Early-initiated childhood reading for pleasure: associations with better cognitive performance, mental well-being and brain structure in young adolescence

Yun-Jun Sun1,2,3,*, Barbara J. Sahakian1,2,3,4,5,*, Christelle Langley1,2,3,4,5, Anyi Yang1,2,3, Yuchao Jiang1,2,3, Jujiao Kang1,2,3, Xingming Zhao1,2,3,7, Chunhe Li1,2,3, Wei Cheng1,2,3,6 and Jianfeng Feng1,2,3,6,7

1Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, Shanghai, China; 2Key Laboratory of Computational Neuroscience and Brain Inspired Intelligence (Fudan University), Ministry of Education, Shanghai, China; 3MOE Frontiers Center for Brain Science, Fudan University, Shanghai, China; 4Department of Psychiatry, University of Cambridge, Cambridge, UK; 5Behavioural and Clinical Neuroscience Institute, University of Cambridge, Cambridge, UK; 6Department of Computer Science, University of Warwick, Coventry CV4 7AL, UK and 7Zhangjiang Fudan International Innovation Center, Shanghai, China

Abstract

Background. Childhood is a crucial neurodevelopmental period. We investigated whether childhood reading for pleasure (RfP) was related to young adolescent assessments of cognition, mental health, and brain structure.

Methods. We conducted a cross-sectional and longitudinal study in a large-scale US national cohort (10 000 + young adolescents), using the well-established linear mixed model and structural equation methods for twin study, longitudinal and mediation analyses. A 2-sample Mendelian randomization (MR) analysis for potential causal inference was also performed. Important factors including socio-economic status were controlled.

Results. Early-initiated long-standing childhood RfP (early RfP) was highly positively correlated with performance on cognitive tests and significantly negatively correlated with mental health problem scores of young adolescents. These participants with higher early RfP scores exhibited moderately larger total brain cortical areas and volumes, with increased regions including the frontal, temporal, insula, supramarginal, frontal, angular and para-hippocampal; right middle-occipital, anterior-cingulate, orbital areas; and subcortical ventral-diencephalon and thalamus. These brain structures were significantly related to their cognitive and mental health scores, and displayed significant mediation effects. Early RfP was longitudinally associated with higher crystallized cognition and lower attention symptoms at follow-up. Approximately 12 h/week of youth regular RfP was cognitively optimal. We further observed a moderately significant heritability of early RfP, with considerable contribution from environments. MR analysis revealed beneficial causal associations of early RfP with adult cognitive performance and left superior temporal structure.

Conclusions. These findings, for the first time, revealed the important relationships of early RfP with subsequent brain and cognitive development and mental well-being.

Introduction

Reading for pleasure (RfP) is one of the important and enjoyable childhood learning initiatives. Basically, reading is a cognitively enriching activity of gaining language and information in written form, which sets the stage and contributes broadly to knowledge acquisition (Castles, Rastle, & Nation, 2018; Mol & Bus, 2011). Distinct from language ability that allows children to naturally develop language systems even in linguistically deficient environments (Senghas, Kita, & Ozyurek, 2004), reading is a learnt skill requiring step-by-step systematic instructions, and is generally acquired and developed through explicit practice over months/years (Seymour, Aro, & Erskine, 2003). Moreover, since children can learn well through play (Palagi, Stanyon, & Demuru, 2015), it’s important to awaken joy in their early learning, e.g., by utilizing well-illustrated picture reading materials to help them understand better. During RfP, children not only practice cognitive phonological and orthographic reading processes, but also enjoy assimilating wisdom of interest, which may help in the development of long-term reading habits. Regarding early childhood RfP, with instructions and assistance from caregivers/teachers, children can understand initial printed information, learn initial reading skills including alphabetic decoding and phonological processes, engage in interactive discussions of the educational texts and images, and strengthen the bond with caregivers while...
enjoying reading together. Hearing books read aloud also helps young children’s language skills develop (Montag, Jones, & Smith, 2015), and it’s known that written language has differences with spoken language, and even books for beginner readers contain language that is significantly different in terms of content and complexity from oral conversations (Castles et al., 2018; Montag et al., 2015). Multiple benefits of early reading activities on future advanced reading and language abilities have been reported (Lonigan, Shanahan, & Washington, 2008; Mol & Bus, 2011; Mol, Bus, & de Jong, 2009). Furthermore, earlier interventions for dyslexia have been found to have a more positive response than those delivered in later primary school years (Shaywitz, Morris, & Shaywitz, 2008).

The critical period for brain development and experience-dependent plasticity in learning is childhood, which helps foster good cognition and well-being throughout life (Black et al., 2017; Klingberg, 2014; Shonkoff et al., 2012). The benefits of brain health and the development and establishment of behaviour in childhood that promote this is key to healthy development and mental well-being into adolescence and adulthood, providing resilience in times of stress (Beddington et al., 2008). Our brains are in the key developmental stage during early childhood. Neuroimaging revealed that growth in the brain cortical surface area and grey matter volume was greatest and had reached 80% of adult size by age 2. Then they expanded gradually until approximately ages 8–12, after which they declined progressively (Gilmore, Knickmeyer, & Gao, 2018; Sowell et al., 2003). The later slowdown and decreases may be due in part to subsequent synaptic pruning and functional specialization (Casey, Tottenham, Liston, & Durston, 2005; Spear, 2018). Cognitive progression milestones parallel the sequence of early brain maturation: brain regions associated with primary functions, such as the sensory and motor systems, are the earliest to mature. Subsequently, regions linked to basic language and spatial attention functions mature, including the temporal and parietal association regions, with the developmental curve peaking at 9 months after birth and lasting until late childhood. The prefrontal and lateral temporal cortices that function in modulating language and attention mature last, peaking at 2 years and lasting until approximately ages 14–16 (Casey et al., 2005).

Neuroscience findings indicate that cortical area development entails a large array of signals: there are substantial interactions between intrinsic mechanisms (to cortical progenitors and neurons) and extrinsic mechanisms (to the cortex via neural activity) (Sur & Rubenstein, 2005). Sensory experiences, in addition to internally spontaneous neural activity, are critical to the development and plasticity of neural circuit formation (Sur & Rubenstein, 2005; Zheng et al., 2014). Notably, brain plasticity peaks during the early-life childhood period throughout its lifelong dynamical regulations (Reh et al., 2020), and studies have indicated the importance of the early stage in learning, such as language (Kuhl, 2010) and mathematics learning (Clements & Sarama, 2011). Comparable to the childhood sensitive period for learning and extensive brain maturation (Shonkoff et al., 2012), adolescence is also a continued period of structural and functional brain development, particularly in the brain systems related to social–emotional, cognitive and motivational processes, with mixed influences from pubertal maturation (Crone & Dahl, 2012). Based on current existing data, our study focused on the childhood stage and the beginning of adolescence.

