time points via cross-lagged panel models and c) exploratory associations between change in brain structure (increases/reduction) and more internalizing/externalizing symptoms via bivariate correlations. As developmental trajectories indicate global reductions and increases, global brain volume and white matter topography were chosen as measures. It was hypothesized that the relationship between GM volume and internalizing and externalizing symptoms would be characterized by an interaction with age, i.e., negative associations in younger adolescents and positive associations in older adolescents. Regarding WM volume, it was hypothesized that less WM volume would be associated with more internalizing and externalizing symptoms, independent of age. For WM tracts, it was hypothesized that adolescents with more internalizing/externalizing symptoms would show reduced FA. The general relation was examined via linear mixed effect models, whereas specificities of cohort and age at the different time points were examined via cross-lagged panel models. Sex was expected to be a significant predictor of brain volume and internalizing and externalizing symptoms, as previous studies have shown large volume differences between boys and girls (Ruigrok et al., 2014) and boys show higher externalizing symptoms than girls (Mayes et al., 2020). Finally, IQ was included as a control variable as higher IQ has been associated with larger brain volume (Pietschnig et al., 2015). No specific hypotheses regarding bidirectional relationships at the separate time points were formulated.

2. Material and methods

The current study is a longitudinal three-phase investigation of the relationship between brain morphology (brain volume and WM tractography) and internalizing/externalizing symptoms in healthy children and adolescents. Data have been acquired as part of a larger study concerned with brain maturation and neuropsychological development. Children within the age range of interest were recruited via post after contacting the citizens registration office in Heidelberg. The study was approved by the Ethics Committee of the Medical Faculty at the University Heidelberg, Germany (S-604/2011). Caregivers and adolescents provided written informed consent and the procedure was in accordance with the Declaration of Helsinki (World Medical Association, 2013).

2.1. Procedure

During two appointments, clinical and cognitive assessments took place at the Department of Child and Adolescent Psychiatry, University Hospital Heidelberg and magnetic resonance imaging (MRI) sessions were conducted at the Division of Radiography at the German Cancer Research Center in Heidelberg, Germany. Both appointments were repeated each year for three years (T1, T2, T3). At the first appointment, a clinical interview, cognitive assessments and questionnaires were administered. At the second appointment, participants took part in an MRI examination. Participants received 25 € per session. See Fig. 1 for participants' age range. Data on cognitive assessments (Mürner-Lavanchy et al., 2020) and the effect of pubertal age on WM tractography (Ando et al., 2021) were published previously. The focus of the current report is the assessment of participants' internalizing/externalizing

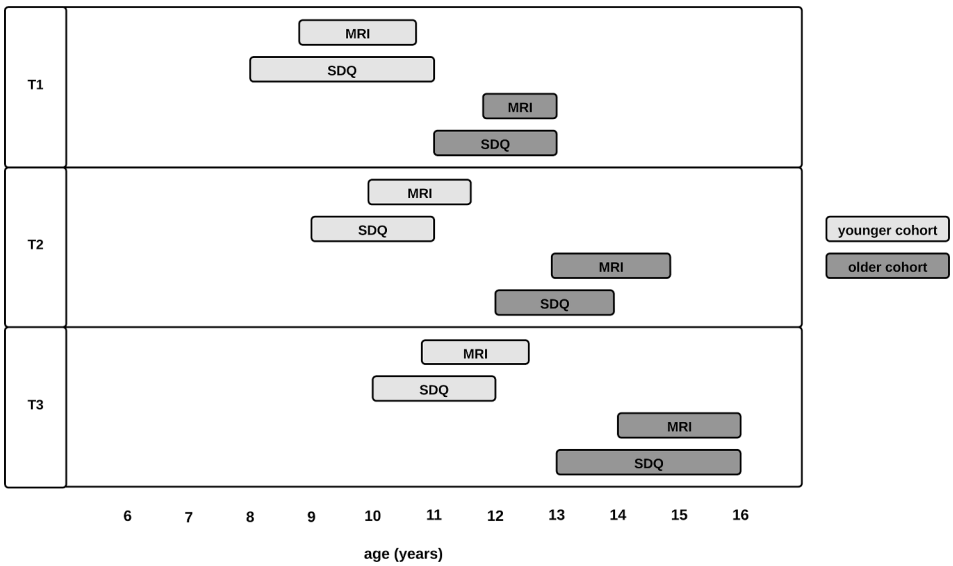

Fig. 1. Time point and measurement overview. An overview of participants' age ranges for the clinical (SDQ) and the MRI assessment across time points (T1, T2, T3). The younger cohort (9- to 10-year-olds) is depicted in light gray and the older cohort (12- to 13-year-olds) in dark gray. MRI = Magnetic resonance imaging, SDQ = Strengths and Difficulties Questionnaire.

symptoms and its relation to brain morphology.

2.2. Participants

The final sample consisted of $n = 98$ participants, split over two cohorts (cohort 1: $n = 49$, $M_{age} = 9.10$, $SD = 0.55$, range = 8–11; cohort 2: $n = 49$, $M_{age} = 12.00$, $SD = 0.41$, range = 11–13). Only participants that took part in at least two interview assessments and two MRI assessments were included in the final sample. On average the time between the interview assessment and the MRI assessment was 51.56 days ($SD = 55.83$ days) for T1, 31.71 days ($SD = 34.18$ days) for T2 and 28.43 days ($SD = 40.74$ days) for T3. The detailed exclusion criteria and recruitment process are described in the *Appendix*.

2.3. Assessment

The German version of the Mini-International Neuropsychiatric Interview for Children and Adolescents (M.I.N.I. KID) was used to rule out psychiatric disorders among participants (Sheehan et al., 1998). IQ was estimated at T1 using the General Ability Index from the German Version of the Wechsler Intelligence Scale for Children (Wechsler, 2003).

Externalizing/internalizing symptoms were assessed via a 25-item self-report, using the German version of the Strength and Difficulties Questionnaire (SDQ) for 4- to 16-year-olds (Goodman, 1997). A value between 17 and 40 for the total score is considered a clinical cut-off, whereas values between 14 and 16 are considered at-risk. An externalizing and internalizing scale was computed from the existing scales and has been shown to be more reliable in low risk samples (Goodman et al., 2010). See the *Appendix* for the detailed scale construction.

2.4. MRI acquisition and image processing

Whole brain images were acquired using a 3 T Siemens Magnetom Biograph system with a 16-channel head coil. Sequences used are elaborated on in the *Appendix*. Two types of sequences were acquired, 1) a T_1-weighted MP-RAGE sequence (Mugler and Brookeman (1990)) with repetition time (TR) = 2300 ms, echo time (TE) = 2.98 ms, field of view = 256 mm, voxel size = 1 × 1 × 1 mm, acquisition matrix = 256 and flip angle = 9°, and 2) an echo planar diffusion sequence with TR = 12100 ms, TE = 112 ms, field of view = 240 mm, acquisition matrix = 240, slice thickness = 2.5 mm and 64 gradient directions with b-values up to 3000 s/mm^2. The first sequence was used for the acquisition of brain morphology measures and the second sequence was used for the acquisition of diffusion tensor imaging (DTI) metrics.

T_1 images were reconstructed and segmented using the Freesurfer image analysis suite, version 6.0. Global metrics were subcortical GM volume (SubCortGM), cortical GM volume (CortGM) and cerebral WM volume (CerebWM). Hippocampus volume was added across hemispheres and included as it has been shown to relate to externalizing symptoms (Bos et al., 2018).

Diffusion-weighted images were preprocessed using TRActs Constrained by UnderLying Anatomy (TRACULA), a tool implemented in Freesurfer for probabilistic path reconstruction of WM pathways (Yendiki et al., 2011). An advantage of TRACULA is longitudinal tractography in which a single subject's pathways are reconstructed simultaneously for all time points (Yendiki et al., 2016). The following tracts were included in analyses: the corticospinal tract, the inferior longitudinal fasciculus, the unicinate fasciculus, the corpus callosum (forceps major, forceps minor), anterior thalamic radiation, cingulum (cingulate, angular) and the superior longitudinal fasciculus (parietal, temporal). Excluded tracts were either a) very specific (i.e., acoustic radiation) or b) described the same area via multiple tracts (i.e., 8 tracts for the corpus collosum). No particular hypotheses for specific tracts or laterality were formulated, therefore in a first step global measures were computed (by averaging over hemispheres and tracts) subsequently measures were considered by tract (by averaging over hemispheres for each tract). Specific tracts were examined if models with average values across tracts were significant. FA, MD and average white matter path length (AvgPLen) were investigated.

