

# Educational Environment and White Matter Development in Early Adolescence

Ethan Roy<sup>1\*</sup>, Amandine Van Rinsveld<sup>1</sup>, Pierre Nedelec<sup>2</sup>, Adam Richie-Halford<sup>1,3</sup>, Andreas M. Rauschecker<sup>2</sup>, Leo P. Sugrue<sup>2</sup>, Ariel Rokem<sup>4</sup>, Bruce D. McCandliss<sup>1</sup>, Jason D. Yeatman<sup>1,3</sup>

<sup>1</sup>Graduate School of Education, Stanford University, Stanford, CA, USA.

<sup>2</sup>Department of Radiology and Biomedical Imaging, University of California San Francisco, San Francisco, CA, USA.

<sup>3</sup>Division of Developmental-Behavioral Pediatrics, Stanford University, Stanford, CA, USA.

<sup>4</sup>Department of Psychology and eScience Institute, University of Washington, Seattle, WA, USA.

\*Corresponding author

## Abstract:

Coarse measures of socioeconomic status, such as parental income or parental education, have been linked to differences in white matter development. However, these measures do not provide insight into specific aspects of an individual's environment and how they relate to brain development. On the other hand, educational intervention studies have shown that changes in an individual's educational context can drive measurable changes in their white matter. These studies, however, rarely consider socioeconomic factors in their results. In the present study, we examined the unique effect of educational opportunity on white matter development, even when controlling other known socioeconomic factors. To explore this question, we leveraged the rich demographic and neuroimaging data available in the ABCD study, as well the unique data-crosswalk between ABCD and the Stanford Education Data Archive (SEDA). We find that educational opportunity is related to accelerated white matter development, even when accounting for other socioeconomic factors, and that this relationship is most pronounced in white matter tracts associated with academic skills. These results suggest that the school a child attends has a measurable impact on brain development for years to come.

## Introduction:

Students who attend high-quality schools demonstrate higher academic achievement both in terms of reading and math scores<sup>1-4</sup>, as well as long-term outcomes such as college admissions and social mobility<sup>5-7</sup>. A common suggestion in the scientific literature<sup>8,9</sup>, and popular press<sup>10,11</sup>, is that the relationship between educational opportunity and academic outcomes reflects the influence that high-quality educational experiences might exert on brain development. Despite the well-established links between school quality and academic achievement, the specific relationship between educational opportunity and brain development remains unexplored.

Past studies, however, have demonstrated a relationship between brain development and various non-academic socioeconomic and environmental factors<sup>12</sup>. For example, diffusion MRI has revealed that higher family income predicts differences in white matter properties in adulthood<sup>13</sup> and that parental income moderates the relationship between cognitive flexibility and tissue properties across a range of white matter tracts<sup>14</sup>. Furthermore, the influence of genetic heritability on white matter structure has been shown to be higher for individuals from high income backgrounds<sup>15</sup>. Together these findings suggest that the financial resources available to an individual during childhood influence and interact with brain development in complex ways that are not fully understood.

Additionally, lower levels of parental education have been linked to differences in white matter structure across multiple fiber tracts, including those purportedly underlying academic skills, such as the left arcuate fasciculus, left superior longitudinal fasciculus (SLF) and left inferior longitudinal fasciculus (ILF)<sup>16-18</sup>. These studies offer a variety of different (and sometimes conflicting) accounts of the relationship between white matter, parental education, and cognitive skills, with some results finding a brain-behavior relationship in individuals with lower levels of parental education<sup>17</sup> and others suggesting that the relationship between parental education and cognitive behaviors is completely mediated by white matter properties<sup>16,18</sup>. What none of these studies address is why parental education affects white matter structure. In other words, are differences in parental education a proxy for a variety of environmental factors (broadly encompassed by the construct of SES) that influence white matter development? Or are there specific aspects of a child's environment that are responsible for the link between measures of SES and brain structure?

Typical measures of socioeconomic status (SES) do not elucidate the specific aspects of an individual's environment that drive differences in white matter development. The hormone cortisol has been shown to mediate the relationship between stressful life experiences and white matter properties<sup>19</sup>, suggesting that environmental stress, which has been linked to aspects of SES<sup>20</sup>, plays a role in white matter development. Furthermore, other specific environmental factors, such as home language<sup>21</sup>, screen use<sup>22</sup>, early childhood nutrition<sup>23,24</sup>, and adverse childhood events<sup>25,26</sup> have also been linked to differences in white matter properties. Together, these findings suggest that typical measures of SES, such as parental income or education, may act as proxies for a confluence of other environmental factors that directly impact white matter development.

However, across these studies examining the link between SES and white matter development, an individual's academic environment has gone unexplored. Although traditional measures of SES are highly correlated with school achievement<sup>27</sup>, it remains unclear the extent to which the specific school, and more generally the educational environment, that a child ends up in influences white matter development above and beyond the myriad of correlated factors that are wrapped up in indices of SES.

It bears mentioning that educational intervention studies have demonstrated that educational experiences can drive changes in brain structure and function over remarkably short timescales. These studies have shown that a short-term, intensive reading intervention led to changes across a range of white matter tracts, including left arcuate and left ILF, and that these changes correspond to changes in reading skill<sup>28-30</sup>. In the domain of mathematics, intensive learning experiences have been shown to “normalize” functional activity in students with mathematical learning difficulties<sup>31</sup> and participation in specific math curricula drives changes in neurotransmitter concentration in the middle frontal gyrus and predicts longitudinal changes in mathematical reasoning<sup>32</sup>.

Although these findings serve as a proof-of-concept that educational experience can shape brain development in a manner that facilitates the development of academic skills, the samples used in these studies included a small number of participants in an intervention setting and were not representative of the population at large<sup>33</sup>. Moreover, the intensive and highly controlled interventions employed by these studies are far from representative of the typical differences among American schools. Furthermore, these studies did not include measures of SES in their analyses; It is possible or even likely that interventions have variable effects on brain development and learning depending on a child’s sociodemographic background<sup>34,35</sup>.

Although careful recruitment strategies can lead to sociodemographically diverse study populations, it is nearly impossible to capture the vast range of educational experiences of students across the United States in a typical brain imaging study. Recent efforts to collect and share large-scale neuroimaging datasets<sup>36-40</sup> have now opened the door for researchers to explore the interplay between brain development, cognitive skills, and environmental and demographic factors. The ongoing ABCD study<sup>38</sup> is particularly well positioned to examine the relationship between school quality and brain development in a large and representative sample. This study is following a cohort of approximately 10,000 children from across the United States longitudinally to understand brain development throughout adolescence. In addition to neuroimaging data, the ABCD study collects rich demographic and behavioral data on each participant, including traditional measures of SES, household and neighborhood cohesion, and educational opportunity, as measured by the Stanford Education Data Archive (SEDA; see Methods for overview)<sup>27</sup>. This rich set of neuroimaging and demographic data presents the first opportunity to understand the relationship between brain development and the diversity of educational environments experienced by students across the United States.

In the current study, we test the hypothesis that differences in white matter development are related to the quality of an individual’s educational environment, while controlling for the multitude of factors indexed by traditional measures of SES. We first leverage the individual white matter tract data generated through automated-fiber quantification (AFQ)<sup>41,42</sup> to test the hypothesis that educational opportunity relates to white matter development in specific tracts underlying academic skill, such as reading and math. We find that FA in the bilateral arcuate

fasciculus, left posterior arcuate, and the corpus callosum are related to educational opportunity, even when controlling for other measures of socioeconomic status. We then train a brain-age model to test the hypothesis that educational opportunity relates to accelerated white matter development. This model suggests that an individual's educational opportunity may influence white matter development throughout the brain, though this relationship may be more pronounced in white matter tracts associated with academic skills.

## Results:

### *Exploring the relationship between educational opportunity and white matter development*

Intervention studies have shown that an individual's academic environment can drive measurable changes in brain structure and function<sup>28,29,31</sup>. However, it remains unclear the extent to which educational opportunities uniquely influence brain development above and beyond other socioeconomic factors<sup>12,16,19</sup>. Here we operationalize educational opportunity based on data from the Stanford Educational Data Archive (SEDA)<sup>27</sup>. SEDA leverages standardized test scores from schools across the United States to calculate two indices of an individual's educational environment. "Early educational opportunities" provided by parents, neighborhoods, pre-K, and early elementary school education<sup>43</sup> are operationalized by the average 3rd grade test scores for a district (SEDA intercept). Year-to-year learning opportunities afforded by an individual school is operationalized by the average change in test scores for a school between grade 3rd and 8th (SEDA slope). Both measures are in z-score units and relative to national norms. Thus, a school with a SEDA intercept and slope of zero performs at the national average in terms of third grade test scores and in terms of how quickly students grow from year to year. A school with a SEDA intercept of -1 and slope of zero performs 1 standard deviation below the national average, and students progress at the average rate meaning the discrepancy in achievement is maintained throughout schooling. Because the ABCD study begins in 4th or 5th grade, SEDA intercept is the most relevant measure of the educational opportunity that a participant has experienced up until the first ABCD measurement.