Compatible with the rapid rate of brain development and learning during the early years, evidence has shown the benefits of early-stage learning or intervention initiatives. The US Child Parent Center education programme has reported long-lasting benefits for children who began preschool education earlier at 3–4 years old compared with those who began kindergarten at older ages (Reynolds, Temple, Ou, Arteaga, & White, 2011): The former preschool participants have higher educational levels and socioeconomic status (SES) and lower rates of substance abuse in adulthood (Melhuish, 2011; Reynolds et al., 2011). European research also reports the long-term beneficial effects of high-quality preschool and home learning environments for children, that are associated with literacy and numeracy at age 11 (Melhuish, 2011; Sammons, Sylva, Melhuish, Siraj-Blatchford, & Hunt, 2008): this research shows that the magnitudes of the benefits are equal to or exceed the effects of several negative factors, including early developmental problems (Melhuish, 2011). Furthermore, the importance of children’s early stages for behavioural intervention in the typical neurodevelopmental disorder, autism spectrum disorder (ASD) is clear. Behavioural interventions launched in young children using the Early-Start-Denver Model (Estes et al., 2015), and infant preemptive (Whitehouse et al., 2021) interventions after the long-term developmental course of autism, with core ASD symptoms severity reduced (Estes et al., 2015; Whitehouse et al., 2021).

These suggest that the early childhood stage is critical for neurodevelopment, learning, cognition, and behaviour. For optimal typical development, it’s important to seize this critical period for early learning activities (Hines, McCartney, Mervis, & Wible, 2011) and make these experiences appropriate, informative and enriching. However, the relationships between specific forms of early learning experiences, such as early RfP, and the brain, cognition and mental health later-in-life still remain unclear.

Several studies have shown that reading acquisition induces brain functional and anatomical signatures (Dehaene, Cohen, Morais, & Kolinsky, 2015; Dehaene-Lambertz, Monzalvo, & Dehaene, 2018), however, to the best of our knowledge, no previous study has investigated the relationships of early RfP experiences with young adolescent brain structure, cognition and mental well-being. The recent adolescent brain and cognitive development (ABCD) project, which started from 2018, recruited more than ten thousand US young adolescents and provided us an opportunity to conduct such a study. We analysed a wide range of clinical interviews and measurement data across multiple domains related to the cognitive, mental, behavioural, neuroimaging and genotype indices of developing young adolescents. We aimed to investigate whether, in comparison with peers who began RfP quite late or never had it, those who began it earlier and had accumulated early RfP experiences would show positive associations with better subsequent young adolescent performance, including higher cognitive assessment scores, lower psychopathological or behavioural problems scores, and increased school academic achievement during typical development. We also aimed to examine the related brain structures and determine whether our findings were mediated by brain structure. We also performed twin study and Mendelian randomization (MR) analyses to investigate the genetic heritability and potential causal relationships, respectively.

Materials and methods

Participants

Study design is shown in online Supplementary diagram.1. Data from the 3.0 (for baseline data collected between 2016–2018,
ages 9–11) and latest 4.0 (for follow-up data collected between 2017–2020, ages 11–13) annual curated data releases of the ABCD project (https://abcdstudy.org/about/) were included. It’s recognized that adolescence begins with the onset of pubertal maturation that typically occurs between the ages of 9 and 12, and is usually 1–2 years earlier in girls (Crone & Dahl, 2012). Since the World Health Organization (WHO) defined adolescence as starting at the 10th birthday, and because the majority of the ABCD participants were close to or older than this age, we, therefore, refer to them as young adolescents. Participants were recruited from 21 research sites throughout the US. The details of participant recruitment and more information are provided in online Supplementary Method S1. The ABCD project conforms to the ethical guidelines of each research site’s Institutional Review Board (IRB). The IRB committee of the University of California, San Diego, is responsible for the ethical oversight of the ABCD research. All participants provided written consent forms and informed consent forms. The demographic characteristics are summarized in Table 1.

**Childhood reading for pleasure (RfP) measurement**

The two key childhood RfP measurement scores we investigated in the current study were obtained from the ABCD summary scores sports and activities (abcd_spaccss01) scale completed by participants’ parents or caregivers regarding their children. (1) The first measurement reflects participants’ previous RfP experiences throughout childhood, which is question_1, namely ‘How many years has your child read for pleasure (RfP)?’ (data ID: sai_ss_read_years_p). A higher score in question_1 indicates that the young participant has enjoyed RfP for more years during childhood, who relatively began to have this activity at earlier ages (representing early-initiated long-standing childhood RfP experiences, which is abbreviated as ‘early RfP’ in the study). Raw scores of question_1 ranged from [0–2.5] years (late and short-term RfP) to [3–10] years (higher levels of early RfP). 1-year and 2-year follow up data were available, with the question: ‘Since we last saw you on [asnt_timestamp_c], has your child read for pleasure?’ While young children’s early RfP needs assistance from caregivers or teachers substantially, the RfP of young adolescents is a relatively more active and independent activity (youth hobby). (2) The second measurement represents participants’ regular RfP durations per week, which is question_2, namely ‘Approximately how many hours per week does your child spend reading for pleasure?’ (data ID: sai_ss_read_hours_p). A higher score in question_2 indicates that the participant usually spent, on average, more time per week on RfP. The 2-year follow-up data of RfP duration was available and the question was the same as its baseline recording.

**Cognitive and psychopathology symptoms assessment scales**

An extensive battery of clinical interviews had been completed by young adolescents and their caregivers. For neurocognitive assessments, we focused on the well-validated National Institutes of Health (NIH) TB cognition summary scale (abcd_tbs01), which was designed to harmonize data collection across NIH-funded projects and to facilitate cross-study comparisons. It contains two standardized indices of crystallized and fluid composites of cognition that are comparable to widely used IQ measures (Akshoomoff et al., 2013). It also evaluates the neurocognition of the following seven major subdomain components (Weintraub et al., 2013).

For psychopathology symptoms and behavioural assessments, we focused on the standard parent-reported Child_Behaviour_CheckList (CBCL) summary scale (abcd_cbcls01), which contains the 20 subdomain assessments of dimensional psychopathology and adaptive functioning (Achenbach & Rescorla, 2004) scores that applied to ABCD participants.

Details of all the subdomains (subscales) and their average scores from the cognitive and psychopathological summary scales in the ABCD dataset are listed in online Supplementary Table S1.

**Brain structural neuroimaging data and preprocessing**

In this study, quality-controlled (QC) high-resolution neuroimaging data processed with FreeSurfer were analysed. The ABCD neuroimaging data were obtained using 3T scanners with 32-channel head coil and high resolution T1-weighted structural magnetic resonance imaging (MRI). Both the methods and evaluations of these MRI images have been harmonized and optimized across all ABCD research sites (Casey et al., 2018; Hagler et al., 2019). According to the ABCD standard pipeline and protocol, pre-processed T1 sMRI data were obtained. The pre-processing processes were completed by the ABCD research teams, with details described in the image processing paper (Hagler et al., 2019). Whole brain and regional morphometric structure evaluations containing cortical volume, thickness, and surface areas were obtained using FreeSurfer processing, including 148 cortical regions from the FreeSurfer Destrieux (h.aparc.a2009s) atlas, and 40 subcortical segmentations from the ASEG atlas (https://surfer.nmr.mgh.harvard.edu/fswiki/CorticalParcellation), which were used to conduct subsequent analysis of the associations between brain structures and RfP measurements. Online Supplementary Method S2 contains additional information on sMRI.