2.5. Statistical procedure

All statistical analyses were conducted using the R statistical package, version 4.0.2 (R Core Team, 2020) and SPSS 28 (IBM Corp, 2021). As the sample was healthy and thus showed a right-skew on the SDQ, both the externalizing score and the internalizing score were square-root transformed to achieve normal distribution. At first, bivariate correlations (Kendall's τ) between the main variables at the different time points were examined, especially relationships between main variables and covariates such as IQ, age and socio-economic status (SES). In the following analyses, three analytic techniques were used to characterize

the relationship between internalizing/externalizing symptoms and brain morphology: 1) linear mixed effect models (LMEs) were used to examine predictors of internalizing/externalizing symptoms and brain morphology (separately) across all examined time points, 2) cross-lagged panel models (CLPMs) were used to examine the identified relationships in 1) in more detail, by examining predictions for each time point while controlling for cross-lagged associations of the same variable. Finally, 3) exploratory bivariate correlations were computed between changes in brain volume (T2-T1 and T3-T1) and subsequent internalizing and externalizing symptoms. The third analytic strategy assessed the impact of volume increases/reductions over time instead of total volume alone. See Fig. 2 for an overview of the three techniques and which relationships they address. LMEs have the advantage of robustness against missing data, the inclusion of random effects (i.e., effects of subject) and considerations over several longitudinal time points. CLPMs on the other hand, are able to examine simultaneous predictive relationships in both directions and for specific time points.

2.5.1. Linear mixed effect models

LMEs for 1) were computed using the *lme4* package in R (Bates et al., 2015). All linear mixed effect models followed best practice recommendations for model fitting (Barr et al., 2013), starting with a null model including a random intercept and comparing the null model to a maximized model including all predictors, followed by a reduced model if applicable. No maximized model with a random slope was included, as the model was over-specified for the available sample size. The best fitting model was determined by several criteria, the Akaike Information Criterion (AIC; Akaike (1974)), the Bayesian Information Criterion (BIC; Schwarz (1978)) and the log likelihood ratio (LR) statistics. The model that was significantly different from the null model while having lower AIC and BIC values (>2 units lower) was chosen.

Several models were computed to assess the predictive power of brain morphology (SubCortGM, CortGM, CerebWM, FA, MD and Avg-PLen) for externalizing/internalizing symptoms (SDQ scores) and vice versa. The package *lmerTest* (Kuznetsova et al., 2017) was used in order

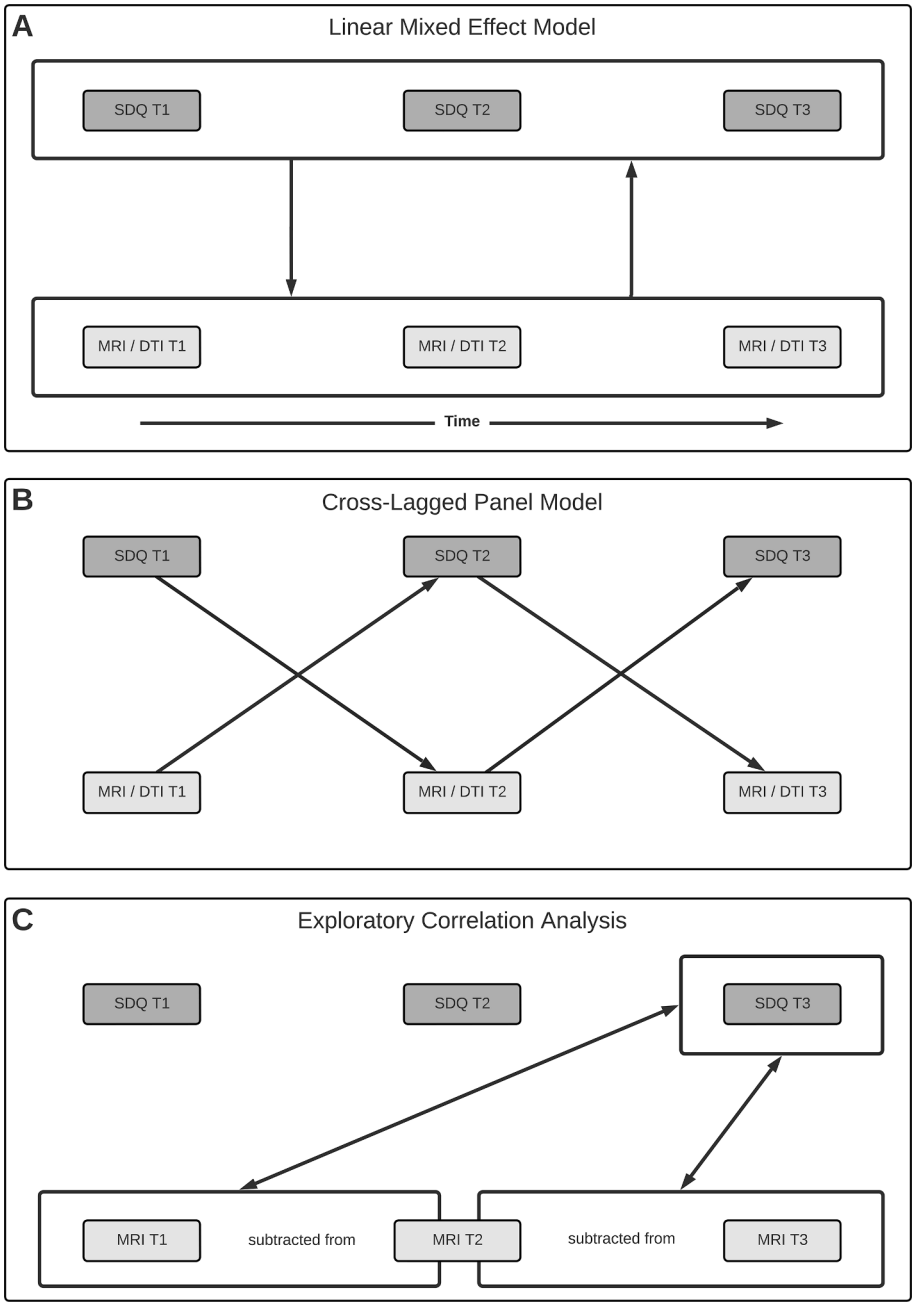

Fig. 2. Overview of analytic techniques. Three analytic techniques were used, **A.** Examining whether brain morphology predicts internalizing and externalizing symptoms (and vice versa) across all three time points, **B.** examining specific bidirectional (cross-lagged) associations between brain morphology and internalizing and externalizing symptoms from one time point to the next, and **C.** computing bivariate correlations between change in brain morphology (T2 – T1; T3 – T2) and internalizing and externalizing symptomatology at T3. MRI = Magnetic resonance imaging, DTI = Diffusion tensor imaging, SDQ = Strengths and Difficulties Questionnaire.

to compute estimated *p*-values via *t*-tests using the Satterthwaite approximations to degrees of freedom. The two basic model equations for the two directions were as follows:

$MRI_{struc} \sim Age_{MRI} + Sex + IQ + Cohort + (SDQ_{ext}\ or\ SDQ_{int}) + (SDQ_{ext} * Age_{MRI}\ or\ SDQ_{int} * Age_{MRI}) + Total\ Brain\ Volume + Sex{*}Age_{MRI} + (1|Subject)$.

$SDQ_{ext}/SDQ_{int} \sim Age_{SDQ} + Sex + IQ + Cohort + MRI_{struc} + MRI_{struc} * Age_{SDQ} + Total\ Brain\ Volume + Sex{*}Age_{MRI} + (1|Subject)$.

Where MRI_{struc} refers to both global structural MRI brain metrics and DTI metrics, Age_{MRI} and Age_{SDQ} is age at the MRI or SDQ appointment, *Cohort* is the group split for 9-year-olds and 12-year-olds, SDQ_{ext} and SDQ_{int} are the square-root-transformed externalizing and internalizing SDQ scores respectively and $(1|Subject)$ the random effect of subject. The variable of interest is the interaction between SDQ scores (SDQ_{int} and SDQ_{ext}) and age at the MRI appointment (Age_{MRI}) for equation a) and the interaction between structural MRI (MRI_{struc}) and the age at the SDQ appointment (Age_{SDQ}) for equation b). A random effect of subject was added to each model to adjust for individual differences in brain/behavior outcomes per subject. For as long as predictors did not correlate stronger than $r = 0.50$, they were included simultaneously. As different measures of brain morphology were strongly correlated, and internalizing and externalizing symptoms correlated strongly, separate models were computed for 1) internalizing symptoms, 2) externalizing symptoms and 3) every brain morphology measure as both, predictor and outcome variable. Any significant interaction effects in DTI measures were followed up with additional models involving the individual tracts as predictors. To keep models as similar as possible, age, IQ, sex, cohort, and total brain volume were included as covariates for all models. Figure A2 and A6 in the Appendix depict all computed models. The false discovery rate (FDR) was used to correct for multiple comparisons (Benjamini and Hochberg, 1995). In relation to LMEs the FDR was applied to the interaction of the independent variable with age across models testing for the same direction. This was done as only the interactions terms were relevant for confirming or rejecting the main hypotheses. Reported p-values correspond to the correction. See the Appendix for the detailed procedure of model selection.

For individual terms, the semi-partial (marginal) R^2 ($R^2_{\beta^*}$) is reported including confidence intervals (Jaeger et al., 2017). Similarly to η^2, 0.01 is considered a small effect, 0.06 a medium sized effect and 0.14 a large effect (Cohen, 1988).