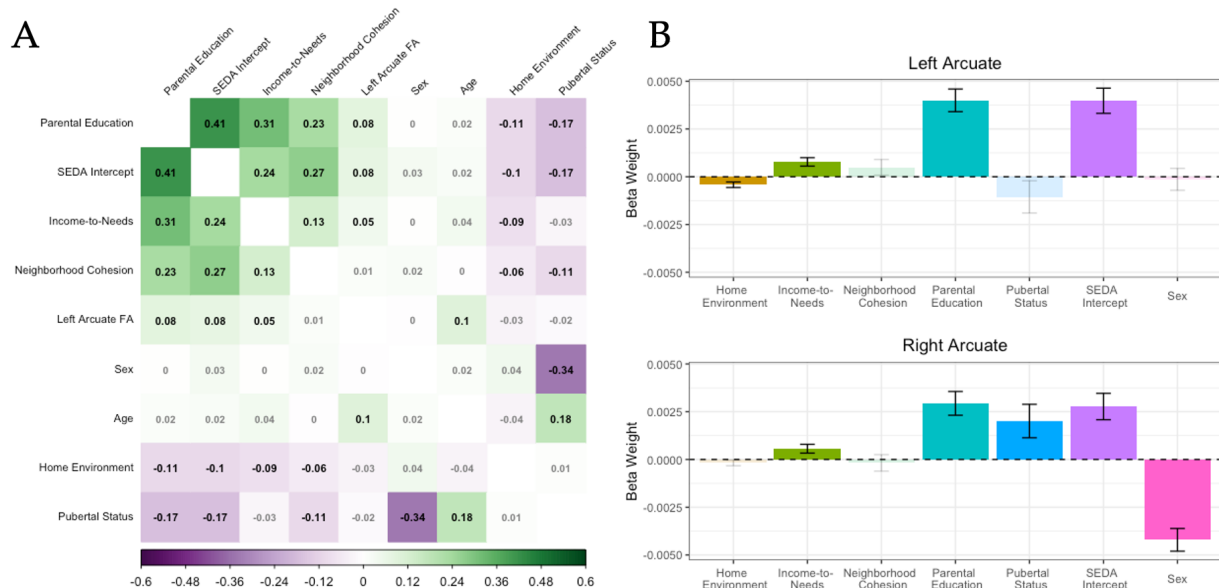
As expected, SEDA intercept is highly correlated with traditional measures of socioeconomic status (Figure 1A). The relationships between these various indices of SES raise the possibility that measures such as parental education and household income could act as proxies for other factors, like educational opportunity, that might more directly influence brain development. To this end, we leverage the rich demographic and neuroimaging data available in the ABCD dataset<sup>38</sup> to separate the unique contributions made by a student's academic environment towards their white matter development, compared to other known socioeconomic factors.

### **Diffusion Properties of the Left Arcuate are Correlated with Socioeconomic Factors**

Past studies have demonstrated a relationship between white matter development and socioeconomic factors, such as parental income, environmental stress and home context<sup>14,19,21</sup>. We began by attempting to replicate the results from past studies that observed univariate relationships between white matter and socioeconomic factors. We focus these initial analyses on the left arcuate, as past studies have shown this tract to relate to both reading skill and measures of SES<sup>18,28,44</sup>. We first calculated the correlation between mean FA in the left arcuate and a range of demographic and developmental factors, including age, pubertal status, sex, family cohesion, log-transformed income-to-needs ratio, parental education, SEDA intercept, and neighborhood cohesion (Figure 1A). This revealed small, yet significant correlations between mean FA in the left arcuate and parental education ( $r = 0.085$ ;  $p_{\text{corrected}} < 0.001$ ), income-to-needs ratio ( $r = 0.046$ ;  $p_{\text{corrected}} = 0.019$ ), SEDA intercept ( $r = 0.076$ ;  $p_{\text{corrected}} < 0.001$ ) and age ( $r = 0.096$ ;  $p_{\text{corrected}} < 0.001$ ).

However, these correlations do not account for individual and site-level random effects that might relate to white matter properties. To explore univariate relationships between FA in the left arcuate and demographic, developmental, and socioeconomic factors while controlling for these random effects, we constructed a series of linear mixed-effects models. These models included family structure nested within scanner site as random effects<sup>45</sup> and either log-transformed income-to-needs ratio, parental education, family cohesion, neighborhood cohesion, sex, SEDA intercept or pubertal status as a sole fixed-effect predictor. As a control measure, we also fit the same sequence of models predicting FA in the right arcuate to assess whether these univariate relationships occur across the white matter or are specific to the left arcuate.

Similar to the correlation analyses, these models revealed significant relationships between mean FA in the left arcuate and parental education, SEDA intercept, and income-to-needs ratio (all  $p_{\text{corrected}} < 0.001$ ; Figure 1B). Additionally, these models identified a slight yet significant relationship between home environment and mean FA in the left arcuate ( $B = -0.0004$ ,  $p_{\text{corrected}} = 0.006$ ). Interestingly, the models predicting FA in the right arcuate also found significant relationships between mean FA and parental education, income-to-needs ratio, and SEDA intercept (all  $p_{\text{corrected}} < 0.05$ ) as well as significant effects of pubertal status ( $B = 0.002$ ,  $p_{\text{corrected}} = 0.023$ ) and biological sex ( $B = -0.004$ ,  $p_{\text{corrected}} < 0.001$ ). Taken together, these models suggest that, in a univariate setting, FA and socioeconomic predictors, like income or parental education, relate to white matter properties.



**Figure 1: A.** Correlation matrix illustrating the univariate relationships between mean FA in the left arcuate, SEDA intercept, and other demographic and socioeconomic factors. Coefficients in bold represent correlations where FDR-corrected  $p < 0.05$ . **B.** Beta-weights for linear mixed-effects models predicting mean FA in the left and right arcuate from a single predictor, specified on the x-axis. Each model included a random effects structure of family structure nested within scanner site. The colors of each bar denote each predictor variable. Error bars represent the standard error of each beta-coefficient. Bars that are bolded illustrate the beta-weights with FDR-corrected  $p < 0.05$ .

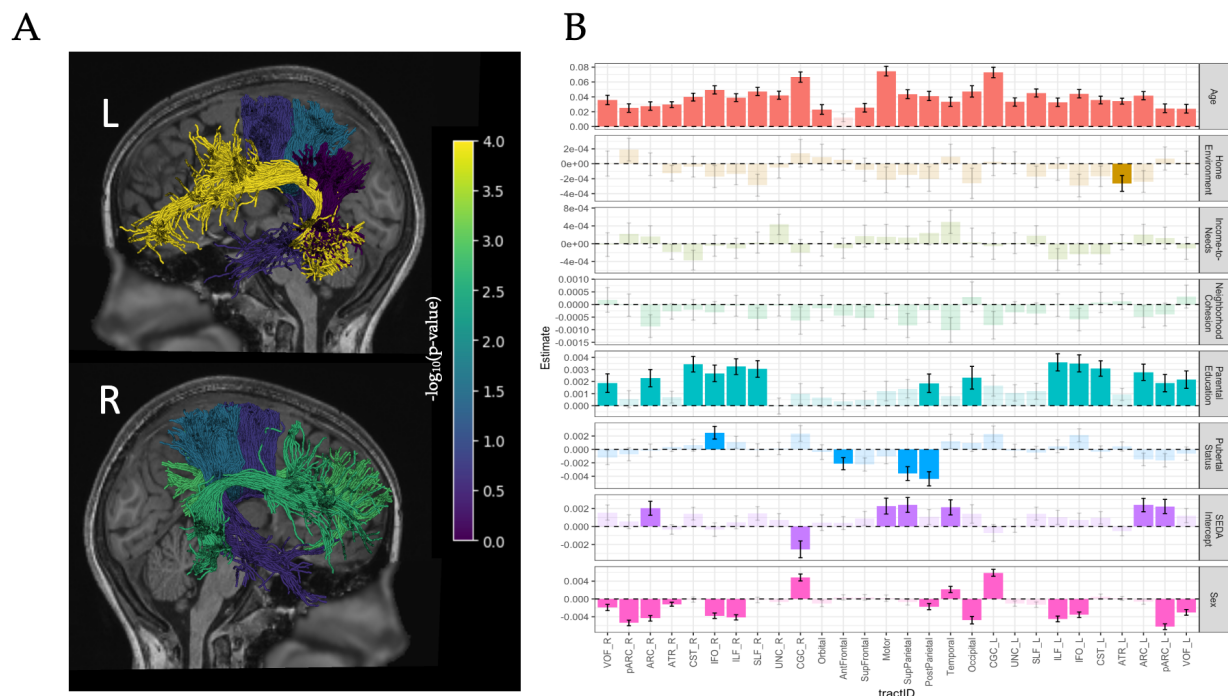
## The relationship between educational opportunity and tissue properties varies across the white matter

Although univariate analyses suggest a relationship between FA and socioeconomic factors, SES measures are also highly correlated with each other. Thus, the observed relationship between FA and any one index of SES may actually be driven by a separate, yet correlated measure. To test the hypothesis that there is a specific relationship between educational opportunity and white matter properties, while accounting for other developmental and socioeconomic effects, we modeled the relationship between SEDA intercept and tissue properties across all the white matter tracts identified by pyAFQ (see Methods of overview). Briefly, pyAFQ identifies 28 major white matter tracts and calculates diffusion properties at 100 nodes along the length of each tract (see Supplemental Figure 1 for overview of the tracts identified by pyAFQ). To account for scanner differences between the 21 ABCD study neuroimaging sites, we also performed ComBat harmonization<sup>46-48</sup> across each node included in the pyAFQ outputs.