The QC procedure of the processed neuroimages was checked by the ABCD team both automatically and manually. According to the FreeSurfer reconstruction QC measures (freesqc01), 529 participants who failed QC were removed from structural analyses.

**Two-sample MR analysis**

To evaluate the potential causal relationship, we conducted the MR analysis, utilizing independent genetic variants from the largest available nonoverlapping genome-wide association studies (GWAS). The analysis was conducted using the ‘TwoSampleMR’ package (Choi et al., 2019; Rosoff et al., 2021) in R4.0.3. The method requires that all genetic variants be valid instrumental variables based on the MR assumptions (online Supplementary Fig. S1).

The two independent exposure and outcome data for MR that contained the following indices were harmonized using the ‘TwoSampleMR’ package, which included single-nucleotide polymorphisms (SNP)s, p values, effect alleles (with frequencies), reference alleles, estimated effect sizes, and standard errors (s.e.s). For potential causal association, the MR estimates were combined using inverse variance weighted (IVW) meta-analysis, which was equivalent to a weighted regression of the SNP-outcome coefficients on SNP-exposure coefficients to make an overall estimate of causal effect, while the intercept was constrained to zero (Choi et al., 2019).

We used the standard IVW method of MR analysis to determine potential causal relationships. Two other established MR methods were also tested, including: (1) weighted median (WM) and (2) the MR–Egger regression that is recognized as
being more robust to horizontal pleiotropy but at the expense of decreased statistical power (Choi et al., 2019; Hemani, Bowden, & Davey Smith, 2018). MR-Egger intercept, MR-PRESSO, and MR heterogeneity tests were conducted for sensitivity analyses to detect pleiotropy and heterogeneity.

Detailed information on the RfP-related GWAS analysis, as well as the exposure and outcome data preparation for MR analyses can be found in online Supplementary Method S3.

### Statistical analysis

#### Association analysis

A linear mixed model (LMM) was applied to identify the associations between early RfP, and the wide range of assessment scales including cognition and mental problem scores, and brain structure. The LMM model was recommended by the ABCD research consortium and has been used in previous publications (Hackman et al., 2021; Marshall et al., 2020) to consider the correlated observations within family structures due to siblings and twins and at different sites. In this study, LMM was performed using the `fitlme` function in MATLAB (R2019b), which was specialized to model families nested within study sites. Generally, in the LMM model, the dependent variables were scores related to cognitive, mental assessments and brain structural measurements. The target fixed effects were the early RfP and RfP durations. The nuisance covariates and random effects were the family structures nested within different sites. Other covariates that were controlled for LMM included age, sex, BMI, puberty, race/ethnicity, and family social-economic status (SES, including parental education and family income). For brain structure related analysis, the different types of MRI scanners were also added as covariates.

### Table 1. Demographic characteristics of participants in the study

<table>
<thead>
<tr>
<th>Variables</th>
<th>Late RfP</th>
<th>Middle- to early-initiated long-standing RfP (N = 4172, having RfP for 3–5.5 years)</th>
<th>Early- to very early-initiated long-standing RfP (N = 1129, having RfP for 6–10 years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total participants of ABCD cohort, N = 11 878</td>
<td>118.98 (±7.49)</td>
<td>118.39 (±7.49)</td>
<td>119.21 (±7.5)</td>
</tr>
<tr>
<td>Age (months)</td>
<td>5682 (47.84)</td>
<td>2004 (40.55)</td>
<td>2251 (53.95)</td>
</tr>
<tr>
<td>Female</td>
<td>6196 (52.16)</td>
<td>2938 (59.45)</td>
<td>1921 (46.05)</td>
</tr>
<tr>
<td>Missing/undefined</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Race Ethnicity</td>
<td>252 (2.12)</td>
<td>69 (1.40)</td>
<td>111 (2.66)</td>
</tr>
<tr>
<td>Asian</td>
<td>1784 (15.02)</td>
<td>750 (15.18)</td>
<td>407 (9.76)</td>
</tr>
<tr>
<td>Black</td>
<td>2411 (20.31)</td>
<td>1150 (23.27)</td>
<td>664 (15.92)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1247 (10.50)</td>
<td>504 (10.20)</td>
<td>432 (10.35)</td>
</tr>
<tr>
<td>Other</td>
<td>6182 (52.05)</td>
<td>2469 (49.95)</td>
<td>2558 (61.31)</td>
</tr>
<tr>
<td>Missing/undefined</td>
<td>2 (0.00)</td>
<td>2 (0.00)</td>
<td>2 (0.00)</td>
</tr>
<tr>
<td>Highest Parental Education</td>
<td>1725 (14.52)</td>
<td>806 (16.30)</td>
<td>314 (7.53)</td>
</tr>
<tr>
<td>College and Bachelor</td>
<td>6095 (51.31)</td>
<td>2751 (55.67)</td>
<td>2034 (48.75)</td>
</tr>
<tr>
<td>Post Graduate Degree</td>
<td>4044 (34.04)</td>
<td>1385 (28.03)</td>
<td>1824 (43.72)</td>
</tr>
<tr>
<td>Missing/undefined</td>
<td>15 (0.13)</td>
<td>15 (0.13)</td>
<td>15 (0.13)</td>
</tr>
<tr>
<td>Family Income</td>
<td>3224 (27.14)</td>
<td>1759 (35.59)</td>
<td>904 (21.67)</td>
</tr>
<tr>
<td>&lt;$50 000</td>
<td>3071 (25.85)</td>
<td>1438 (29.10)</td>
<td>1182 (28.33)</td>
</tr>
<tr>
<td>≥$50 000 &amp; &lt;100 000</td>
<td>4565 (38.43)</td>
<td>1745 (35.31)</td>
<td>2086 (50.00)</td>
</tr>
<tr>
<td>Missing/undefined</td>
<td>1019 (8.58)</td>
<td>1019 (8.58)</td>
<td>1019 (8.58)</td>
</tr>
</tbody>
</table>

RfP, reading for pleasure; ABCD, adolescent brain cognitive development study.

Of the total 11 878 participants in the ABCD cohort, we included 10 243 participants (86.2%) who had full data recordings for the RfP measurements. While 48.2% of the participants began RfP quite late or never had it (0–2.5 years, N = 4942), 40.8% of the participants had higher levels of early RfP (3–5.5 years, N = 4172), and had spent more early years enjoying this activity compared to the former. The remaining 11% began RfP activities very early (more than 6 years, N = 1129).
Parental education was defined by the highest education level achieved by both parents. RfP, age, BMI, family income and parental education were continuous factors. Sex, race/ethnicity and MRI scanners were categorical factors. Bonferroni-correction was used for multiple comparisons. More details of the LMM model are shown in online Supplementary Method.S4.