2.5.2. Cross-lagged panel models

Significant interactions within LMEs were followed up with cross-lagged panel models (CLPMs) for the specific predictor variables computed through the R package *lavaan* (Rosseel, 2012). CLPMs were used as a post-hoc procedure in order to determine the directionality of the relationship between externalizing and internalizing symptoms and brain morphology. In case of significant covariates, these covariates were added as additional predictors into the model. See Fig. 3A for the basic CLPM. In order to determine model fit, the Comparative Fit Index (CFI; >0.90; Bentler, 1990), the Tucker-Lewis Index (TLI; >0.90; Tucker and Lewis, 1973), the Root Mean Square Error of Approximation (RMSEA; <0.08; Steiger, 1990) and the Standardized Root Mean Square Residual (SRMR; <0.08) were used. Here it is important to note, that the RMSEA has been shown to be less reliable with few degrees of freedom and small samples sizes (Kenny et al., 2015).

2.5.3. Exploratory bivariate correlations

As change in brain volume over time has been associated with externalizing symptoms (Bos et al., 2018), a final exploratory step included Kendall's τ correlations between a metric of change in brain volume (computed by subtracting volume at T1 from volume at T2 and T3) and externalizing symptoms as outcome variable. Also, an exploratory correlation with internalizing symptoms was computed. Correlations were computed separately for the two cohorts and controlled for total brain volume. Significance was defined as $p < .05$.

3. Results

3.1. Participants

See Table 1 for detailed participant characteristics over the three time points. Sex differences were present in brain volume and externalizing symptoms (see *Appendix*). Socio-economic statis (SES) was measured via three categories, low, middle and high. In cohort 1, 0 (0.00 %) participants had a low SES, 6 (12.50 %) had a medium SES, 42 (87.50 %) had a high SES and for 1 (2.00 %) participant SES was unknown. In cohort 2, 1 (2.00 %) participant had a low SES, 11 (22.40 %) had a medium SES and 37 (75.50 %) had a high SES. There was no significant difference between cohorts in SES at either of the three time points ($p > .05$).

3.2. Predicting brain morphology and DTI metrics

LMEs were computed to predict SubCortGM, CortGM, CerebWM, global FA, global MD and global AvgPLen from externalizing/internalizing symptoms (including age at clinical assessment, IQ, sex, cohort, total brain volume as covariates). Detailed model selection is described in the *Appendix* and depicted in Table A2. As no variables of interest were significant predictors (interaction between internalizing/externalizing symptoms and age), no reduced model was computed. Neither externalizing nor internalizing score were significant predictors for structural measures (see Table 2A and 2B). See Table A3, A4 and A5 in the *Appendix* for a complete model overview including non-significant effects.

3.3. Predicting externalizing and internalizing symptoms

LMEs were computed to predict externalizing/internalizing symptoms from brain morphology measures or DTI measures (including age, IQ, sex, total brain volume and cohort as covariates). As brain volume measures were highly correlated, separate models were computed. Model comparisons are reported in Table A6.

Among models significantly different from the null model, SubCortGM and AvgPLen were significant predictors of externalizing symptoms (see Table 3). An additional LME including the AvgPLen of all tracts separately as predictors (averaged over hemispheres), revealed only the forceps major (FMajor) and age as significant predictors, showing a negative association with externalizing symptoms (see Table 3). Fig. 3 shows the interaction between SubCortGM (A), AvgPLen (B) and the FMajor (C) in predicting externalizing symptoms.

For internalizing symptoms as outcome variable, none of the predictors resulted in better model fit. An additional analysis with hippocampus volume was not significant and is presented in the *Appendix*.

3.4. Bidirectional relationship between brain morphology and externalizing/ internalizing symptoms

LMEs with significant interaction terms were followed up with CLPMs to determine relationships between variables across different time points. Fig. 4 shows an overview, including the general model strategy and fit indices. As each cohort was a different age at the specific time points investigated, CLPM analyses were conducted separately for each cohort. IQ was added as control variable for each model.

For models examining SubCortGM and externalizing symptoms, neither externalizing symptoms at T1 predicted SubCortGM at T2, nor vice versa. However, larger SubCortGM at T2 significantly predicted more externalizing symptoms at T3 for the older cohort ($M_{ageT2} = 13.32$; see Fig. 4C), but not for the younger cohort ($M_{ageT2} = 10.29$; see Fig. 4B). This relationship was not present in the other direction, i.e., more externalizing symptoms at T2 did not predict larger SubCortGM at T3.

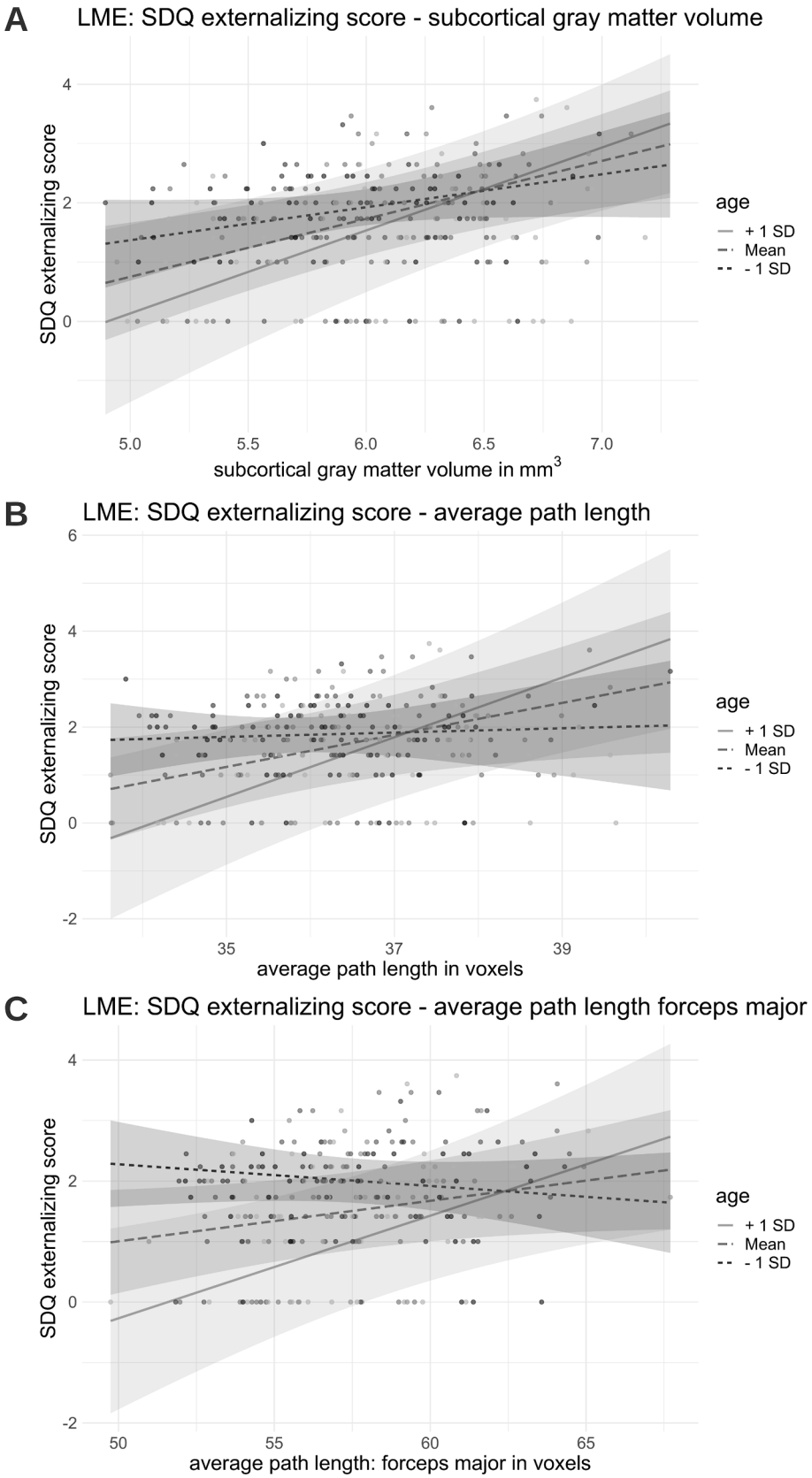

Fig. 3. Interactions between significant predictors and age in LMEs predicting SDQ externalizing scores. An overview of the interaction between **A.** subcortical GM volume and age, **B.** average path length and age, and **C.** the average path length of the FMajor and age in LMEs predicting square-root transformed externalizing SDQ scores. LME = Linear mixed effect model, SDQ = Strengths and Difficulties Questionnaire.

Table 1
Participant characteristics over the three time points.