Based on previous research in smaller samples linking specific white matter tracts to academic skills, we hypothesized that educational opportunity would specifically relate to tissue properties in the the left arcuate fasciculus (ARC), left posterior arcuate fasciculus (pARC), bilateral inferior longitudinal fasciculus (ILF), bilateral superior longitudinal fasciculus



(SLF), and uncinate fasciculus (UNC), as these tracts have all been previously implicated in academic skills, such as reading and arithmetic<sup>49–52</sup>. Within each tract, we then fit a linear mixed-effects model predicting mean harmonized fractional anisotropy (FA; averaged over the length of tract) at the first ABCD observation from SEDA intercept while also controlling for age, log-transformed income-to-needs ratio, parental education, family cohesion, neighborhood cohesion, sex, and pubertal status. Family structures nested within scanner site were included as random effects in each model<sup>45</sup>.



**Figure 2: A.** Renderings of the six white matter tracts significantly related to SEDA intercept. These include the left and right arcuate fasciculus, the left posterior arcuate, and the motor, superior parietal, and temporal bundles of the corpus callosum. Shading represents the  $-\log_{10}(p\text{-value})$  for the beta-weight on SEDA intercept from the models predicting FA in each tract (1.301 corresponds to a  $p\text{-value}$  of 0.05). This association was strongest in the left arcuate (yellow in the top panel). **B.** Beta-coefficients for the fixed effects of the models predicting FA in each major white matter tract. Each row in the figure refers to the fixed-effect in each model and each color represents a specific bundle. Error bars represent the standard error of each beta-coefficient. Bars that in bold illustrate the beta-weights with FDR-corrected  $p < 0.05$ .

Examining the beta-coefficients from these models revealed significant relationships between FA and SEDA intercept in the left and right arcuate, the left posterior arcuate, right cingulate cingulum (CGC), and three colossal tracts. The strongest SEDA relationship between SEDA intercept and FA was observed in the left arcuate (all FDR-corrected  $p < 0.05$ ; Figure 4, second row from bottom). All together, these tract-wise models demonstrate that the school a child attends influences the development of some white matter tracts above and beyond the myriad of other socioeconomic variables that characterize a child's environment.

However, as expected, FA in many tracts was also related to other environmental, developmental, and demographic factors (see Figure 2 for overview of these relationships). Parental education was linked to FA in the bilateral arcuate, CST, IFOF, ILF, right VOF, and the right SLF, whereas pubertal status was negatively related to FA in a collection of callosal bundles. Furthermore, age was positively related to higher FA across the entirety of the white matter and males, on average, demonstrated lower FA compared to females across most white matter tracts, with the exception of the left and right CGC and temporal portion of the corpus callosum. These results illustrate that, although educational environment is uniquely linked to FA in some tracts, other environmental and developmental factors are also related to white matter tissue properties across the white matter.

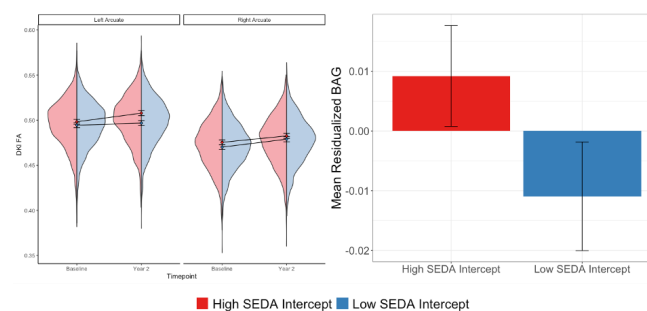
### **Development of the left and right arcuate is moderated by educational opportunity**

Based on the observed cross-sectional link between FA and SEDA intercept in both the left and right arcuate, we then focused on the longitudinal development of these two white matter tracts. Because past longitudinal studies have linked the development of left arcuate and gains in reading skill<sup>28,44,53</sup>, we chose to center this analysis around the left arcuate and its right hemisphere counterpart. To investigate the developmental dynamics of these two white matter tracts, we constructed a series of linear growth models<sup>54</sup> predicting mean FA over time (operationalized as years since initial MRI scan). In our growth models, we again included individuals nested within family structures nested within scanner sites as random effects. To control for known developmental, demographic, and socioeconomic effects, we also included initial age, pubertal status, sex, log-transformed income-to-needs ratio, parental education, family cohesion, and neighborhood cohesion, sex as fixed-effects. Our main predictors of interest in these models were time, SEDA intercept, tract (right or left arcuate), and their interactions.

We began by fitting separate growth models predicting mean FA in the left and right arcuate. These models revealed significant changes in FA in both tracts within each participant across the two observations (both  $p < 0.001$ ; See Supplemental Materials for full model outputs). These models also revealed a significant relationship between SEDA intercept and mean FA across both tracts, suggesting that, on average, individuals with greater educational opportunities have higher FA in both the left and right arcuate. We then tested the hypothesis that educational opportunity relates to interindividual differences in FA development over time. To do so, we added a SEDA intercept by time interaction to both growth models. Wald tests comparing the full and reduced models revealed that the addition of the interaction term significantly improved the fit for model predicting FA in the left arcuate ( $\chi^2(1) = 35.603$ ,  $p < 0.001$ ) but not the right arcuate ( $\chi^2(1) = 1.773$ ,  $p = 0.183$ ), suggesting that educational opportunity may have developmental effects specific to the left arcuate.



To directly test the hypothesis that educational opportunity is related to differences in FA development between the left and right arcuate, we then constructed a combined growth model based on data from both the left and right arcuate fasciculus. This model included the same fixed and random effects as the previous growth models with the addition of a three-way hemisphere by SEDA intercept by time interaction, allowing us directly compare FA development across the left and right arcuate. This model revealed that, on average, FA was lower in the right arcuate compared to the left ( $B = -0.023$ ;  $p < 0.001$ ) and that FA increased over time in both tracts ( $B = 0.002$ ;  $p < 0.001$ ). Furthermore, the average rate of FA development was slightly, yet significantly, higher in the right arcuate compared to the left ( $B = 0.0004$ ;  $p < 0.001$ ). Interestingly, individuals in environments with higher SEDA intercept scores, on average, demonstrated significantly faster rates of FA development across both the right and left arcuate ( $B = 0.001$ ;  $p < 0.001$ ) and that this interaction was significantly more pronounced in the left arcuate compared to the right ( $B = -0.0013$ ;  $p < 0.001$ ; Figure 3). For a full summary of the longitudinal growth model, see Supplementary Table 1.



**Figure 3: Left:** Growth trajectories for Diffusion Kurtosis (DKI) FA in the left and right arcuate across the first two observations of the ABCD study. The red and blue lines represent the average DKI FA growth trajectories for individuals in high (Intercept = 1) or low SEDA (Intercept = -1) intercept schools, respectively. Gray lines represent the observed changes in FA in the left and right arcuate for each individual present in the dataset. **Right:** Mean residual values for the model predicting Brain-Age Gap from a reduced model that excludes SEDA intercept as a predictor, but retains all other random and fixed-effects. Each bar represents either the top (red) or bottom (blue) 20% of participants based on their SEDA intercept scores. Error bars represent one standard error from the mean.

### *Educational Opportunity Accelerates White Matter Development*

Although the bundle-wise analyses suggest that educational opportunity is related to white matter development in the left and right arcuate, cingulate cingulum, and corpus callosum, linear-mixed effects models cannot capture these relationships while considering the entirety of the white matter. To better understand the link between educational opportunity and global white matter development, we trained a series of convolutional neural networks to predict an individual's age from the white matter properties derived using pyAFQ (see Methods for overview of training procedure). Our brain-age model trained on white matter data from both timepoints, was able to explain roughly 22% of the variance in age in the validation

set with an MAE of 0.834 years. It should be noted that the variance explained by this model is smaller than other brain age models<sup>55</sup> due to the restricted age range in the ABCD sample. Nevertheless, the residuals from this model, or the difference between the model's predicted age and each participant's observed age, can be thought of as the brain-age gap (BAG), a relative measure of how accelerated or delayed an individual's brain is maturing.