**Longitudinal and mediation analyses and twin study using structural equation modelling**

First, for longitudinal analysis, a classic two-wave cross-lagged panel model (CLPM) was conducted to perform the longitudinal analysis of the RfP measurements with cognitive assessments, and attention problems scores, which was implemented using Mplus 7.4. Subsequently, for mediation analysis, we conducted the standard mediation analysis using the Mediation toolbox (https://github.com/canlab/MediationToolbox) developed by Tor Wager’s team. Thirdly, for the twin study, heritability analysis was performed using the OpenMx V 2.20.6. R statistical package with structural equation modelling of the standard twin ACE (A, additive genetic factors; C, shared/common environmental component; E, unique/specific unique environmental component) statistical model controlled for cofounders. Details are shown in online Supplementary Method.S5.

**Results**

**The associations of early RfP with young adolescent cognitive performance and mental problems scores in the ABCD dataset**

Of the 10,243 participants with complete RfP recordings in the ABCD cohort, 4942 (48.2%) never had RfP or began it until much later (late short-term RfP, for 0–2.5 years), and the remaining half had higher levels of early RfP (for 3–10 years) (Table 1, demographic characteristics).

As shown in Fig. 1a and online Supplementary Table S2, covariates-adjusted LMM association analyses indicated that early RfP was significantly positively correlated with young adolescents’ standard cognitive performance scores (Bonferroni-corrected $p$, $P_{\text{Bonferroni}} < 0.05$), including: (1) the cognition summary scale of NIH-TB (abcd_tbs01, $r_{\text{LMM}}$ values ranging from 0.041 to 0.331), and (2) the verbal-learning and immediate-memory scale (abcd_ps01, $r_{\text{LMM}}$: 0.035–0.154). Additionally, early RfP was also correlated with their sleep development ($r_{\text{LMM}} = 0.140$, $p < 0.001$) and school academic achievement (school performance: $r_{\text{LMM}} = 0.269$, $p < 0.001$; course grades: $r_{\text{LMM}} = 0.199$, $p < 0.001$) (online Supplementary Table S3), meanwhile was less significantly correlated with how early they said their first words ($r_{\text{LMM}} = 0.044$, $p < 0.001$).

In contrast, results showed significant negative correlations between early RfP and young adolescents’ mental problems scores ($P_{\text{Bonferroni}} < 0.05$), especially: (1) dimensionality psychopathology and adaptive functioning assessed by the CBCL (abcd_cbcl01, $r_{\text{LMM}}$: −0.034 to −0.139), and its summary – the psychopathology symptoms scale (abcd_cbcls01, $r_{\text{LMM}}$: −0.040 to −0.105); (2) brief problems monitor-reported by teachers (abcd_hpmto1, $r_{\text{LMM}}$: −0.047 to −0.118), and its summary scale (abcd_ssbmto1, $r_{\text{LMM}}$: −0.057 to −0.110); (3) UPPS-P impulsivity (abcd_upps01, $r_{\text{LMM}}$: −0.032 to −0.072). Besides, ABCD parent diagnostic interview for mental disorders (abcd_ksd01, $r_{\text{LMM}}$: −0.034 to −0.130) also showed negative associations.

Importantly, early RfP was significantly negatively correlated with youth total screen time (total time spent on electronic devices including TV, cellphone and tablet etc., during weekdays: $r_{\text{LMM}} = −0.097$, $p < 0.001$; during weekends: $r_{\text{LMM}} = −0.062$, $p < 0.001$) (online Supplementary Fig. S2), and positively correlated with their sleep duration ($r_{\text{LMM}} = 0.068$, $p < 0.001$).

These results indicated that the cognitive performance was better and the mental problems were lower in young adolescents with higher levels of early RfP. For the core constituent subdomains of the summary scales of cognition (abcd_tbs01) and psychopathology symptoms (abcd_cbcl01) (Fig. 1b, c), the most positively correlated cognitive subscales included: (1) the cognition crystallized composite ($r_{\text{LMM}} = 0.329$, $p < 0.001$) that consisted of oral reading and picture vocabulary, which was more dependent upon learning experiences (Akshoomoff et al., 2013); (2) the cognition total composite ($r_{\text{LMM}} = 0.267$, $p < 0.001$) and (3) cognition fluid composite ($r_{\text{LMM}} = 0.150$, $p < 0.001$) scores. Meanwhile, the top-ranked negatively correlated psychopathological subdomains were: (1) attention problems ($r_{\text{LMM}} = −0.106$, $p < 0.001$) and ADHD ($r_{\text{LMM}} = −0.098$, $p < 0.001$) scores; (2) conduct ($r_{\text{LMM}} = −0.085$, $p < 0.001$) and (3) externalizing ($r_{\text{LMM}} = −0.082$, $p < 0.001$) problems scores. (4) The total problems ($r_{\text{LMM}} = −0.070$, $p < 0.001$) was also negatively correlated with early RfP (Fig. 1b–i, online Supplementary Table S4, $P_{\text{Bonferroni}} < 0.05$). All covariates were regressed. Scatter-plots showing the consistent representative correlations based on raw data are in online Supplementary Fig. S3). Other significant subdomains included positively correlated list sorting, picture sequence memory of cognition, and negatively correlated rule-breaking, aggressive, stress, social and depressive problems scores etc. These findings were consistent in both young female and male subgroups (online Supplementary Table S5, $P_{\text{Bonferroni}} < 0.05$).

In all analyses, accounting for potential confounds, we adjusted age, sex, BMI, puberty, race/ethnicity, family SES (including parental education and family income) and family structures nested within research sites in the LMM association model as covariates. Furthermore, sensitivity analysis was performed on three different groups of confounders to examine if there were potential influences of covariate adjustments, and the results confirmed no significant influence of covariates on the results (online Supplementary Table S6).