		Cohort 1	Cohort 2				
		T1	T2	T3	T1	T2	T3
	Age	9.10 (0.55)	10.29 (0.61)	11.11 (0.61)	12.00 (0.41)	13.32 (0.56)	14.31 (0.56)
	IQ	119.86 (14.14)	–	–	118.57 (11.04)	–	–
Externalizing / internalizing symptoms	N	47	48	43	49	47	46
	SDQ Total Score	6.13 (3.24)	6.13 (3.82)	6.63 (4.77)	6.76 (4.99)	5.74 (4.63)	5.96 (4.67)
	SDQ Ext	3.49 (2.13)	3.77 (2.42)	3.81 (2.85)	4.04 (3.34)	3.36 (2.95)	3.30 (2.93)
	SDQ Int	2.64 (2.38)	2.35 (2.45)	2.81 (2.62)	2.71 (2.55)	2.38 (2.44)	2.65 (2.43)
Volumetric metrics	N	48	49	40	49	47	42
	SubCortGM	59622.63 (4388.70)	60200.31 (4448.51)	59738.47 (3939.94)	60409.94 (5463.19)	60827.06 (5150.71)	60995.38 (5542.09)
	CortGM	567525.98 (53344.51)	563943.00 (5237.46)	557143.85 (48719.02)	556109.03 (48871.35)	548585.13 (46424.04)	547531.58 (50478.40)
	CerebWM	405009.50 (45111.75)	413263.57 (45803.71)	413533.78 (44079.08)	418059.02 (50530.67)	423003.32 (47399.83)	427572.48 (52472.24)
	N	49	48	35	49	48	37
WM microstructure	FA	0.48 (0.03)	0.49 (0.02)	0.48 (0.02)	0.50 (0.03)	0.50 (0.02)	0.50 (0.02)
	MD	0.45 (0.02)	0.45 (0.01)	0.44 (0.01)	0.43 (0.02)	0.44 (0.01)	0.43 (0.01)
	AvgPLen	36.41 (1.18)	36.41 (1.30)	36.38 (1.12)	36.60 (1.59)	36.77 (1.37)	36.60 (1.40)

Note. Values correspond to means with standard deviations in brackets. Reported SDQ scores are raw values before the square-root-transformation. Volumetric metrics are reported in mm^3. FA is a unitless ratio with a range between 0 and 1. MD is reported in 10^{-3} mm^2/s. AvgPLen is reported in voxels. Ext = externalizing, Int = internalizing, SubCortGM = subcortical GM volume, CortGM = cortical GM volume, CerebWM = cerebral WM volume, FA = fractional anisotropy, MD = mean diffusivity, AvgPLen = average path length.

Table 2A
Significant fixed effects of interest within LMEs predicting structural brain morphology (SubCortGM, CortGM, CerebWM).

Outcome Variable	Predictor	Fixed Effect	Estimate	SD	*t*-value	p-value	$R^2_{\beta^*}$	$R^2_{\beta^*}$ *CI*
SubCortGM	Externalizing symptoms	Total Brain Volume	0.04	0.00	12.76	< 0.001	0.60	0.50 - 0.70
SubCortGM	Internalizing symptoms	Total Brain Volume	0.04	0.00	12.59	< 0.001	0.57	0.46 - 0.67
CortGM	Externalizing symptoms	Age	−0.48	0.19	−2.58	0.01	0.01	0.00 - 0.09
		Total Brain Volume	0.47	0.02	24.27	< 0.001	0.85	0.80 - 0.88
CortGM	Internalizing symptoms	Age	−0.70	0.16	−4.32	< 0.001	0.05	0.00 - 0.15
		Total Brain Volume	0.46	0.02	24.22	< 0.001	0.84	0.80 - 0.88
CerebWM	Externalizing symptoms	Age	0.35	0.17	2.07	0.09	0.01	0.00 - 0.07
		Total Brain Volume	0.40	0.02	21.75	< 0.001	0.81	0.76 - 0.86
CerebWM	Internalizing symptoms	Age	0.57	0.14	3.96	0.001	0.03	0.00 - 0.13
		Total Brain Volume	0.41	0.02	22.18	< 0.001	0.82	0.77 - 0.86

Table 2B
Significant fixed effects of interest within LMEs predicting structural brain morphology (FA, MD, AvgPLen).

Outcome Variable	Predictor	Fixed Effect	Estimate	SD	*t*-value	p-value	$R^2_{\beta^*}$	$R^2_{\beta^*}$ *CI*
FA	Externalizing symptoms	IQ	0.49e-4	0.23e-4	2.15	0.034	0.05	0.00 - 0.15
		Cohort	21.78e-4	4.24e-4	−5.14	< 0.001	. 09	0.02 - 0.22
FA	Internalizing symptoms	Age	6.18e-4	2.87e-4	−2.15	0.045	0.02	0.00 - 0.09
		IQ	0.46e-4	0.22e-4	2.43	0.044	0.04	0.00 - 0.14
		Cohort	20.96e-4	4.30e-4	−4.87	< 0.001	0.09	0.02 - 0.22
MD	Externalizing symptoms	Cohort	0.01e-4	0.00e-4	3.94	0.003	0.045	0.00 - 0.15
MD	Internalizing symptoms	Cohort	0.01e-4	0.00e-4	3.66	0.005	0.04	0.00 - 0.14
AvgPLen	Externalizing symptoms	IQ	0.01	0.00	−2.25	0.028	0.05	0.00 - 0.15
		Total Brain Volume	0.11	0.01	21.01	< 0.001	0.71	0.66 - 0.76
AvgPLen	Internalizing symptoms	Age	0.12	0.04	2.72	0.007	0.01	0.00 - 0.07
		Total Brain Volume	0.11	0.01	20.69	< 0.001	0.75	0.68 - 0.81

Additionally, longer AvgPLen at T2 was a significant predictor of more externalizing symptoms at T3 and not vice versa. This relationship applied to the older cohort (see Fig. 4E), but not to the younger cohort (see Fig. 4D).

Finally, as it was particularly longer AvgPLen of the FMajor serving as a significant predictor for more externalizing symptoms in previous LMEs, an additional CLPM was computed with the FMajor as outcome and predictor over the three time points. There was a significant cross-lagged relationship between the longer AvgPLen of the FMajor at T2 and more externalizing symptoms at T3 for the older cohort (see Fig. 4G), but not for the younger cohort (see Fig. 4F).

3.5. Exploratory analysis of change in brain volume in relation to externalizing symptoms

As exploratory analyses, bivariate correlations between internalizing

Table 3
Fixed effects within LMEs predicting externalizing symptoms.

Outcome Variable	Predictor	Fixed Effect	Estimate	SD	*t*-value	p-value	$R^2_{\beta^*}$	$R^2_{\beta^*}$ *CI*
Externalizing symptoms	SubCortGM	Age	−1.34	0.58	−2.29	0.024	0.05	0.00 - 0.15
		IQ	−0.01	0.01	−2.12	0.024	0.05	0.00 - 0.15
		SubCortGM	−1.83	1.05	−1.74	0.084	0.03	0.00 - 0.12
		SubCortGM*Age	0.22	0.10	2.30	0.023	0.045	0.00 - 0.145
	DTI metrics	Age	−4.69	1.41	−3.32	0.001	0.09	0.02 - 0.21
		IQ	−0.01	0.01	−2.11	0.037	0.07	0.01 - 0.18
		AvgPLen	−1.25	0.43	−2.95	0.004	0.07	0.01 - 0.19
		AvgPLen*Age	0.13	0.04	3.32	0.003	0.09	0.02 - 0.21
	AvgPLen	Age	−3.01	0.91	−3.30	0.001	0.09	0.02 - 0.21
		IQ	−0.01	0.01	−1.92	0.058	0.04	0.00 - 0.14
		Sex	−0.23	0.09	−2.61	0.010	0.07	0.01 - 0.18
		FMajor	−0.52	0.17	−3.13	0.002	0.08	0.01 - 0.20
		FMajor*Age	0.05	0.02	3.31	0.001	0.09	0.02 - 0.21

Note. For individual model terms, the semi-partial (marginal) R^2 ($R^2_{\beta^*}$) is reported including confidence intervals. Reduced models are depicted. Internalizing symptoms are not reported as no predictors were shown to be significant. SubCortGM = subcortical GM volume, CortGM = cortical GM volume, CerebWM = cerebral WM volume, FA = fractional anisotropy, MD = mean diffusivity, AvgPLen = average path length, FMajor = forceps major.

and externalizing symptoms at T3 and change in brain volume between 1) T1 and T2 and 2) T2 and T3 were examined in order to identify whether increases or decreases in brain volume over time are relevant for externalizing or internalizing symptoms. As it was hypothesized that the younger cohort shows GM increases, whereas the older cohort shows GM decreases, correlations for the two cohorts were examined separately.

For the younger cohort, there was a significant negative correlation between externalizing symptoms and change between T1 and T2 in SubCortGM (τ = -0.35, p =.002) and CerebWM (τ = -0.41, p <.001). These remained significant when controlling for total brain volume at T1. No other correlations were significant (p-value range = 0.053 - 0.907). See Table 4 for an overview of all computed correlations.

See Fig. 5 for a depiction of the correlational relationship between changes in SubCortGM and CerebWM between T1 and T2 and externalizing symptoms at T3, depicted separately for the two cohorts.