After training our brain age model, we fit a sequence of linear-mixed effects models predicting interindividual differences in BAG from a range of environmental and demographic factors. Our baseline model included age, log-transformed income-to-needs ratio, parental education, family cohesion, neighborhood cohesion, sex, and pubertal status as fixed-effects (and the same random effects structure as previous models), allowing us to control for known demographic, socioeconomic, and developmental factors related to brain development<sup>19,56</sup>.

To test the relationship of educational opportunity and white matter development, we fit a model including SEDA intercept as an additional predictor, and compared the full and reduced models using a Wald test. This tests revealed that SEDA intercept significantly improved model fit ( $\chi^2(1) = 5.483$ ,  $p = 0.02$ ). The coefficients of the SEDA intercept model revealed a significant negative relationship between BAG and age ( $B = -0.475$ ;  $p < 0.001$ ; See Supplemental Table 2 for full model output), suggesting that our model underestimates the brain-age of older participants and over estimates the brain age of younger participants. This is a known phenomenon with brain age models<sup>57</sup> and can be interpreted as regression to the mean. Additionally, the model revealed significant relationships between BAG and both parental education ( $B = 0.038$ ;  $p = 0.02$ ) and pubertal status ( $B = 0.0041$ ;  $p < 0.001$ ). The effect of parental education is in line with past findings<sup>13,16,18</sup> suggesting that parental education is linked with differences in white matter development and the effect of pubertal status suggests that puberty accelerates white matter development.

Interestingly, this model also revealed a small, yet significant relationship between the BAG and SEDA intercept ( $B = 0.019$ ;  $p = 0.019$ ; Figure 3). The observed relationship between SEDA intercept and BAG suggests that the educational opportunities afforded to a learner before 3<sup>rd</sup> grade are linked to global differences in white matter maturation. These results, coupled with the insignificant coefficients on other demographic and environmental predictors (all  $p > 0.05$ ), indicate that an individual's early educational opportunities are related to global white matter development even when accounting for factors such as income-to-need ratio, family cohesion, and neighborhood stability.

## Discussion:

In the present study, we leveraged the unique epidemiological sample from the ABCD study to explore the relationship between an individual's white matter development and the educational opportunities provided by their early childhood and elementary school environments. The scale of the ABCD study, coupled with the rich demographic measures

present in the data, provides the very first opportunity to examine how the diverse educational experiences found across the United States' educational system relate to brain development, while also accounting for other environmental factors. We leveraged data from the Stanford Education Data Archive alongside diffusion MRI data to test the hypothesis that an individual's white matter development is related to the quality of their educational environment.

These analyses showed that SEDA intercept, a measure of early educational environment, was associated with fractional anisotropy in the left and right arcuate fasciculi, left posterior arcuate, and the corpus callosum. These relationships held even when accounting for other socioeconomic factors known to relate to white matter, such as parental education or household income. When examining the longitudinal relationship between white matter properties and educational opportunity over the course of 2 years, the rate of FA development in both the right and left arcuate was associated with SEDA intercept and this relationship was slightly, yet significantly, more pronounced in the left arcuate. Furthermore, a global analysis of the white matter using a brain-age modeling approach revealed a relationship between SEDA intercept and levels of white matter maturation.

To the best of our knowledge, these results are the first to show a specific link between educational environment and white matter development in a sample of this size. As observed in the present study, measures of socioeconomic status are highly correlated with one another. Thus, some indices of SES, such as household income, may confound other measures of SES in studies of brain development. The fact that SEDA intercept predicts tissue properties of specific white matter tracts, as well as global brain-age measures, even when controlling for other developmental and environmental factors, demonstrates a specific relationship between an individual's educational environment and their white matter development.

Because SEDA intercept can be thought of as a measure of the educational opportunities available to a learner in early childhood and elementary school<sup>43,58</sup>, these results suggest that early educational experiences impact the development of white matter tracts throughout elementary school and into middle school. This parallels behavioral and educational policy research that has shown that gaps in reading and mathematics at the onset of elementary school, on average, persist throughout the course of K-12 education<sup>59-61</sup> and that early measures of academic skills serve as strong predictors of later academic success and life outcomes<sup>5,6</sup>.

However, it remains unclear whether early childhood opportunities continue to influence white matter development throughout late elementary school and into adolescence or if the year-to-year educational opportunities afforded by a school also shape white matter later in development. Future work with the full longitudinal ABCD sample will have to explore this question. Nevertheless, the findings that SEDA intercept is related to both increased rates of FA development in the left arcuate and overall brain-age indicate that differences in early educational opportunity may not only lead to differences in academic outcomes but also

influence the developmental trajectories of the white matter throughout childhood and into adolescence.

Furthermore, the observed relationship between SEDA intercept and the brain-age gap suggest that early educational opportunities may not only influence the development of the specific white matter tracts underlying academic skills, but rather relate to white matter development more broadly throughout the brain. Studies in animal models have shown that environmental enrichment leads to an increase in cellular activities related to myelination, such as the proliferation of oligodendrocyte progenitor cells and alterations of the oligodendrocyte transcriptome in a broad range of brain regions<sup>62-64</sup>. Furthermore, evidence from human neuroimaging data suggests that environmental stress and caregiving settings relate to differences in white matter properties throughout the brain<sup>65,66</sup>. Taken together, these results suggest that general enrichment of an individual's educational environment may drive global changes in white matter, whereas opportunities to meaningfully engage in specific subject areas may impact the white matter tracts subserving academic skills.

Unfortunately, because SEDA scores are derived from school-level standardized test scores, they do not provide any insight into specific aspects of each individual's learning experience within a given environment, thereby limiting the specificity of the questions we can answer using the ABCD dataset. A student's experience in the classroom can be impacted by socio-cultural equity, language use, student-teacher relationships, the curriculum adopted by the school district, and classroom organization<sup>67-70</sup>, all of which do not necessarily manifest in their school's standardized test scores. These factors may differentially impact global white matter development and tissue properties in white matter connections subserving different academic skills. Future intervention studies and additional measures of educational environment are needed to better understand these relationships. These will serve to isolate the white matter changes driven by opportunities to engage in an academic subject area, such as reading or mathematics, from those due to aspects of the educational environment that are independent of the specific academic content matter.

Furthermore, because each participant has at most two observations, our longitudinal models are limited in the types of relationships captured by difference scores. White matter properties have been shown to follow non-linear growth trajectories over the lifespan<sup>66,71</sup>, however, with only two observations, one cannot effectively model non-linear relationships. Future research using the full longitudinal ABCD sample will have to explore the developmental dynamics of the white matter over the course of adolescence and determine whether the observed relationship between FA development and SEDA is best described by a linear or non-linear trajectory.

In summary, these results suggest that the educational opportunities provided to a learner in early elementary school are related to subtle differences in white matter maturation, even when accounting for other socioeconomic factors. We observe a brain-wide link between

white matter development and educational environment and find that this relationship is strongest in the white matter tracts typically associated with academic skills. Future research is needed to inform the design of interventions and policies addressing educational inequities from a neuroscientifically-informed perspective. The current study provides the first direct evidence for the relationship between educational opportunity and brain development at scale and shed light on the complex interaction between environmental factors, brain development, and learning.

## Methods

### *Participants*

The participants in the present study come from the ABCD study, a ten-year longitudinal study that includes both neuroimaging and behavioral data collected from children aged 9-10 from 21 study sites across the United States<sup>38</sup>. The data used in the present analysis come from the baseline and 2-year follow up visits of the ABCD study and can be found in the ABCD curated annual data release 4.0 (<https://nda.nih.gov/abcd/>). The baseline observation included 6,410 individuals who had the necessary neuroimaging and demographic data and the longitudinal data included 4,770 individuals with the necessary data at both time points.

### *Covariates of Interest*

We included a range of demographic and developmental factors as covariates including participant age, log-transformed income-to needs ratio, parental education, family cohesion, neighborhood cohesion, biological sex, and pubertal status. All of these measures are readily available or calculated using the data present in the ABCD data release 4.0.

### *Income-to-Needs Ratio*

Log-transformed income-to-needs ratio was calculated using the approach outlined in Wiessman et al. (2023), which combines family income and household size data. Briefly, in the ABCD study, parents report family income on a scale of 1-10, where each interval represents an income range. The midpoint of the reported range was then calculated for each participant. This dollar amount was then divided by the poverty threshold for a household of a given size. The thresholds used in this calculation come from the 2017 report by the U.S Census Bureau<sup>72</sup>. This value was then log-transformed.

### *Parental Education*

The ABCD study records parental reports of the highest level of education they have completed. This is measured on an ordinal scale ranging from “Never attended school” to “Doctoral Degree”. In the present analysis, we converted this ordinal scale into a binary variable

indicating whether or not a participant's parent completed any sort of post-secondary education (associates degree/bachelors degree).