Considering the potential high rates of comorbidity (25–40%) in developmental dyslexia (DD) and ADHD (Mascheretti et al., 2017), we carried out a series of post-hoc analyses to evaluate the relationships between early RfP and cognitive and mental well-being and how they were associated with ADHD scores. The 2 main sub-groups of interest were young adolescents who were typically developing (TD) and had no comorbidities (the TD group) and those with ADHD (the ADHD group). For each subgroup, we regressed cognitive or psychopathological scores on early RfP via LMM models with confounds adjusted. Of the 10,243 participants we investigated, none was currently diagnosed with ASD or schizophrenia, and 930 (9.1%) met the diagnostic criteria for ADHD (Dworkin et al., 2022). Among the ADHD group, there’s a higher proportion of late short-term RfP (62.22% had RfP for 0–2.5 years, compared with 46.71% in TD group). The significant correlations of early RfP with youth higher comprehensive cognitive and lower psychopathological scores remained robust in the TD group ($N = 9313$), whereas only positive correlations with cognition were observed in the ADHD group (online Supplementary Table S7, $P_{\text{Bonferroni}} < 0.05$). Further, in the TD group, for those who frequently read on screen devices, early RfP didn’t show significant negative correlations.
Figure 1. Early RfP, cognitive and mental health assessments in young adolescents. (a) The correlations of early RfP with young adolescents’ comprehensive assessments across multiple domains related to cognition, behaviour and health in the ABCD dataset. Assessments were classified into 6 main categories: physical health, cognition, mental problems, substance use, culture and environment, and screen time (details of all assessment scales are defined in online Supplementary material S1). Each data point represents an $r_{LMM}$ value calculated from association analysis between early RfP and a measurement subscale. The 10 assessment scales that were most significantly correlated with early RfP were highlighted and listed below: (1) Physical health (positive correlated): abcd_saiq02 ABCD parent-reported sports and activities involvement regarding their children (including RfP-related measurements). (2) Cognition (positive correlated): abcd_ps01 ABCD Pearson scores on verbal-learning and immediate-memory. abcd_tbss01 ABCD youth NIH-TB cognition summary scale. (3) Mental problems (negative correlated): abcd_bpmt01, and abcd_ssbpmt01 ABCD brief problem monitor teacher-reported form, and its summary scale. abcd_cbcl01, and abcd_cbcls01 dimensional psychopathology and adaptive functioning scores assessed by the CBCL, and the CBCL summary scale of psychopathology symptoms. abcd_ksad01 ABCD parent diagnostic interview for DSM5 full. abcd_upps01 UPPS-P for children short form (for impulsivity). (4) Screen time (negative correlated): abcd_stq01 ABCD youth screen time surveys. (b, c) The correlations of early RfP with the core constituent subdomains of the NIH-TB cognition summary (abcd_tbss01, all core subscales were significant) (b), and the CBCL psychopathology scores summary (abcd_cbcls01, 11 out of the total 20 core subscales were significant) (c), respectively. (d–f) Representative density scatter plots showing that the crystallized composite (nihtbx_cryst), total composite (nihtbx_totalcomp) and fluid composite (nihtbx_fluidcomp) were the top 3 cognitive subscales that were positively correlated with early RfP. Meanwhile, (g–i) the attention problems (cbcl_scr_syn_attention), conduct (cbcl_scr_dsmd_conduct) and total problems (cbcl_scr_syn_totprob) were the top-ranked negatively correlated psychopathological subscales. Each individual datapoint is coloured by the number of neighbouring datapoints (n_neighbour_points) to display the overall data distribution. Covariates were all adjusted. Bonferroni-corrected $p$ ($\alpha_{\text{Bonferroni}}$) < 0.05.
with dimensional psychopathological symptoms, whereas results in those who usually read printed materials remained robust (online Supplementary Table S8, \( p_{\text{Bonferroni}} < 0.05 \)).

**Early RfP, increased cognition and decreased psychopathology scores are associated with brain structural signatures in young adolescents**

To investigate whether there were potential brain structural signatures related to early RfP in young adolescents, we analysed measures of cortical surface area, cortical thickness and volume which were obtained using FreeSurfer (Destrieux cortical parcellation atlas) on participants’ T1-weighted MRI volumes, as well as the participants’ whole brain measures. The results revealed significant moderate positive correlations between early RfP and youth total brain volume (TBV) (\( r_{\text{LMM}} = 0.060, p < 0.001 \) (Fig. 2a, online Supplementary Table S9)), intracranial volume (ICV) (\( r_{\text{LMM}} = 0.061, p < 0.001 \)), total brain cortical area (\( r_{\text{LMM}} = 0.059, p < 0.001 \)) and volume (\( r_{\text{LMM}} = 0.055, p < 0.001 \)) (online Supplementary Table S9, upper panel). Among all the cortical parcellations, we found the most moderately significantly increased cortical area regions in young adolescents with higher early RfP levels (\( r_{\text{LMM}} = 0.038–0.064, p_{\text{Bonferroni}} < 0.05 \)), which included language-relevant regions and parts of the visual system. These early RfP related cortical areas included, in particular: (1) bilateral: superior and middle temporal regions, insula circular sulci, occipital regions, supramarginal, temporal pole and middle frontal gyri; (2) left-hemisphere (lh): angular, superior frontal and para-hippocampal gyri, pre- and postcentral, and inferior temporal regions; (3) right-hemisphere (rh): anterior cingulate (ACC) region, orbital gyri, and lingual sulcus, etc (Fig. 2b, online Supplementary Table S9, middle panel). These structural findings remained robust in the TD group (\( r_{\text{LMM}} = 0.037–0.061 \)) (online Supplementary Fig. S4, \( p_{\text{Bonferroni}} < 0.05 \)). Early RfP was also moderately positively correlated with increased subcortical regions, including the ventral diencephalon (DC), thalamus, brain stem, and putamen regions (\( r_{\text{LMM}} = 0.039–0.058 \) (Fig. 2c, online Supplementary Table S9, bottom panel). \( p_{\text{Bonferroni}} < 0.05 \).

The majority of these early RfP-related brain structures were important overlapping cortical and subcortical regions that had both significant positive associations with young adolescents’ cognition crystallized composite and total cognition scores, and negative associations with their attention and total psychopathological problems scores (Fig. 2d–i, online Supplementary Table S10, \( p_{\text{Bonferroni}} < 0.05 \)).

Additionally, young adolescents from both low- and high-income families showed significant positive associations of early RfP with increased brain structures (online Supplementary Table S11, \( p_{\text{Bonferroni}} < 0.05 \)).

Further, despite the above observed positive correlations between youth total brain ICV and early RfP, one ROI, the left superior temporal cortical region, remained significantly correlated with early RfP after adding the ICV as a covariate (\( r_{\text{LMM}} = 0.0374, p = 0.026 \), FDR-corrected), while no subcortical structures remained significant under this condition.

Structural diffusion tensor imaging analysis also revealed positive associations of early RfP with white-matter fibre tract volumes of the specific brain regions, including the corticospinal/pyramidal, temporal and superior longitudinal fasciculus, fornix, and inferior-fronto-occipital and anterior thalamic radiations (\( r_{\text{LMM}} = 0.045–0.060 \) (online Supplementary Table S12, \( p_{\text{Bonferroni}} < 0.05 \)).

**The longitudinal and mediation analyses**

We performed longitudinal analysis using CLPM structural equation modelling to analyse the changes of assessment scores between the baseline recording and 2-years later, with covariates adjusted. Results indicated that higher early RfP levels recorded at baseline year were significantly associated with improved crystallized cognition (\( \beta = 0.238, \text{s.e.} = 0.024, p < 0.001 \)) and reduced attention problems scores (\( \beta = -0.031, \text{s.e.} = 0.012, p = 0.012 \) at the follow-up in young adolescents, and the reverse directions were also significant (crystallized cognition: \( \beta = 0.011, \text{s.e.} = 0.002, p < 0.001; \) attention problems: \( \beta = -0.008, \text{s.e.} = 0.002, p = 0.001 \)) (Fig. 3a, b).