4. Discussion

The aim of this study was to investigate the longitudinal relationship between internalizing/externalizing symptoms and brain development in adolescence in the absence of a psychiatric disorder. Two analytic techniques were used to differentiate between associations across time (LMEs) and cohort specific associations at the separate time points (CLPMs). There was no predictive relationship between internalizing symptoms and brain morphology. LMEs revealed that larger SubCortGM and longer AvgPLen predicted more externalizing symptoms in older, but not in younger ages. CLPMs added to this finding by confirming that specifically in the older cohort (11–13 years) larger SubCortGM and longer AvgPLen (particularly of the FMajor) at T2 predicted more externalizing symptoms at T3 (i.e., at the time point when participants were oldest). While the two main analytic techniques examined relationships with absolute volume, additional exploratory correlations examined changes in brain volume and their relationship to externalizing and internalizing symptoms. Results showed that negative change in brain volume between T1 and T2 (SubCortGM and CerebWM) was associated with more externalizing symptoms at T3 for the younger cohort (8–11 years), whereas there were no significant correlations for the older cohort.

Previous research on the longitudinal relationship between brain morphology and internalizing and externalizing symptoms has been very mixed in relation to brain areas (frontal, limbic) and changes (increases, decreases) in volume or thickness that have been identified. The present study identified that decreases in SubCortGM and CerebWM volume are associated with more externalizing symptoms in the younger cohort, whereas larger absolute SubCortGM volume and longer AvgPLen are associated with more externalizing symptoms in the older cohort. This is in line with the developmental trajectory of gray matter volume, where increases are expected in the first decade of life, whereas decreases are expected in the second (Barnea-Goraly et al., 2005; Mills et al., 2016). In the younger cohort, decreases can therefore be considered a sign of delayed maturation, whereas in the older cohort larger volumes can be considered a sign of delayed maturation. Ducharme et al. (2014) have identified a similar age effect with children showing a thinner prefrontal cortex in relation to internalizing symptoms and adolescents showing a thicker prefrontal cortex, both suggesting delayed maturation. Whittle et al. (2020) showed a similar relationship but in the other causal direction, identifying that increases in externalizing symptoms were associated with less cortical thinning later on, once again suggesting delayed maturation. An important difference is that the present study identified the predictive effect for gray matter volume of limbic structures. Limbic structures have been implicated in many psychiatric disorders, among them depression and ADHD (Hamilton et al., 2008; Shaw and Rabin, 2009). Additionally, gray matter reductions have been observed in conduct disorder which serves as a prime example for externalizing symptomatology (Huebner et al., 2008). A commonly identified function of limbic areas is emotion regulation (Lewis and Stieben, 2004), a frequent challenge for children and adolescence with externalizing disorders (Crowell et al., 2015). Thus, delayed gray matter development of limbic areas, especially in older adolescents for whom maturation should already be completed, may result in delayed development of emotion regulation strategies. This in turn may result an increase in externalizing symptoms. Future research should examine the role of emotion regulation in regard to gray matter development and externalizing symptoms.

Not many longitudinal studies have examined brain morphology for older adolescents, two exception being Bos et al. (2018) who found reductions in hippocampal volume in relation to more aggressive symptoms in adolescents and young adults, and Dennis et al. (2019) who identified that in 9–14-years-olds higher irritability was related to gray matter increases later on. Bos et al. (2018) identified the opposite developmental direction (reductions) compared to the present study, whereas Dennis et al. (2019) identified the opposite causal direction. One of the reasons for differences in results may be that Bos et al. (2018) also included young adults and the sample was older on average while Dennis et al. (2019) did not examine the bidirectional relationship between irritability and gray matter increases. Furthermore, previous studies have not explicitly excluded psychiatric disorders among their healthy samples. Examining a healthy sample with explicitly excluded psychiatric disorders allows inferences that are not based on possible

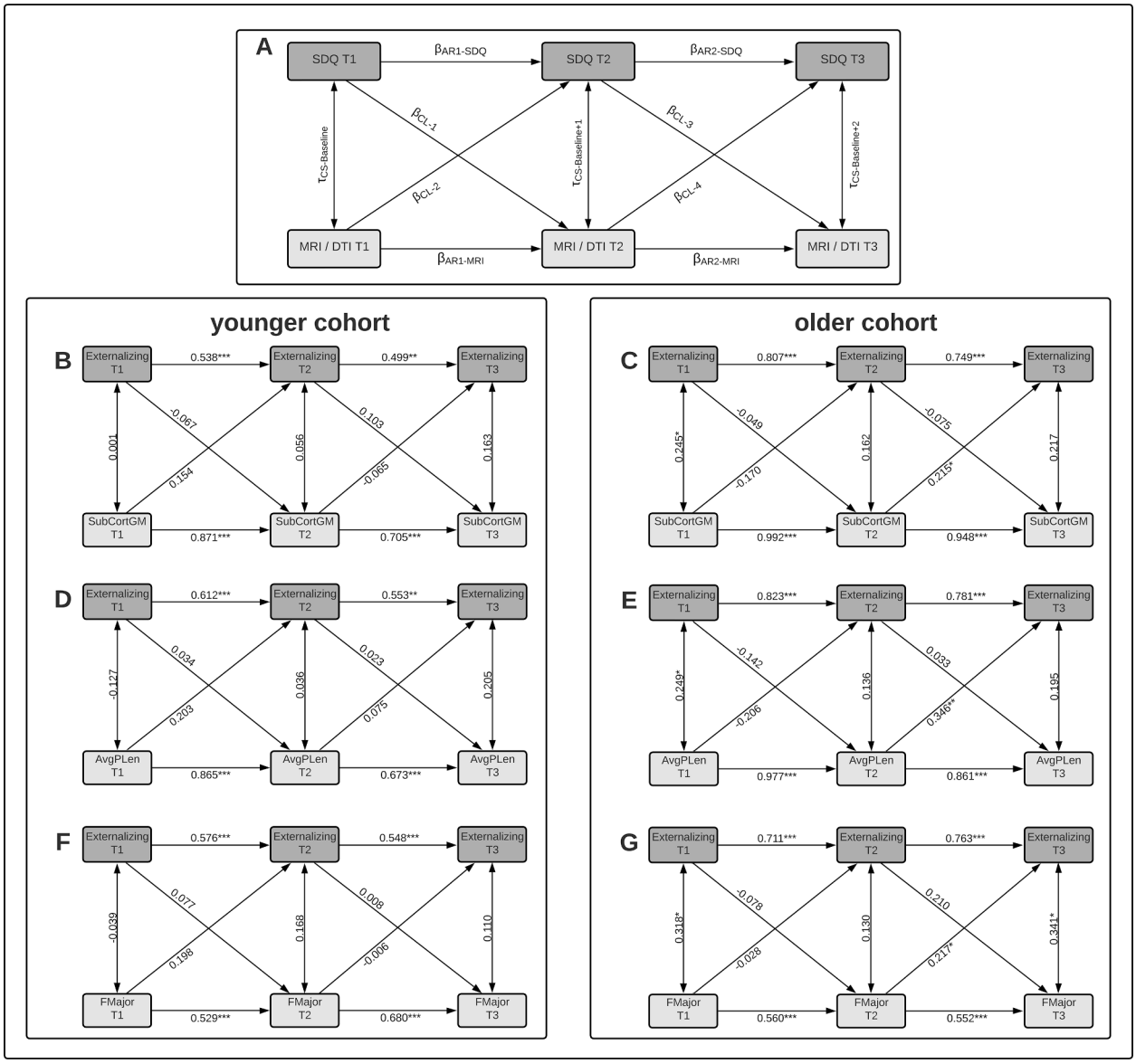

Fig. 4. Cross-lagged panel model overview. A. General model strategy used for computing the CLPMs. β refers to standardized regression coefficients and τ refers to Kendall's τ correlations. Auto-regressive coefficients are marked as 'AR' and cross-lagged coefficients are marked as 'CL'. Cross-sectional correlations are marked as 'CS'. **B.** CLPM depicting the relationship between externalizing symptoms and SubCortGM for the younger cohort, CFI = 0.887, TLI = 0.595, RMSEA = 0.291, SRMR = 0.051. **C.** CLPM depicting the relationship between externalizing symptoms and SubCortGM for the older cohort, CFI = 0.988, TLI = 0.958, RMSEA = 0.120, SRMR = 0.023. **D.** CLPM depicting the relationship between externalizing symptoms and AvgPLen for the younger cohort, CFI = 0.989, TLI = 0.960, RMSEA = 0.081, SRMR = 0.038. **E.** CLPM depicting the relationship between externalizing symptoms and AvgPLen for the older cohort, CFI = 0.972, TLI = 0.899, RMSEA = 0.165, SRMR = 0.029. **F.** CLPM depicting the relationship between externalizing symptoms and the FMajor (AvgPLen) for the younger cohort, CFI = 0.828, TLI = 0.517, RMSEA = 0.236, SRMR = 0.086. SubcortGM = subcortical GM volume, AvgPLen = average path length, FMajor = FMajor. **G**. CLPM depicting the relationship between externalizing symptoms and the FMajor (AvgPLen) for the older cohort, CFI = 0.885, TLI = 0.677, RMSEA = 0.242, SRMR = 0.087. MRI: volumetric measures; DTI: white matter topography measures.

Table 4
Correlations between change in brain volume between 1) T1 and T2 and 2) T2 and T3 and internalizing and externalizing symptoms at T3.