### *Non-Academic Environment*

Family home environment was measured using the average of the nine questions present on Family Environment Scale-Family Conflict<sup>73</sup>. A higher score on this measure indicates higher levels of conflict within an individual's household and family environment. Neighborhood cohesion was assessed by taking the average of the ten items present on the ABCD Parent PhenX Community Cohesion measure<sup>74</sup>. For this measure, a higher score indicates that the participant perceives their neighborhood and surrounding community as safer and more cohesive.

### *Pubertal Status*

Pubertal status was assessed using the PDS<sup>75</sup>, a measure designed to mimic the Tanner scale to assess the development of secondary sex characteristics during the onset of puberty. In line with past research using PDS in the ABCD sample<sup>56,76</sup>, pubertal status was calculated by taking the average of the seven PDS items present on the parental PDS survey collected at each time point.

### *Educational Opportunity*

Educational opportunity was measured using linked data from the Stanford Education Data Archive (SEDA)<sup>27</sup>. This dataset leverages standardized test scores from 3rd to 8th grade students in nearly every single school district in the United States to generate measures of the educational opportunities provided by a given school or district relative to the national average. The SEDA database breaks educational opportunity into two distinct measures: intercept and slope. SEDA intercept refers to the average standardized test score for third graders from a given school or district relative to the national average and can be thought of as an index of the pre-school and early elementary school educational opportunities provided by a school catchment area<sup>43</sup>. SEDA slope is a measure of year-to-year growth in standardized scores in students from a given school or district relative to the national average. This can be thought of as the educational opportunity provided to students in a school or district between 3rd and 8th grade.

### *Diffusion MRI Acquisition and Processing*

The neuroimaging data used in this analysis come from the baseline and Year 2 follow-up sessions collected across the 21 ABCD study sites. An overview of the data acquisition and preprocessing protocols can be found in Casey et al. (2018)<sup>38</sup> and Hagler Jr. et al. (2019)<sup>77</sup>. Briefly, multi-shell, high angular-resolution imaging scans were collected on each



participant during each scan session. These data underwent manual quality control and were then minimally preprocessed using a pipeline that included eddy-current correction, motion correction,  $B_0$  distortion correction, and gradient warp correction<sup>77</sup>.

These preprocessed diffusion images were then processed with pyAFQ<sup>41</sup>. Briefly, fiber orientation distributions were estimated in each voxel using constrained spherical deconvolution<sup>78</sup> implemented in DIPY<sup>79</sup> before probabilistic tractography was used to generate streamlines throughout the white matter. As originally described in Yeatman et al. (2012)<sup>42</sup>, 30 major white matter tracts were identified from these streamlines. Each tract was then sampled to 100 nodes. At each node, fractional anisotropy (FA), mean diffusivity (MD), radial diffusivity (RD), and axial diffusivity (AD) were calculated using the diffusion kurtosis model (DKI)<sup>80,81</sup>.

To account for potential non-biological variance in the diffusion MRI signal introduced by scanner differences across the 21 ABCD sites, ComBat harmonization<sup>46,47,82</sup> was performed on the diffusion metrics calculated by pyAFQ. Harmonization was performed using the neurocombat\_sklearn Python library<sup>47,82</sup>.

### *Brain Age Gap Analysis*

To generate age predictions for the brain-age gap analysis, we trained a ResNet model<sup>83</sup> implemented in the Python library AFQ-Insight<sup>55,84</sup> on the harmonized pyAFQ outputs from both the baseline and Year 2 follow-up scans. The data used to train and evaluate this model were split into three splits: a training set, a test set, and a validation set. To prevent peeking, longitudinal observations from the same participant were placed in the same split. The validation set contained 20% of the observations, while the remaining 80% was distributed across the training and test sets. To prevent overfitting, the model was then trained on varying proportions of the training and test sets, allowing us to determine the point at which the model performance did not improve with the addition of more training data (Supplemental Figure 2).

We found that model performance plateaued when the model was trained on 56% of the overall sample. This model attained an  $R^2$  score of 0.22 on the unseen validation set. To prevent data leakage in our brain age models, we then generated two additional train-test splits to ensure that brain age predictions for each individual were generated from a model that was trained on data that did not include that individual. We then trained two additional models on 56% of the overall sample to generate brain age predictions for the individuals used as the training set of the initial brain age model. Across these three models, the average  $R^2$  score was 0.19 on the unseen data. We then calculated the residuals of these brain age predictions for each individual, using the prediction from the model that was not trained on that individual's data. The residuals of these brain age model were then used as the outcome measure of a linear-mixed effects model to explore the relationship between educational opportunity and brain-age, while controlling for other developmental and socioeconomic factors. All

linear-mixed effects models were carried out with R version 4.2.1<sup>85</sup> using the *lme4* package (version 1.1.30)<sup>86</sup>.

### *Growth Modeling*

To investigate intraindividual change in the white matter properties of the left and right arcuate and the relationship of this change with educational opportunity, we constructed a series of growth models<sup>54</sup> specified as follows:

$$FA_{tijk} = \beta_{0ijk} + \beta_{1ijk}(\text{time in study}_{tijk}) + e_{tijk}$$

where each participant's FA at a given scan session,  $t$ , is modeled as a function of a participant specific intercept ( $\beta_{0i}$ ), a participant specific slope ( $\beta_{1i}$ ), and a residual error term ( $e_{ti}$ ). To examine interindividual differences, the participant specific coefficients were modeled as:

$$\begin{aligned}\beta_{0ijk} &= \gamma_{0ij} + \gamma_{01}(\text{SEDA Intercept}_{ijk}) + u_{0ijk} \\ \beta_{1ijk} &= \gamma_{1ij} + \gamma_{11}(\text{SEDA Intercept}_{ijk}) + u_{1ijk}\end{aligned}$$

where the  $\gamma$  coefficients on SEDA Intercept refer to, on average, how baseline FA and FA development differ with SEDA intercept and  $u_{0jk}$  and  $u_{1jk}$  refer to residual error at the individual level. These models also included initial age, log-transformed income-to-needs ratio, parental education, family home environment, neighborhood cohesion, and pubertal status as covariates.

Our final growth model, which examined FA development in the left and right arcuate simultaneously, included two additional parameters,  $\gamma_{02}(\text{Hemisphere}_{ijk})$  and  $\gamma_{12}(\text{Hemisphere}_{ijk})$ , which allowed us to compare FA development across the two tracts. In this model  $\gamma_{0ijk}$  and  $\gamma_{1ijk}$  are modeled at the family structure level as:

$$\begin{aligned}\gamma_{0ij} &= \alpha_{0i} + v_{0ij} \\ \gamma_{1ij} &= \alpha_{0i} + v_{1ij}\end{aligned}$$

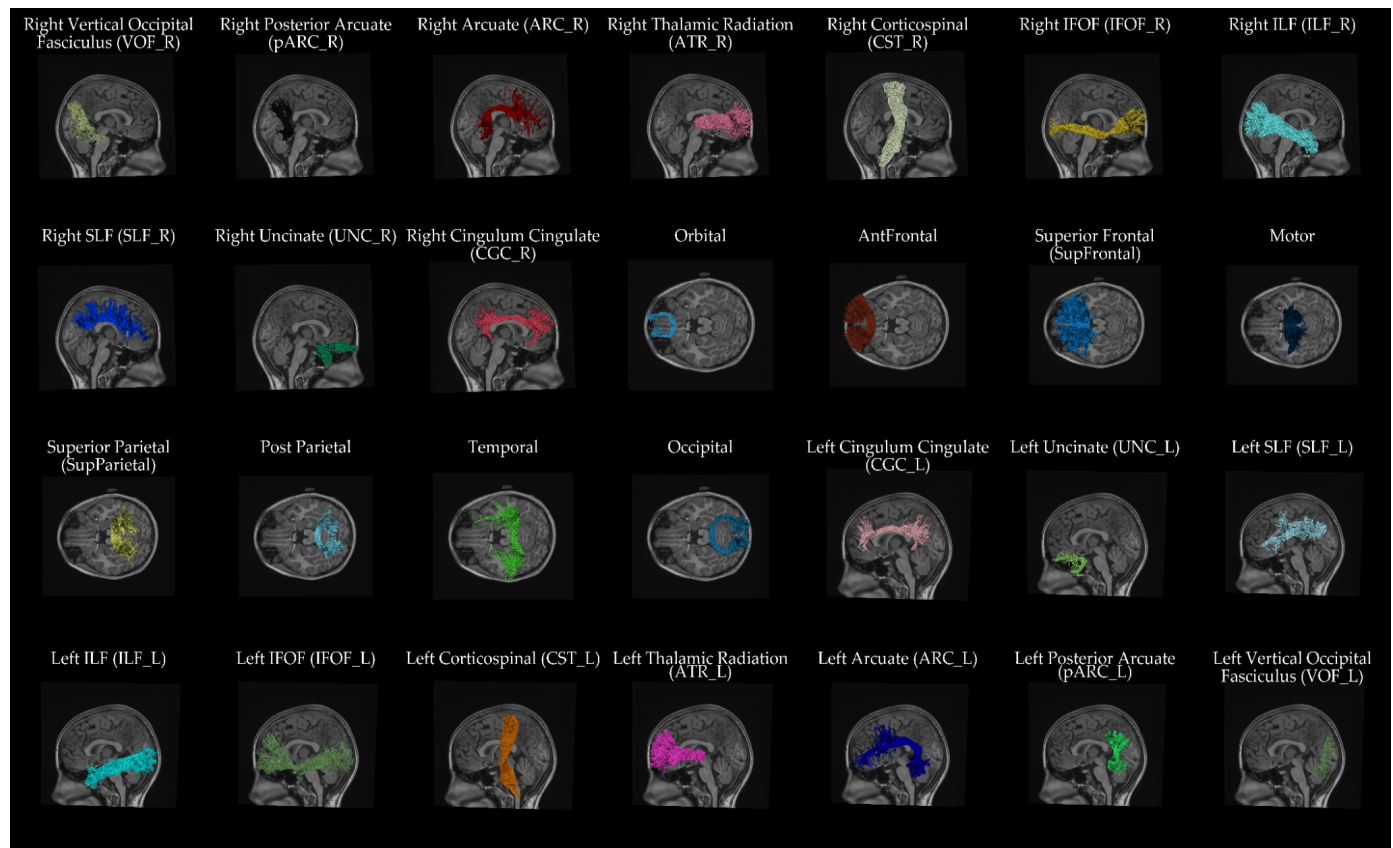
where  $v_{0ij}$  and  $v_{1ij}$  refer to residual error at the family structure level and  $\alpha_{00}$  and  $\alpha_{01}$  are modeled at the level of scanner site as:

$$\begin{aligned}\alpha_{0i} &= a_0 + w_{0i} \\ \alpha_{1i} &= a_1 + w_{1i}\end{aligned}$$

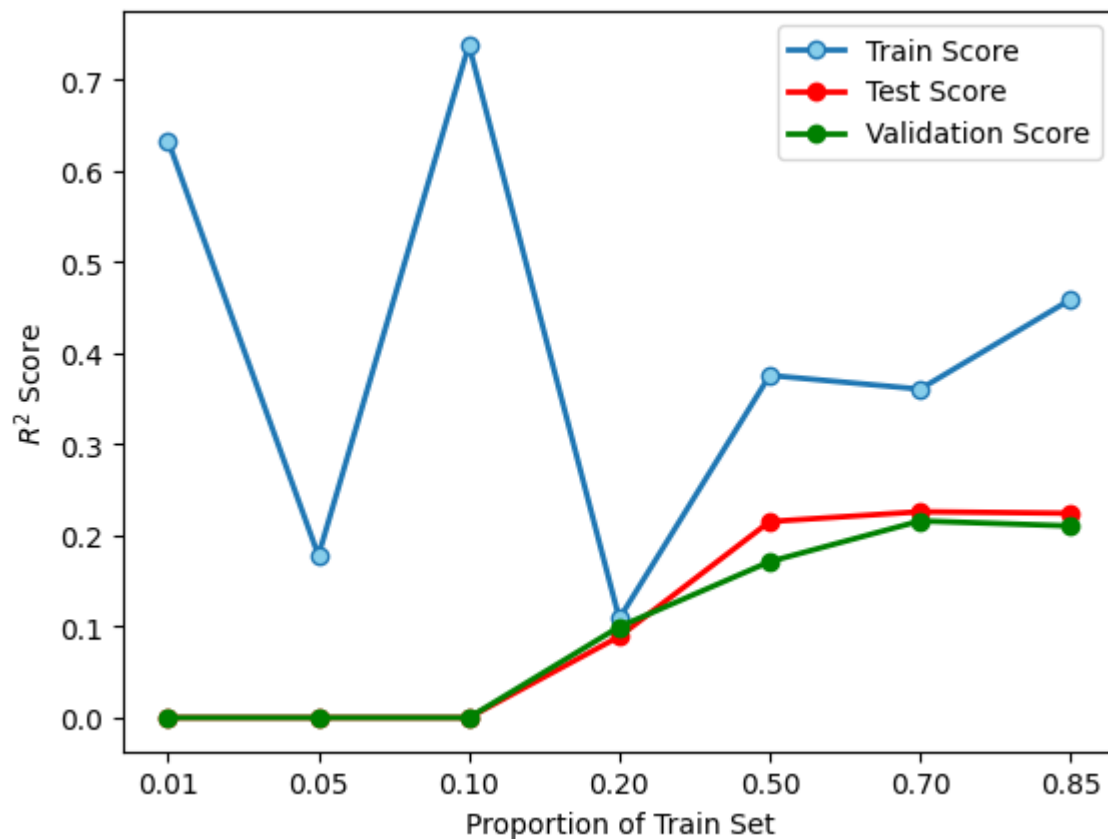
where  $w_{0i}$  and  $w_{1i}$  refer to residual error at each scanner site and  $\alpha_{0i}$  and  $\alpha_{1i}$  are the mean FA and rate of FA development, respectively, at each scanner site.

These models were fit in R version 4.2.1<sup>85</sup> using the *lme4* package (version 1.1.30)<sup>86</sup>. Code to replicate the analyses and figures presented in this manuscript can be found at: [https://github.com/earoy/white\\_matter\\_education](https://github.com/earoy/white_matter_education)

## Supplemental Figures



**Supplemental Figure 1:** Bundle renderings of the 28 white matter tracts identified by pyAFQ



**Supplemental Figure 2:** An overview of brain-age model performance trained on varying proportions of the test split. The x-axis represents the proportion of the training set used to train the brain age model and the y-axis represents the  $R^2$  score of each model.

## Supplemental Tables

**Supplemental Table 1:** Output for the growth model predicting FA development in the left arcuate, right arcuate, and left and right arcuate together. In the bilateral model, the left arcuate is used as the reference predictor.

<i>Predictors</i>	<b>Left Arcuate</b>		<b>Right Arcuate</b>		<b>Bilateral Arcuate</b>	
	<i>Estimates</i>	<i>p</i>	<i>Estimates</i>	<i>p</i>	<i>Estimates</i>	<i>p</i>
(Intercept)	0.46885	<b>&lt;0.001</b>	0.45206	<b>&lt;0.001</b>	0.47247	<b>&lt;0.001</b>
Initial Age	0.00284	<b>&lt;0.001</b>	0.00255	<b>&lt;0.001</b>	0.00267	<b>&lt;0.001</b>
Log(Income-to-Needs)	0.00007	0.757	0.00022	0.361	0.00014	0.503
Parental Ed.	0.00284	<b>&lt;0.001</b>	0.00249	<b>&lt;0.001</b>	0.00267	<b>&lt;0.001</b>
Home Env.	-0.00020	0.166	0.00005	0.759	-0.00009	0.500
Sex	-0.00041	0.471	-0.00410	<b>&lt;0.001</b>	-0.00226	<b>&lt;0.001</b>
Pubertal Status	-0.00019	0.661	-0.00003	0.943	-0.00014	0.733
Neighborhood Cohesion	-0.00042	0.325	-0.00091	<b>0.038</b>	-0.00066	0.089
SEDA	0.00232	<b>0.003</b>	0.00243	<b>0.002</b>	0.00195	<b>0.010</b>
Age	0.00198	<b>&lt;0.001</b>	0.00269	<b>&lt;0.001</b>	0.00216	<b>&lt;0.001</b>
SEDA × Age	0.00131	<b>&lt;0.001</b>	-0.00027	0.183	0.00119	<b>&lt;0.001</b>
Hemisphere					-0.02340	<b>&lt;0.001</b>
SEDA × Hemisphere					0.00084	0.094
Age × Hemisphere					0.00035	<b>0.048</b>



(SEDA× Age) × Hemisphere			-0.00134	<0.001
<b>Random Effects</b>				
$\sigma^2$	0.00	0.00	0.00	
$\tau_{00}$	0.00	0.00	0.00	
	subjectID:(rel_family_id:site_id_l)	subjectID:(rel_family_id:site_id_l)	subjectID:(rel_family_id:site_id_l)	
	0.00 rel_family_id:site_id_l	0.00 rel_family_id:site_id_l	0.00 rel_family_id:site_id_l	
	0.00 site_id_l	0.00 site_id_l	0.00 site_id_l	
ICC	0.81	0.85	0.70	
N	6129 subjectID	6129 subjectID	6130 subjectID	
	5345 rel_family_id	5345 rel_family_id	5346 rel_family_id	
	21 site_id_l	21 site_id_l	21 site_id_l	
Observations	14032	14030	28062	
Marginal R <sup>2</sup> / Conditional R <sup>2</sup>	0.022 / 0.813	0.023 / 0.849	0.180 / 0.751	