Subsequent mediation analysis showed that the mean cortical area of the above identified brain regions which were significantly correlated with both early RfP and the cognitive assessments (path AB, regions of Fig. 2f) significantly partially mediated the associations between early RfP and youth cognitive scores: (1) for the cognition crystallized score, \( \beta_{\text{path AB}} = 0.025, p < 0.001, 95\% \text{ CI 0.019–0.033, } 3\% \text{ of the total effect size (TES) (Fig. 3c); } \) (2) for total cognition score: \( \beta_{\text{path AB}} = 0.026, p < 0.001, 95\% \text{ CI 0.019–0.035, } 3\% \text{ of TES (Fig. 3d). } \) A similar finding was also observed for psychopathological symptoms scores, which indicated a significant partial mediation effect of the cortical area (regions of Fig. 2i): (1) for attention problems score: \( \beta_{\text{path AB}} = -0.008, p < 0.001, 95\% \text{ CI } -0.011 \text{ to } -0.005, 6\% \text{ of TES (Fig. 3e); } \) (2) for total problem score: \( \beta_{\text{path AB}} = -0.046, p < 0.001, 95\% \text{ CI } -0.065 \text{ to } -0.031, 8.6\% \text{ of TES (Fig. 3f). } \) The total brain measurements of TBV (online Supplementary Fig. S5, \( p < 0.001 \)). The ICV displayed closely similar results to the TBV) and subcortical structures (online Supplementary Fig. S6, \( p < 0.001 \)) also showed significant mediation effects. Further ROI analysis showed that the mediation effect of the left superior temporal cortical area region was significant after controlling for ICV as a covariate (online Supplementary Fig. S7, \( p < 0.001 \)). Our hypotheses were confirmed by the mediation analysis that the increased brain cortical and subcortical structures significantly mediated the associations between higher early RfP scores with increased youth cognition, meanwhile, these brain signatures also significantly mediated the prediction of decreased young adolescent psychopathological scores based on higher early RfP scores. All models indicated the brain was a significant mediator.

**The nonlinear associations between youth regular weekly RfP durations and cognition and brain structure**

Using nonlinear fitting models, we further observed the optimal RfP duration for youth cognitive assessment scores was approximately 12 h per week (h/week) (online Supplementary Fig. S8). For crystallized composite: piecewise linear \( R_{\text{adj}} = 0.162, \) polynomial \( R^2_{\text{adj}} = 0.173, \) GAM \( R^2_{\text{adj}} = 0.178; \) for total cognition composite: piecewise linear \( R_{\text{adj}} = 0.120, \) polynomial \( R^2_{\text{adj}} = 0.128, \) GAM \( R^2_{\text{adj}} = 0.133; \) all above \( p < 0.001 \)). The increasing rates of cognitive scores accompanied by the growth in RfP durations were most notable within the optimal 12 h/week RfP (crystallized composite, \( r_{\text{LMM}} = 0.323, p < 0.001; \) total cognition composite, \( r_{\text{LMM}} = 0.253, p < 0.001 \)). Whereas exceeding 12 h/week RfP, no additional benefits but gradual declines in cognition including crystallized (\( r_{\text{LMM}} = -0.137, p = 0.025 \)) and total composite (\( r_{\text{LMM}} = -0.125, p = 0.073 \)) were observed, accompanied with increases in RfP durations. Excessive weekly RfP duration may be detrimental.
regarding cognition because long time sedentary behaviour may be disadvantageous, and also time for other cognitively enriching activities, including sports and social activities, may possibly be reduced. Meanwhile, we observed mild negative linear associations of RfP durations with attention ($r_{LMM} = -0.045, p < 0.001$) and conduct problems scores ($r_{LMM} = -0.038, p = 0.007$) (online Supplementary Fig. S9, $p_{Bonferroni} < 0.05$). Further, a positive correlation between early RfP and optimal regular RfP durations was also observed ($r_{LMM} = 0.304, p < 0.001$).

Participants with increased weekly RfP durations (within 12 h/week) showed increased cortical volumes in the similar brain regions defined above, including the temporal, frontal, and subcortical putamen, ventral DC and thalamus regions ($r_{LMM} = 0.041–0.052$) (online Supplementary Table S13, $p_{Bonferroni} < 0.05$), while in conditions of more than 12 h/week RfP, no significant change in brain structure was observed in association with increased RfP durations.

The heritability ($h^2$) of early RfP, cognition, attention problems and brain structure

In order to evaluate the relative proportional significance of both genetic and environmental factors for the identified key behaviours and phenotypes, we applied the twin study analysis and measured heritability. Embedded within the ABCD project, its Twin Hub recruited approximately 800 pairs of twins (Iacono et al., 2018), of which 711 pairs had RfP measurements collected, including 317 monozygotic (MZ) twin pairs (150 females, 47.3%), and 394 dizygotic (DZ) twin pairs (203 females, 51.5%). Using this sub-project dataset, we measured heritability of target variables including early RfP under the best fit model of structural equation modelling (online Supplementary Table S14). Moderate heritability of early RfP ($h^2: 0.315, 95\% CI 0.272–0.358$) was observed, and twins’ common environmental factors contributed significantly to their early RfP ($C: 0.505, 95\% CI 0.465–0.545$). Heritability of cognition composite scores ranged from moderate for the crystallized composite ($h^2: 0.461, 95\% CI 0.456–0.467$) to high for the fluid composite ($h^2: 0.608, 95\% CI 0.602–0.614$). In addition, attention problems scores were found to be relatively high in heritability ($h^2: 0.671, 95\% CI 0.652–0.691$), which was consistent with previous observations related to attention problems (Rietveld, Hudziak, Bartels, van Beijsterveldt, & Boomsma, 2004). For brain structure, consistent with previous research using different methodological approaches, genetic factors contributed moderately significantly to brain cortical area (also measured in FreeSurfer.aparc.2009s

Figure 2. Young adolescent brain structures with their cortical areas and subcortical regions linked to early RfP, cognition, and psychopathology assessment scores. (a) Associations of early RfP with total cortical volume (left panel, mm$^3$) and total brain volume (TBV) (right panel, mm$^3$). (b, c) Brain maps showing the specific cortical areas and subcortical regions that were moderately significantly increased in young adolescents with higher levels of early RfP ($r_{LMM}$ values ranging from 0.038–0.064). Brain regions with larger areas/volumes positively associated with early RfP are represented by the red colour. (d, e) Cortical areas that had significant positive associations with the cognition crystallized composite score (in d), and total cognition score (in e). Regions where a larger area was positively associated with a higher cognition score are represented by the red colour. (Only the regions with $r_{LMM} > 0.055$ were shown here). (f) Most of the increased cortical areas related to early RfP shown in (b) were overlapping regions that were also positively associated with the cognition crystallized score and total cognition score. (g, h) Cortical areas that had significant negative associations with attention problems (in g), and total problems score (in h). Regions having a negative association between brain area and psychopathological assessment (i.e. a reduced cortical area was associated with increased attention problems score) are represented by the blue colour. (i) The overlapping brain regions with their areas positively associated with early RfP and negatively associated with the attention problems and total problems scores are shown. Covariates were all adjusted. $p_{Bonferroni} < 0.05$. 

https://doi.org/10.1017/S0033291723001381 Published online by Cambridge University Press
atlas) (Elliott et al., 2018) including the ROI superior temporal area (h²: 0.261, 95% CI 0.220–0.301), revealing the significant roles of environmental factors.