	Younger cohort		Older Cohort	
	Internalizing symptoms T3	Externalizing symptoms T3	Externalizing symptoms T3	Internalizing symptoms T3
Change in SubCortGM T1-T2	-0.15	-0.35**	-0.12	-0.01
Change in SubCortGM T2-T3	0.06	0.18	-0.02	-0.14
Change in CerebWM T1-T2	-0.26	-0.41***	-0.03	-0.06
Change in CerebWM T2-T3	0.10	0.11	0.02	-0.06

Note. SubCortGM = subcortical GM volume, CerebWM = cerebral WM volume. *** $p < .001$, ** $p < .01$.

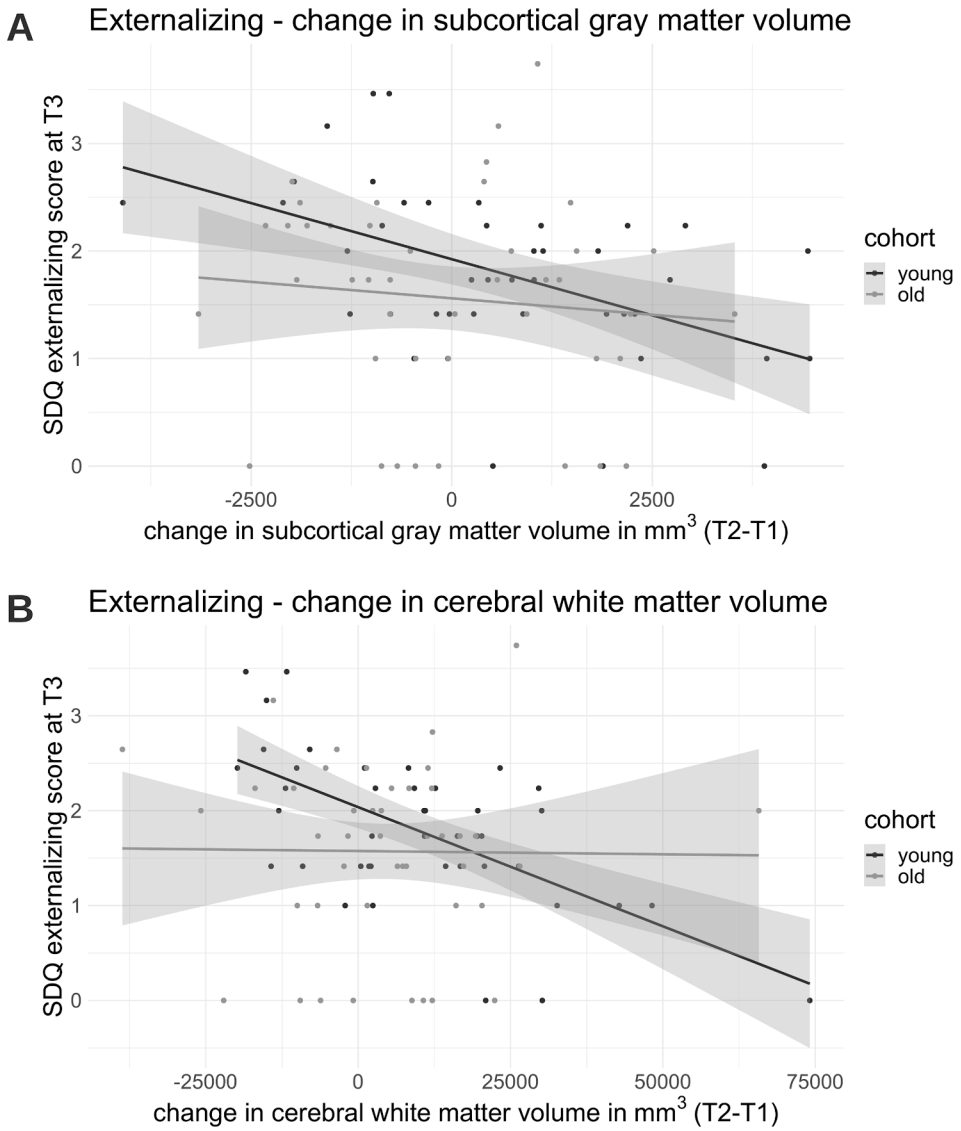

Fig. 5. Correlational relationship between SDQ externalizing scores and changes in brain volume. Overview of the correlational relationship between externalizing symptoms at T3 and **A.** change in SubCortGM from T1 to T2, split according to cohort and **B.** change in CerebWM from T1 to T2, split according to cohort. SDQ = Strengths and Difficulties Questionnaire.

morphological differences within single individuals above clinical cut-off. The present study identified associations between gray matter volume and subsequent externalizing symptoms, suggesting that increased externalizing symptoms (in a healthy range) may be a side effect of delayed maturation. Overall, more longitudinal studies about developmental trajectories of brain matter and their relation to the emergence of internalizing and externalizing symptoms are necessary to clarify relationships.

Until now, Muetzel et al. (2018) is the only longitudinal study that examined the bidirectional relationship between symptoms and brain morphology with healthy participants and found that externalizing/internalizing symptoms predicted SubCortGM decreases one year later. The present study identified the opposite causal direction with larger SubCortGM predicting more externalizing symptoms, specifically for the older cohort. This may be because Muetzel et al. (2018) investigated children below the age of 10, when GM volume increases are part of healthy development (Mills et al., 2016). In the current study, the relationship was identified for older adolescents, when GM reductions are to be expected. A reason for the current study observing the opposite causal direction compared to Muetzel et al. (2018) may be differences in recruitment periods (i.e. for Muetzel et al. externalizing symptoms and SubCortGM at T1 were examined with delays of up to 2 years). Thus, in the study by Muetzel et al. (2018), externalizing symptoms at T1 and SubCortGM at T1 may not have been assessed at the same developmental age of participants. In the current study, the predictive effect of SubCortGM for externalizing symptoms was small to medium with a broad confidence interval that ranged from a small/no effect to a large effect. As variations over the span of a year were commonly large, meaningful increases of externalizing symptoms are very well feasible. A meaningful increase can be considered a score difference that shows in adolescents' behavior, even if no at-risk or cut-off scores are reached. Overall, in younger adolescents, volume reductions for SubCortGM and CerebWM are related to more externalizing symptoms, whereas for older adolescents, it is larger absolute SubCortGM predicting subsequently more externalizing symptoms. Both associations are likely to reflect delayed brain maturation.

Apart from larger SubCortGM, also longer AvgPLen at T2 significantly predicted more externalizing symptoms at T3 for the older cohort. The predictive effect of AvgPLen was medium with a confidence interval that ranged from a small to a large effect. Overall, in the older cohort, longer AvgPLen, particularly of the FMajor, predicted more externalizing symptoms, whereas in the younger cohort, CerebWM volume

decrease was found to predict more externalizing symptoms. As white matter maturation continues into late adulthood and increases in volume are expected throughout adolescence (Mills et al., 2016), it may be surprising why the length of the FMajor is associated with more externalizing symptoms. Opposed to measures like FA, AvgPLen assesses the entire WM tract (Pannek et al., 2011) and may be influenced by differences in brain volume. In the present analysis, total brain volume was controlled for, but specific volume dependencies may nevertheless be present. The FMajor connects the occipital lobes across the corpus callosum and may be longer (contain more voxels) in individuals with greater cortex volume. It is thus possible that length of the FMajor is merely an additional indicator of delayed GM maturation. Alternatively, a longer AvgPLen may be less a sign of larger brain volumes and more a sign of delayed white matter specification. Supekar et al. (2009) have argued that as a part of white matter maturation, functional integration increases whereas average path length of networks decreases with age. Longer AvgPLen in older adolescents may therefore also indicate a delay in development that reflects in subsequent externalizing symptoms.

There was no association between internalizing/externalizing symptoms and WM microstructure measures. Muetzel et al. (2018) identified reductions in FA to be predicted by externalizing/internalizing symptoms. Similar results were obtained in clinical samples (Lichenstein et al., 2016) or cross-sectional studies with healthy participants (Andre et al., 2019), but participants were not thoroughly screened for psychiatric disorders. Muetzel et al. (2018) included participants with scores above clinical cut-off values and when removing these cases, results did not reach statistical significance. This either suggests small effect size or differences in relationships above and below diagnostic cutoffs. It is important to note that associations present across samples may change direction or disappear within samples (Kievit et al., 2013). Overall, variations of FA present in the healthy population may be minimal and too small to predict externalizing/internalizing symptoms.