**Supplemental Table 2:** Output for the growth model predicting brain-age gap

<b>Brain-Age Gap</b>		
<i>Predictors</i>	<i>Estimates</i>	<i>p</i>
(Intercept)	-0.02398	0.105
Initial Age	-0.47519	<b>&lt;0.001</b>
Log(Income-to-Needs)	-0.00334	0.663
Parental Ed.	0.03827	<b>0.023</b>
Home Env.	-0.00007	0.991
Age	-0.64366	<b>&lt;0.001</b>
Sex	0.00871	0.571
Pubertal Status	0.04126	<b>&lt;0.001</b>
Neighborhood Cohesion	-0.01293	0.086
SEDA Intercept	0.01929	<b>0.019</b>
<b>Random Effects</b>		
$\sigma^2$	0.23	
$\tau_{00}$ subjectID:(rel_family_id:site_id_1)	0.10	
$\tau_{00}$ rel_family_id:site_id_1	0.07	
$\tau_{00}$ site_id_1	0.00	
ICC	6201	
N <sub>subjectID</sub>	5399	
N <sub>rel_family_id</sub>	21	
N <sub>site_id_1</sub>	0.23	

---

Observations	10659
Marginal $R^2$ / Conditional $R^2$	0.724 / 0.854

1. Baker, D. P., Akiba, M., LeTendre, G. K. & Wiseman, A. W. Worldwide Shadow Education: Outside-School Learning, Institutional Quality of Schooling, and Cross-National Mathematics Achievement. *Educ. Eval. Policy Anal.* **23**, 1–17 (2001).
2. Chetty, R., Hendren, N. & Katz, L. F. The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment. Working Paper at <https://doi.org/10.3386/w21156> (2015).
3. Heck, R. H. Examining the Impact of School Quality on School Outcomes and Improvement: A Value-Added Approach. *Educ. Adm. Q.* **36**, 513–552 (2000).
4. Rivkin, S. G. & Schiman, J. C. Instruction time, Classroom Quality, and Academic Achievement. *Econ. J.* **125**, F425–F448 (2015).
5. Chetty, R. *et al.* How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star \*. *Q. J. Econ.* **126**, 1593–1660 (2011).
6. Duncan, G. J. *et al.* School readiness and later achievement. *Dev. Psychol.* **43**, 1428–1446 (2007).
7. Murnane, R., Willet, J. & Levy, F. The Growing Importance of Cognitive Skills in Wage Determination. *Natl. Bur. Econ. Res.* (1995).
8. Ng, B. The neuroscience of growth mindset and intrinsic motivation. *Brain Sci.* **8**, 20 (2018).
9. Yeager, D. S. & Dweck, C. S. Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educ. Psychol.* **47**, 302–314 (2012).
10. McCandliss, B. & Toomarian, E. Putting neuroscience in the classroom: How the brain changes as we learn. *Pew Charit. Trusts Retrieved June 29, 2022* (2020).

11. Boaler, J. Ability and mathematics: The mindset revolution that is reshaping education. in (Forum, 2013).
12. Brito, N. H. & Noble, K. G. Socioeconomic status and structural brain development. *Front. Neurosci.* **8**, (2014).
13. Dufford, A. J. *et al.* Prospective associations, longitudinal patterns of childhood socioeconomic status, and white matter organization in adulthood. *Hum. Brain Mapp.* **41**, 3580–3593 (2020).
14. Ursache, A., Noble, K. G., & the Pediatric Imaging, Neurocognition and Genetics Study. Socioeconomic status, white matter, and executive function in children. *Brain Behav.* **6**, e00531 (2016).
15. Chiang, M.-C. *et al.* Genetics of white matter development: A DTI study of 705 twins and their siblings aged 12 to 29. *NeuroImage* **54**, 2308–2317 (2011).
16. Noble, K. G., Korgaonkar, M. S., Grieve, S. M. & Brickman, A. M. Higher education is an age-independent predictor of white matter integrity and cognitive control in late adolescence. *Dev. Sci.* **16**, 653–664 (2013).
17. Ozernov-Palchik, O. *et al.* The relationship between socioeconomic status and white matter microstructure in pre-reading children: A longitudinal investigation. *Hum. Brain Mapp.* **40**, 741–754 (2019).
18. Vanderauwera, J., Setten, E. R. H. van, Maurits, N. M. & Maassen, B. A. M. The interplay of socio-economic status represented by paternal educational level, white matter structure and reading. *PLOS ONE* **14**, e0215560 (2019).
19. Simon, K. R. *et al.* Socioeconomic factors, stress, hair cortisol, and white matter

- microstructure in children. *Dev. Psychobiol.* **63**, e22147 (2021).
20. Luby, J. *et al.* The Effects of Poverty on Childhood Brain Development: The Mediating Effect of Caregiving and Stressful Life Events. *JAMA Pediatr.* **167**, 1135–1142 (2013).
21. Turesky, T. K. *et al.* Home language and literacy environment and its relationship to socioeconomic status and white matter structure in infancy. *Brain Struct. Funct.* **227**, 2633–2645 (2022).
22. Hutton, J. S., Dudley, J., Horowitz-Kraus, T., DeWitt, T. & Holland, S. K. Associations Between Screen-Based Media Use and Brain White Matter Integrity in Preschool-Aged Children. *JAMA Pediatr.* **174**, e193869 (2020).
23. Isaacs, E. B. *et al.* Impact of Breast Milk on Intelligence Quotient, Brain Size, and White Matter Development. *Pediatr. Res.* **67**, 357–362 (2010).
24. Ottolini, K. M., Andescavage, N., Keller, S. & Limperopoulos, C. Nutrition and the developing brain: the road to optimizing early neurodevelopment: a systematic review. *Pediatr. Res.* **87**, 194–201 (2020).
25. Hanson, J. L. *et al.* Early Neglect Is Associated With Alterations in White Matter Integrity and Cognitive Functioning. *Child Dev.* **84**, 1566–1578 (2013).
26. Choi, J., Jeong, B., Rohan, M. L., Polcari, A. M. & Teicher, M. H. Preliminary Evidence for White Matter Tract Abnormalities in Young Adults Exposed to Parental Verbal Abuse. *Biol. Psychiatry* **65**, 227–234 (2009).
27. Reardon, S. *et al.* Stanford Education Data Archive (Version 4.0). Retrieved February 8, 2021. (2021).
28. Huber, E., Donnelly, P. M., Rokem, A. & Yeatman, J. D. Rapid and widespread white matter



- plasticity during an intensive reading intervention. *Nat. Commun.* **9**, 2260 (2018).
29. Meisler, S. L., Gabrieli, J. D. E. & Christodoulou, J. A. White Matter Microstructural Plasticity Associated with Educational Intervention in Reading Disability. 2023.08.31.553629 Preprint at <https://doi.org/10.1101/2023.08.31.553629> (2023).
30. Huber, E., Mezer, A. & Yeatman, J. D. Neurobiological underpinnings of rapid white matter plasticity during intensive reading instruction. *NeuroImage* **243**, 118453 (2021).
31. Iuculano, T. *et al.* Cognitive tutoring induces widespread neuroplasticity and remediates brain function in children with mathematical learning disabilities. *Nat. Commun.* **6**, 1–10 (2015).
32. Zacharopoulos, G., Sella, F. & Cohen Kadosh, R. The impact of a lack of mathematical education on brain development and future attainment. *Proc. Natl. Acad. Sci.* **118**, e2013155118 (2021).
33. Henrich, J., Heine, S. J. & Norenzayan, A. The weirdest people in the world? *Behav. Brain Sci.* **33**, 61–83 (2010).
34. Romeo, R. R. *et al.* Socioeconomic Status and Reading Disability: Neuroanatomy and Plasticity in Response to Intervention. *Cereb. Cortex* **28**, 2297–2312 (2018).
35. Hermida, M. J. *et al.* Cognitive neuroscience, developmental psychology, and education: Interdisciplinary development of an intervention for low socioeconomic status kindergarten children. *Trends Neurosci. Educ.* **4**, 15–25 (2015).
36. Alexander, L. M. *et al.* An open resource for transdiagnostic research in pediatric mental health and learning disorders. *Sci. Data* **4**, 170181 (2017).
37. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature*