Two-sample MR analysis on the relationships between early RfP, adult cognitive performance and brain structure, and attention disorder

Using the standard IVW approach of MR analysis based on independent SNP genetic variants, we further observed evidence of a beneficial causal association between early RfP (per 1-S.D. unit increase) and better adult cognitive performance (Fig. 4a, online Supplementary Table S15) (β_IVW = 0.026, 95% CI 0.006–0.045, p = 0.009). The WM analysis was also significant (β_WM = 0.035, 95% CI 0.010–0.060, p = 0.006), and the MR-Egger results indicated the similar pattern of effects. Steiger-MR directionality test confirmed a significant causal direction from early RfP to adult cognition (online Supplementary Table S15, bottom panel, p < 0.001). Forest plot showing effect sizes of single and combined SNPs indicated an overall-combined positive effect pattern of early RfP on adult cognitive performance (Fig. 4b), and the estimates remained robust during the leave-one-out sensitivity test.
reading-related early learning experiences encouraged to read by schools and social media (Mol & Bus, 2019), children and adults (Andersen & Nielsen, 2016; Zuckerman, Elansary, & Needelman, 2019), and children are encouraged to read by schools and social media (Mol & Bus, 2019). Here we conducted a US national cohort study on the reading-related experiences — early RfP, and found significant evidence of its associations with three important developmental indicators in young adolescence, including dimensional cognition, mental well-being, and brain structure, which are cornerstones for future learning and well-being.

**Early RfP, cognitive and behavioural assessment scores in young populations**

We first found highly positive associations between early RfP and youth core standard cognitive assessments, with the crystallized composite of cognition (reflecting acquired knowledge) being more significantly correlated with early RfP compared to the fluid composite (reflecting responses to novel situations or problems) (Akshoomoff et al., 2013; Tucker-Drob et al., 2022), after regressing out potential confounds including family SES. We also found increases in overall cognitive performance scores across major neurocognitive subdomains. These are consistent with previous studies which have suggested that overall cognitive capacities including crystallized and fluid composites, despite there being distinctions between them, have modest associations with each other, and can be summarized by a higher-order common factor (general intelligence) (Lovden, Fratiglioni, Glymour, Lindenberger, & Tucker-Drob, 2020; Lubinski, 2004; van der Maas et al., 2006). Subsequently, we found significant negative associations of early RfP with youth attention symptoms scores. Similarly, an observational study using the UK Millennium cohort has reported a link between RfP frequencies in children at age 7 and lower inattention at age 11 (Mak & Fancourt, 2020). Moreover, results have indicated that attention and reading development closely correlate with each other, and learning to read may help promote attention in school-age children (Wang et al., 2022), and this is consistent with our CLPM longitudinal finding that higher early RfP levels were associated with reduced attention problems scores at 2-year follow-up. In addition to the attention, we observed associations between early RfP with other decreased dimensional behavioural and psychopathological symptoms scores, including conduct, externalizing, aggressive, rule-breaking, stress, depressive and total problems. Indeed, many dimensional psychiatric symptoms are not expressed uniquely or diagnosed as a single isolated type (Wise, Robinson, & Gillan, 2022). Interestingly, recent evidence also shows that less-frequent caregiver-reported shared reading in young children is associated with increased risk of their social-emotional problems (Martin et al., 2022).

Importantly, we found a significant negative association of early RfP with youth total screen time (both during weekdays and weekends). Recent studies have indicated that increased screen time is significantly positively associated with multiple psychopathological problems, such as subsequent ADHD symptoms (Ra et al., 2018), internalizing and externalizing problems (Eirich et al., 2022), as well as early developmental problems (Zhao et al., 2022) in young populations. These findings suggest a potential application of early RfP with relevance for interventions on children’s over-extended screen time and electronic device addiction, and related symptoms or problems.

**Early RfP and related youth brain structures**

We further observed that early RfP had moderate significant associations with youth increased cortical area parcellations, including the superior temporal, occipital, supramarginal and left angular regions, which are essential in reading-related cognitive processes (Carreiras et al., 2009; Turkeltaub, Gareau, Flowers, Zeffiro, & Eden, 2003). A neurodevelopmental study of reading (Turkeltaub et al., 2003) has shown that the brain temporoparietal cortices, containing the left superior temporal gyrus, matures early in learning and remains involved in reading throughout adulthood. The superior temporal region functions critically in cross-modal integration, including processes associated with the mapping of print to sound (Turkeltaub et al., 2003). Research in adults reveals that the late-literate group (those who learn to read as adults) have greater grey matter density (GMD) in the left superior temporal and supramarginal, bilateral angular and dorsal occipito regions (Carreiras et al., 2009) than the illiterate groups. Generally, the angular, temporal and supramarginal gyri are activated during reading, and their GMD increases are induced by literacy acquisition (Carreiras et al., 2009; Dehaene et al., 2015). The key brain regions we identified in youth, related to reading acquisition, are similar to those of adults. However, it is not the density of grey matter observed in adults but the measurement of cortical area that is most significantly associated with early reading in the youth population.