Overall, the current study identified several small to medium sized effects for brain morphology influencing externalizing symptoms later on. Advantages were a longitudinal three-wave design, consistent scanner-use and a detailed exclusion of psychiatric disorders among participants. However, the study also had some limitations. First of all, the study employed an accelerated cohort design to compare two age periods (Vijayakumar et al., 2018) instead of considering the same individuals for a longer period of time. Accelerated cohort designs reduce costs, are time efficient and enable direct developmental comparisons between two age groups (Galbraith et al., 2017). In the current study, hypotheses focused on cohort differences and thus a group split according to cohort was beneficial. However, a disadvantage is that development is not regarded as a process, but rather as two stages that are examined separately. Future longitudinal studies should examine whether the same relationships can be identified in a one-cohort longitudinal study over six years. Another limitation is that only volume was investigated as a brain morphology measure in addition to WM tract measures. Common measures of brain development usually include volume, thickness and surface area. Volume and thickness have been found to relate strongly (Tamnes et al., 2017). However, surface area has been reported as the largest factor for individual differences (Tamnes et al., 2017). Future studies should investigate whether all three metrics are able to extensively discern individual differences. An additional limitation was the employment of the SDQ as the only measure of externalizing/internalizing symptoms. The SDQ shows high reliability and validity (Goodman, Lamping, & Ploubidis, 2010). Parent reports were found to be more accurate concerning hyperactivity symptoms (externalizing scale) and should therefore provide accurate estimates for the effects found in the current study. Nevertheless, the SDQ is a parent measure and may be less accurate when assessing older adolescents compared to younger adolescents (Van Roy et al., 2008). Future studies should test the relationship between internalizing/externalizing symptoms and brain morphology using both self- but also other-assessments of symptoms in order to be fully representative. Finally, despite having a sufficiently large sample size to conduct LMEs, small interactions effects could not be identified. Additionally, CLPMs' included splits according to cohort, which reduced the sample size per test substantially. Thus, it is possible that effects may not have been identified despite their presence in the population. Future work with more participants may be able to investigate the specific relationships identified and verify whether volumetric measures show no predictive relationship with internalizing symptoms in healthy adolescents.

5. Conclusion

In conclusion, the current study was able to identify meaningful predictive relationships from brain development to the development of externalizing symptoms in a sample of healthy children and adolescents. Findings suggested that even in the explicit absence of mental health disorders, larger SubCortGM and longer AvgPLen predicted more externalizing symptoms in older adolescents, whereas decreases in SubCortGM and CerebWM predicted more externalizing symptoms in younger adolescents. During early adolescence increases in GM volume are expected for healthy development, whereas in late adolescence decreases are expected for healthy development. For older adolescents, it was particularly a longer FMajor that was associated with more externalizing symptoms later on, a measure that is closely related to brain volume and may be another approximation for gray matter maturation. Overall, the current results support an association between delayed brain maturation and subsequent externalizing symptoms while emphasizing that relationships that hold for at-risk individuals and participants with mental health disorders, only partially hold for a healthy sample. The current study hereby adds to the limited research on predictive relationships between brain development and externalizing/internalizing symptoms in a sample of adolescents without psychiatric disorders.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

We would like to thank Peter Parzer, Dr. Stefan Delorme and Dr. Bram Stieltjes for their support in making this study happen. Additionally, acknowledgement goes out to the work of Susanne Schell and her engagement in participant recruitment. Finally, we are especially grateful to the children and adolescents who participated in this study. Funding: This work was supported by the Dietmar Hopp Foundation, St. Leon-Rot, Germany (Grant number: 23011197).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.nicl.2022.103195.

References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Autom. Control 19, 716–723. https://doi.org/10.1109/TAC.1974.1100705.

Ameis, S.H., Ducharme, S., Albaugh, M.D., Hudziak, J.J., Botteron, K.N., Lepage, C., Zhao, L., Khundrakpam, B., Collins, D.L., Lerch, J.P., et al., 2014. Cortical thickness, cortico-amygdalar networks, and externalizing behaviors in healthy children. Biol. Psychiatry 75, 65–72. https://doi.org/10.1016/j.biopsych.2013.06.008.

Ando, A., Parzer, P., Kaess, M., Schell, S., Henze, R., Delorme, S., Stieltjes, B., Resch, F., Brunner, R., Koenig, J., 2021. Calendar age and puberty-related development of regional gray matter volume and white matter tracts during adolescence. Brain Struct. Funct. 226, 927–937. https://doi.org/10.1007/s00429-020-02208-1.

Andre, Q.R., Geeraert, B.L., Lebel, C., 2019. Brain structure and internalizing and externalizing behavior in typically developing children and adolescents. Brain Struct. Funct. 1–10 https://doi.org/10.1007/s00429-019-01973-y.

Barnea-Goraly, N., Menon, V., Eckert, M., Tamm, L., Bammer, R., Karchemskiy, A., Dant, C.C., Reiss, A.L., 2005. White matter development during childhood and adolescence: a cross-sectional diffusion tensor imaging study. Cereb. Cortex 15, 1848–1854. https://doi.org/10.1093/cercor/bhi062.

Barr, D.J., Levy, R., Scheepers, C., Tily, H.J., 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. J. Mem. Lang. 68, 255–278. https://doi.org/10.1016/j.jml.2012.11.001.

Bates, D., Kliegl, R., Vasishth, S., Baayen, H., 2015. Parsimonious mixed models. ArXiv Prepr. ArXiv150604967.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B Methodol. 57, 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

Bentler, P.M., 1990. Comparative fit indexes in structural models. Psychol. Bull. 107, 238–246. https://doi.org/10.1037/0033-2909.107.2.238.

Blakemore, S.-J., 2012. Imaging brain development: the adolescent brain. Neuroimage 61, 397–406. https://doi.org/10.1016/j.neuroimage.2011.11.080.

Bos, M.G., Wierenga, L.M., Blankenstein, N.E., Schreuders, E., Tamnes, C.K., Crone, E.A., 2018. Longitudinal structural brain development and externalizing behavior in adolescence. J. Child Psychol. Psychiatry 59, 1061–1072. https://doi.org/10.1111/jcpp.12972.

Cohen, J., 1988. Statistical power analysis for the behavioral sciences. Lawrence Erlbaum, Mahwah, New Jersey.

Corp, I.B.M., 2021. IBM SPSS Statistics for Windows Version 28.0. IBM Corp, Armonk, NY.

Crowell, S.E., Puzia, M.E., Yaptangco, M., 2015. The ontogeny of chronic distress: emotion dysregulation across the life span and its implications for psychological and physical health. Curr. Opin. Psychol. Emotion regulation 3, 91–99. https://doi.org/10.1016/j.copsyc.2015.03.023.

Dennis, E.L., Humphreys, K.L., King, L.S., Thompson, P.M., Gotlib, I.H., 2019. Irritability and brain volume in adolescents: cross-sectional and longitudinal associations. Soc. Cogn. Affect. Neurosci. 14, 687–698. https://doi.org/10.1093/scan/nsz053.

Ducharme, S., Albaugh, M.D., Hudziak, J.J., Botteron, K.N., Nguyen, T.-V., Truong, C., Evans, A.C., Karama, S., Ball, W.S., Byars, A.W., et al., 2014. Anxious/depressed symptoms are linked to right ventromedial prefrontal cortical thickness maturation in healthy children and young adults. Cereb. Cortex 24, 2941–2950. https://doi.org/10.1093/cercor/bht151.

Fanti, K.A., Henrich, C.C., 2010. Trajectories of pure and co-occurring internalizing and externalizing problems from age 2 to age 12: findings from the National Institute of Child Health and Human Development Study of Early Child Care. Dev. Psychol. 46, 1159. https://psycnet.apa.org/doi/10.1037/a0020659.

Galbraith, S., Bowden, J., Mander, A., 2017. Accelerated longitudinal designs: An overview of modelling, power, costs and handling missing data. Stat. Methods Med. Res. 26, 374–398. https://doi.org/10.1177/0962280214547150.

Goodman, R., 1997. The Strengths and Difficulties Questionnaire: a research note. J. Child Psychol. Psychiatry 38, 581–586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x.

Goodman, A., Lamping, D.L., Ploubidis, G.B., 2010. When to use broader internalising and externalising subscales instead of the hypothesised five subscales on the Strengths and Difficulties Questionnaire (SDQ): data from British parents, teachers and children. J. Abnorm. Child Psychol. 38, 1179–1191. https://doi.org/10.1007/s10802-010-9434-x.

Goodwin, R.D., Fergusson, D.M., Horwood, L.J., 2004. Early anxious/withdrawn behaviours predict later internalising disorders. J. Child Psychol. Psychiatry 45, 874–883. https://doi.org/10.1111/j.1469-7610.2004.00279.x.

Hamilton, J.P., Siemer, M., Gotlib, I.H., 2008. Amygdala volume in major depressive disorder: a meta-analysis of magnetic resonance imaging studies. Mol. Psychiatry 13, 993–1000. https://doi.org/10.1038/mp.2008.57.

Huebner, T., Vloet, T.D., Marx, I., Konrad, K., Fink, G.R., Herpertz, S.C., Herpertz-Dahlmann, B., 2008. Morphometric brain abnormalities in boys with conduct disorder. J. Am. Acad. Child Adolesc. Psychiatry 47, 540–547. https://doi.org/10.1097/CHI.0b013e3181676545.

Jaeger, B.C., Edwards, L.J., Das, K., Sen, P.K., 2017. An R 2 statistic for fixed effects in the generalized linear mixed model. J. Appl. Stat. 44, 1086–1105. https://doi.org/10.1080/02664763.2016.1193725.

Kaczkurkin, A.N., Sotiras, A., Baller, E.B., Barzilay, R., Calkins, M.E., Chand, G.B., Cui, Z., Erus, G., Fan, Y., Gur, R.E., et al., 2020. Neurostructural heterogeneity in youths with internalizing symptoms. Biol. Psychiatry 87, 473–482. https://doi.org/10.1016/j.biopsych.2019.09.005.