- 562, 203–209 (2018).
38. Casey, B. J. *et al.* The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Dev. Cogn. Neurosci.* **32**, 43–54 (2018).
39. Harms, M. P. *et al.* Extending the Human Connectome Project across ages: Imaging protocols for the Lifespan Development and Aging projects. *NeuroImage* **183**, 972–984 (2018).
40. Jernigan, T. L. *et al.* The Pediatric Imaging, Neurocognition, and Genetics (PING) Data Repository. *NeuroImage* **124**, 1149–1154 (2016).
41. Kruper, J. *et al.* Evaluating the Reliability of Human Brain White Matter Tractometry. *Aperture Neuro* **1**, 10.52294/e6198273-b8e3-4b63-babb-6e6b0da10669 (2021).
42. Yeatman, J. D., Dougherty, R. F., Myall, N. J., Wandell, B. A. & Feldman, H. M. Tract Profiles of White Matter Properties: Automating Fiber-Tract Quantification. *PLOS ONE* **7**, e49790 (2012).
43. Reardon. Educational Opportunity in Early and Middle Childhood: Using Full Population Administrative Data to Study Variation by Place and Age. *RSF Russell Sage Found. J. Soc. Sci.* **5**, 40 (2019).
44. Yeatman, J. D., Dougherty, R. F., Ben-Shachar, M. & Wandell, B. A. Development of white matter and reading skills. *Proc. Natl. Acad. Sci.* **109**, E3045–E3053 (2012).
45. Saragosa-Harris, N. M. *et al.* A practical guide for researchers and reviewers using the ABCD Study and other large longitudinal datasets. *Dev. Cogn. Neurosci.* **55**, 101115 (2022).
46. Fortin, J.-P. *et al.* Harmonization of multi-site diffusion tensor imaging data. *NeuroImage* **161**, 149–170 (2017).
47. Fortin, J.-P. *et al.* Harmonization of cortical thickness measurements across scanners and

- sites. *NeuroImage* **167**, 104–120 (2018).
48. Adjusting batch effects in microarray expression data using empirical Bayes methods | Biostatistics | Oxford Academic.  
<https://academic.oup.com/biostatistics/article/8/1/118/252073>.
49. Tsang, J. M., Dougherty, R. F., Deutsch, G. K., Wandell, B. A. & Ben-Shachar, M.  
Frontoparietal white matter diffusion properties predict mental arithmetic skills in children. *Proc. Natl. Acad. Sci.* **106**, 22546–22551 (2009).
50. Wandell, B. A. & Yeatman, J. D. Biological development of reading circuits. *Curr. Opin. Neurobiol.* **23**, 261–268 (2013).
51. Polspoel, B., Vandermosten, M. & De Smedt, B. Relating individual differences in white matter pathways to children’s arithmetic fluency: a spherical deconvolution study. *Brain Struct. Funct.* **224**, 337–350 (2019).
52. Navas-Sánchez, F. J. *et al.* White matter microstructure correlates of mathematical giftedness and intelligence quotient. *Hum. Brain Mapp.* **35**, 2619–2631 (2014).
53. Van Der Auwera, S., Vandermosten, M., Wouters, J., Ghesquière, P. & Vanderauwera, J. A three-time point longitudinal investigation of the arcuate fasciculus throughout reading acquisition in children developing dyslexia. *NeuroImage* **237**, 118087 (2021).
54. Grimm, K. J., Ram, N. & Estabrook, R. *Growth modeling: Structural equation and multilevel modeling approaches*. (Guilford Publications, 2016).
55. Rokem, A., Qiao, J., Yeatman, J. D. & Richie-Halford, A. Incremental improvements in tractometry-based brain-age modeling with deep learning. *bioRxiv* 2023–03 (2023).
56. Holm, M. C. *et al.* Linking brain maturation and puberty during early adolescence using

- longitudinal brain age prediction in the ABCD cohort. *Dev. Cogn. Neurosci.* **60**, 101220 (2023).
57. Butler, E. R. *et al.* Pitfalls in brain age analyses. *Hum. Brain Mapp.* **42**, 4092–4101 (2021).
58. Drescher, J., Podolsky, A., Reardon, S. F. & Torrance, G. The Geography of Rural Educational Opportunity. *RSF Russell Sage Found. J. Soc. Sci.* **8**, 123–149 (2022).
59. García, E. & Weiss, E. Early Education Gaps by Social Class and Race Start US Children Out on Unequal Footing: A Summary of the Major Findings in "Inequalities at the Starting Gate". *Econ. Policy Inst.* (2015).
60. Reardon, S. F. The widening income achievement gap. *Educ. Leadersh.* **70**, 10–16 (2013).
61. Duff, D. M., Hendricks, A. E., Fitton, L. & Adlof, S. M. Reading and Math Achievement in Children With Dyslexia, Developmental Language Disorder, or Typical Development: Achievement Gaps Persist From Second Through Fourth Grades. *J. Learn. Disabil.* 00222194221105515 (2022) doi:10.1177/00222194221105515.
62. Forbes, T. A. & Gallo, V. All Wrapped Up: Environmental Effects on Myelination. *Trends Neurosci.* **40**, 572–587 (2017).
63. Nicholson, M. *et al.* Remodelling of myelinated axons and oligodendrocyte differentiation is stimulated by environmental enrichment in the young adult brain. *Eur. J. Neurosci.* **56**, 6099–6114 (2022).
64. Goldstein, E. Z., Pertsovskaya, V., Forbes, T. A., Dupree, J. L. & Gallo, V. Prolonged Environmental Enrichment Promotes Developmental Myelination. *Front. Cell Dev. Biol.* **9**, (2021).
65. Bick, J. *et al.* Effect of Early Institutionalization and Foster Care on Long-term White Matter Development: A Randomized Clinical Trial. *JAMA Pediatr.* **169**, 211–219 (2015).

66. Lebel, C. & Deoni, S. The development of brain white matter microstructure. *NeuroImage* **182**, 207–218 (2018).
67. Goldberg, M. J., Lloyd, D. D., Syed, G., Welch, G. W. & Curenton, S. M. A Validation Study of the *Assessing Classroom Sociocultural Equity Scale* (ACSES) in Pre-Kindergarten to Third Grade Classrooms. *Early Educ. Dev.* 1–24 (2022) doi:10.1080/10409289.2022.2146392.
68. Watts, T. W., Nguyen, T., Carr, R. C., Vernon-Feagans, L. & Blair, C. Examining the Effects of Changes in Classroom Quality on Within-Child Changes in Achievement and Behavioral Outcomes. *Child Dev.* **92**, e439–e456 (2021).
69. Limlingan, M. C., McWayne, C. M., Sanders, E. A. & López, M. L. Classroom Language Contexts as Predictors of Latinx Preschool Dual Language Learners' School Readiness. *Am. Educ. Res. J.* **57**, 339–370 (2020).
70. Hamre, B. K. *et al.* A Course on Effective Teacher-Child Interactions: Effects on Teacher Beliefs, Knowledge, and Observed Practice. *Am. Educ. Res. J.* **49**, 88–123 (2012).
71. Yeatman, J. D., Wandell, B. A. & Mezer, A. A. Lifespan maturation and degeneration of human brain white matter. *Nat. Commun.* **5**, 4932 (2014).
72. Bureau, U. C. Poverty Thresholds. *Census.gov*  
<https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-poverty-thresholds.html>.
73. Moos, R. H. & Moos, B. S. Family environment scale manual: Development, applications, research. *No Title* (1994).
74. Hamilton, C. M. *et al.* The PhenX Toolkit: Get the Most From Your Measures. *Am. J. Epidemiol.* **174**, 253–260 (2011).

75. Petersen, A. C., Crockett, L., Richards, M. & Boxer, A. A self-report measure of pubertal status: Reliability, validity, and initial norms. *J. Youth Adolesc.* **17**, 117–133 (1988).
76. Beck, D. *et al.* Puberty differentially predicts brain maturation in male and female youth: A longitudinal ABCD Study. *Dev. Cogn. Neurosci.* **61**, 101261 (2023).
77. Hagler, D. J. *et al.* Image processing and analysis methods for the Adolescent Brain Cognitive Development Study. *NeuroImage* **202**, 116091 (2019).
78. Tournier, J.-D., Calamante, F. & Connelly, A. Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. *NeuroImage* **35**, 1459–1472 (2007).
79. Garyfallidis, E. *et al.* Dipy, a library for the analysis of diffusion MRI data. *Front. Neuroinformatics* **8**, 8 (2014).
80. Jensen, J. H., Helpert, J. A., Ramani, A., Lu, H. & Kaczynski, K. Diffusional kurtosis imaging: The quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magn. Reson. Med.* **53**, 1432–1440 (2005).
81. Henriques, R. N. *et al.* Diffusional Kurtosis Imaging in the Diffusion Imaging in Python Project. *Front. Hum. Neurosci.* **15**, (2021).
82. Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118–127 (2007).
83. He, K., Zhang, X., Ren, S. & Sun, J. Deep Residual Learning for Image Recognition. Preprint at <https://doi.org/10.48550/arXiv.1512.03385> (2015).
84. Richie-Halford, A., Yeatman, J. D., Simon, N. & Rokem, A. Multidimensional analysis and detection of informative features in human brain white matter. *PLOS Comput. Biol.* **17**,



e1009136 (2021).

85. R Core Team. *R: A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, 2022).

86. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).