In addition to the above reading-related cortical structures, other areas moderately related to early RfP were also found,
including the middle frontal, temporal pole, circular insula, left superior frontal, parahippocampal and inferior temporal, right ACC and orbital frontal gyri. We further demonstrated most of the above identified brain structures were important overlapping regions that were positively associated with young adolescents’ cognitive performance and negatively associated with their psychopathological assessments including attention symptoms and total problems, and these brain structures significantly mediated the associations of early RfP with youth cognition and psychopathology scores. Particularly among these, the temporal (Scott, Blank, Rosen, & Wise, 2000), frontal (Badre & Nee, 2018; Duncan, 2013), ACC (Apps, Rushworth, & Chang, 2016) and orbitofrontal cortices (Rolls & Grabenhorst, 2008) play critical roles in cognitive functions. In contrast, abnormal pathological dysfunctions/defects and alterations in these brain regions are significantly related to multiple psychiatric and mental health disorders (Lai, Payne, Byrum, Steffens, & Krishnan, 2000; Tamminga & Buchsbaum, 2004; van Heukelum et al., 2021), including depression (Lai et al., 2000), and aggression (van Heukelum et al., 2021), etc. Impairments in these regions are also closely related to attention disorder (Hoogman et al., 2019): a large cohort study has shown that children with ADHD have decreased cortical areas of the frontal, temporal, and cingulate regions (Hoogman et al., 2019). In addition to the above cortical structures, we also found subcortical regions modestly related to early RfP including ventral DC, thalamus and putamen regions. A previous study has also found that literacy-induced neuroplasticity improves functional connectivity between the occipital cortex and subcortical regions in the midbrain and the thalamus, which is not just restricted to the cortex, after half-a-year of literacy training in adults (Skeide et al., 2017). Analysis across multiple cohorts has also indicated that patients with attention disorder have decreased subcortical regions including the caudate and putamen, as well as whole ICV (Hoogman et al., 2017). We also concurrently observed moderately positive associations of early RfP with youth total brain measures including ICV and TBV. Interestingly, TBV and ICV are driven by brain growth during normal childhood development (Casey et al., 2005; Courchesne et al., 2000; Sgouros, Goldin, Hockley, Wake, & Natarajan, 1995). Children’s TBV and ICV develop in parallel during childhood until early adolescence, after which TBV declines gradually while ICV stabilizes (Casey et al., 2005; Courchesne et al., 2000; Sgouros et al., 1999). Positive association between TBV and intelligence has been found (Luders, Narr,
Thompson, & Toga, 2009). Significantly reduced TBV or ICV are found in patients with neurodevelopmental-related psychiatric disorders, including ADHD (Hoogman et al., 2017) and schizophrenia (Kahn & Sommer, 2015), which might be caused by disruption of early brain development or underdevelopment (Woodward & Heckers, 2015).

It is notable that based on the data measured from the ABCD participants who were mainly healthy TD young adolescents, our brain structural analysis results revealed that there were not large regionally-specific brain structural changes correlated with their early RfP, but a moderately significant whole brain measurement association, as most of these cortical regional structure associations were not significant after controlling for total ICV as a covariate, with the exception of the moderately increased left superior temporal region.

The 2-sample MR analysis for evaluating potential causal associations

Using large-scale GWAS data and independent SNPs, MR analysis showed that early RfP had beneficial causal associations with adult cognitive performance and brain structure, and a trend towards a protective relationship with ADHD disorder in children and adults. Although the bio-molecular mechanisms through which the identified SNPs influence early RfP are not fully understood, the SNPs with known gene consequences are located in genomic regions closely related to neuronal functions: e.g. rs2253202 is an intronic variant in the potassium channel α-subunit protein KCNA1 that plays an important role in neurodevelopment and neurophysiological processes (Liang et al., 2019); and rs10999029 is related to the synaptic extracellular-matrix protein COL13A1, which is involved in neuromuscular synapses (Cruz et al., 2019). Our results also suggested that early RfP showed a similar beneficial association with cognitive development as educational attainment (EA) at certain levels, since EA is positively associated with subsequent cognitive functions later-in-life (Lowden et al., 2020; Ritchie & Tucker-Drob, 2018). This finding is consistent with our MR-analysis result of a beneficial relationship between early RfP and cognitive performance in adulthood. This may partly be due to cognitive enrichment (Mather, 2020; Milgram, Siwak-Tapp, Araujo, & Head, 2006; Oveisghar, Wilson, Yu, Schneider, & Bennett, 2020). Moreover, we observed a positive association between early RfP and youth school academic achievement (including school performance and grades). Taken together, our findings on early RfP extend the importance of formal education to an enjoyable educational and learning activity, which is both applicable in families and school/educational services.

Potential societal relevance

Early RfP in all children may support the best possible chance for good cognitive development and mental health. Making the enjoyable learning initiative and support available to a broad range of young populations, may have beneficial links with subsequent performances in cognition, the brain, well-being and school achievement in young adolescence and possibly beyond. Caregivers play an important role in children’s RfP at early ages. Our findings have implications for caregivers, teachers, and policymakers in supporting early RfP in children from families with different income levels, and for educators and publishers in providing well-designed reading materials for children. In addition, learning initiatives started early may also help in the earlier detection of related cognitive or behavioural abnormalities, and allow for timely interventional strategies (Engle et al., 2007; Shaywitz et al., 2008) and to reduce early deficits (Sasser, Bierman, Heinrichs, & Nix, 2017). Furthermore, taking into account of the impact that the COVID-19 pandemic and lockdowns had on young children's cognitive development, RfP-related activities may help mitigate the negative effects of the pandemic and lockdowns on their emotion, cognition and education.

Strength and limitations

The current study has several strengths: first, the large sample size of the ABCD dataset, which includes many behavioural and phenotypical measurements of young adolescents. Second, the project’s longitudinal design, which allowed us to examine the relationship between early RfP and cognitive and behavioural assessments measured two years later. Third is the twin hub data of the ABCD study, which has enabled us to analyse the influence of genetic heritability and early shared environment on their early RfP as well as other important phenotypes. Fourth, we performed the standard 2-sample MR analysis using genetic instruments obtained from the most recent largest independent GWAS summaries, including the ABCD and UK-Biobank cohorts.

There are still some questions that remain to be investigated. Future analyses that include an investigation of the long-term follow-up data from the ABCD study in later adolescence and young adulthood and the associations of the findings with early RfP data should be conducted when subsequent longitudinal data becomes available. Future studies should also examine data from other important early educational experiences including play- and materials-based physical and social learning.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0033291723001381

Acknowledgements. We thank all the participants and their families for their engagement in the ABCD project and for their time and dedication. We acknowledge that the phenotypic and neuroimaging data used in this study were obtained from the ABCD project (https://abcdstudy.org). The ABCD Project is supported by the NIH and additional federal partners. A list of all supporters can be found at https://abcdstudy.org/federal-partners.html.

Author contributions. Conceptualization: JF and BJS. Conceived the study and designed the analysis: JF, YJS, BJS, WC, CL, AYY, CHL and XMZ. Methodology: JF, BJS, YJS, WC, AYY, CL, JJK. Investigation and visualization: YJS, WC, CL. Supervision: JF and BJS. Writing – original draft: YJS. Writing – review & editing: BJS, CL, JF, YJS, WC, AYY, YCJ, CHL, XMZ and JJK. Project administration: JF and BJS.

Financial support. Prof. Sahakian receives funding from a Wellcome Trust Collaborative Award (200181/Z/15/Z), and the research is conducted within the NIHR Cambridge Biomedical Research Centre (BRC) (Mental Health Theme and Neurodegeneration Theme) and the NIHR MedTech and in vitro diagnostic Co-operative (MIC).

Prof. Feng was supported by National Key R&D Program of China (No. 2019YFA0709502), National Key R&D Program of China (No. 2018YFC1312904), Shanghai Municipal Science and Technology Major Project (No. 2018SHZDZX01), ZJ Lab, Shanghai Center for Brain Science and Brain-Inspired Technology, and the 111 Project (No. B18015), Guangdong Key Research and Development Project (No.2018B030335001)

Dr Sun was supported by grants from the National Natural Science Foundation of China (No. 31700892).