Kenny, D.A., Kaniskan, B., McCoach, D.B., 2015. The Performance of RMSEA in Models With Small Degrees of Freedom. Sociol. Methods Res. 44, 486–507. https://doi.org/10.1177/0049124114543236.

Kievit, R., Frankenhuis, W., Waldorp, L., Borsboom, D., 2013. Simpson's paradox in psychological science: a practical guide. Front, Psychol, p. 4.

Konrad, K., Firk, C., Uhlhaas, P.J., 2013. Brain development during adolescence: neuroscientific insights into this developmental period. Dtsch. Ärztebl. Int. 110, 425–431. https://doi.org/10.3238/arztebl.2013.0425.

Kuznetsova, A., Brockhoff, P.B., Christensen, R.H., 2017. lmerTest package: tests in linear mixed effects models. J. Stat. Softw. 82, 1–26. https://doi.org/10.18637/jss.v082.i13.

Lewis, M.D., Stieben, J., 2004. Emotion Regulation in the Brain: Conceptual Issues and Directions for Developmental Research. Child Dev. 75, 371–376. https://doi.org/10.1111/j.1467-8624.2004.00680.x.

Lichenstein, S.D., Verstynen, T., Forbes, E.E., 2016. Adolescent brain development and depression: a case for the importance of connectivity of the anterior cingulate cortex. Neurosci. Biobehav. Rev. 70, 271–287. https://doi.org/10.1016/j.neubiorev.2016.07.024.

Mayes, S.D., Castagna, P.J., Waschbusch, D.A., 2020. Sex Differences in Externalizing and Internalizing Symptoms in ADHD, Autism, and General Population Samples. J. Psychopathol. Behav. Assess. 42, 519–526. https://doi.org/10.1007/s10862-020-09798-4.

Merikangas, K.R., Nakamura, E.F., Kessler, R.C., 2009. Epidemiology of mental disorders in children and adolescents. Dialogues Clin. Neurosci. 11, 7. https://dx.doi.org/https://dx.doi.org/10.31887/DCNS.2009.11.1/krmerikangas.

Mills, K.L., Goddings, A.-L., Herting, M.M., Meuwese, R., Blakemore, S.-J., Crone, E.A., Dahl, R.E., Güroğlu, B., Raznahan, A., Sowell, E.R., et al., 2016. Structural brain development between childhood and adulthood: Convergence across four longitudinal samples. Neuroimage 141, 273–281. https://doi.org/10.1016/j.neuroimage.2016.07.044.

Muetzel, R.L., Blanken, L.M., van der Ende, J., El Marroun, H., Shaw, P., Sudre, G., van der Lugt, A., Jaddoe, V.W., Verhulst, F.C., Tiemeier, H., et al., 2018. Tracking brain development and dimensional psychiatric symptoms in children: a longitudinal population-based neuroimaging study. Am. J. Psychiatry 175, 54–62. https://doi.org/10.1176/appi.ajp.2017.16070813.

Mugler III, J.P., Brookeman, J.R., 1990. Three-dimensional magnetization-prepared rapid gradient-echo imaging (3D MP RAGE). Magn. Reson. Med. 15, 152–157.

Mürner-Lavanchy, I.M., Koenig, J., Ando, A., Henze, R., Schell, S., Resch, F., Brunner, R., Kaess, M., 2020. Neuropsychological development in adolescents: Longitudinal associations with white matter microstructure. Dev. Cogn. Neurosci. 45, 100812 https://doi.org/10.1016/j.dcn.2020.100812.

Pannek, K., Mathias, J.L., Bigler, E.D., Brown, G., Taylor, J.D., Rose, S.E., 2011. The average pathlength map: a diffusion MRI tractography-derived index for studying brain pathology. Neuroimage 55, 133–141. https://doi.org/10.1016/j.neuroimage.2010.12.010.

Paus, T., Keshavan, M., Giedd, J.N., 2008. Why do many psychiatric disorders emerge during adolescence? Nat. Rev. Neurosci. 9, 947–957. https://doi.org/10.1038/nrn2513.

Pietschnig, J., Penke, L., Wicherts, J.M., Zeiler, M., Voracek, M., 2015. Meta-analysis of associations between human brain volume and intelligence differences: How strong are they and what do they mean? Neurosci. Biobehav. Rev. 57, 411–432. https://doi.org/10.1016/j.neubiorev.2015.09.017.

R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Reef, J., van Meurs, I., Verhulst, F.C., van der Ende, J., 2010. Children's problems predict adults' DSM-IV disorders across 24 years. J. Am. Acad. Child Adolesc. Psychiatry 49, 1117–1124. https://doi.org/10.1016/j.jaac.2010.08.002.

Rosseel, Y., 2012. Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). J. Stat. Softw. 48, 1–36. https://doi.org/10.18637/jss.v048.i02.

Ruigrok, A.N., Salimi-Khorshidi, G., Lai, M.-C., Baron-Cohen, S., Lombardo, M.V., Tait, R.J., Suckling, J., 2014. A meta-analysis of sex differences in human brain structure. Neurosci. Biobehav. Rev. 39, 34–50. https://doi.org/10.1016/j.neubiorev.2013.12.004.

Schmithorst, V.J., Yuan, W., 2010. White matter development during adolescence as shown by diffusion MRI. Brain Cogn. 72, 16–25. https://doi.org/10.1016/j.bandc.2009.06.005.

Schwarz, G., 1978. Estimating the dimension of a model. Ann. Stat. 461–464.

Shaw, P., Rabin, C., 2009. New insights into attention-deficit/hyperactivity disorder using structural neuroimaging. Curr. Psychiatry Rep. 11, 393–398. https://doi.org/10.1007/s11920-009-0059-0.

Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., Dunbar, G.C., et al., 1998. The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J. Clin. Psychiatry 59, 22–33.

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. Clin. Psychol. Sci. 5, 1001–1013. https://doi.org/10.1177/2167702617714563.

Steiger, J.H., 1990. Structural model evaluation and modification: An interval estimation approach. Multivar. Behav. Res. 25, 173–180.

Supekar, K., Musen, M., Menon, V., 2009. Development of Large-Scale Functional Brain Networks in Children. PLOS Biol. 7, e1000157.

Tamnes, C.K., Herting, M.M., Goddings, A.-L., Meuwese, R., Blakemore, S.-J., Dahl, R.E., Güroğlu, B., Raznahan, A., Sowell, E.R., Crone, E.A., et al., 2017. Development of the cerebral cortex across adolescence: a multisample study of inter-related longitudinal changes in cortical volume, surface area, and thickness. J. Neurosci. 37, 3402–3412. https://doi.org/10.1523/JNEUROSCI.3302-16.2017.

Tucker, L.R., Lewis, C., 1973. A reliability coefficient for maximum likelihood factor analysis. Psychometrika 38, 1–10.
Van Roy, B., Veenstra, M., Clench-Aas, J., 2008. Construct validity of the five-factor Strengths and Difficulties Questionnaire (SDQ) in pre-, early, and late adolescence. J. Child Psychol. Psychiatry 49, 1304–1312. https://doi.org/10.1111/j.1469-7610.2008.01942.x.
Verhulst, F.C., Van der Ende, J., 1992. Six-year developmental course of internalizing and externalizing problem behaviors. J. Am. Acad. Child Adolesc. Psychiatry 31, 924–931. https://doi.org/10.1097/00004583-199209000-00022.
Vijayakumar, N., Mills, K.L., Alexander-Bloch, A., Tamnes, C.K., Whittle, S., 2018. Structural brain development: A review of methodological approaches and best practices. Dev. Cogn. Neurosci. 33, 129–148. https://doi.org/10.1016/j.dcn.2017.11.008.
Wechsler, D., 2003. Wechsler intelligence scale for children, German version of F, Petermann. ed. Pearson Assessment.
Whittle, S., Vijayakumar, N., Simmons, J.G., Allen, N.B., 2020. Internalizing and externalizing symptoms are associated with different trajectories of cortical development during late childhood. J. Am. Acad. Child Adolesc. Psychiatry 59, 177–185. https://doi.org/10.1016/j.jaac.2019.04.006.
World Medical Association, 2013. World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. JAMA 310, 2191–2194. https://doi.org/10.1001/jama.2013.281053.
Yendiki, A., Panneck, P., Srinivasan, P., Stevens, A., Zöllei, L., Augustinack, J., Wang, R., Salat, D., Ehrlich, S., Behrens, T., Jbabdi, S.a., Gollub, R., Fischl, B., 2011. Automated probabilistic reconstruction of white-matter pathways in health and disease using an atlas of the underlying anatomy. Front. Neuroinformatics 5, 23. https://doi.org/10.3389/fninf.2011.00023.
Yendiki, A., Reuter, M., Wilkens, P., Rosas, H.D., Fischl, B., 2016. Joint reconstruction of white-matter pathways from longitudinal diffusion MRI data with anatomical priors. Neuroimage 127, 277–286. https://doi.org/10.1016/j.neuroimage.2015.12.003.