eLife

Reviewed Preprint
v1 • December 23, 2024
Not revised

Reviewed Preprint
v2 • October 29, 2026
Revised by authors

For correspondence:
florence.bouhali@univ-amu.fr

* Senior authors

Competing interests: No competing interests declared

Funding: See page 24

Reviewing editor: Saad Jbabdi, University of Oxford, United Kingdom

© 2024, Bouhali et al. This article is distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use and redistribution provided that the original author and source are credited.

Unique longitudinal contributions of sulcal interruptions to reading acquisition in children

Florence Bouhali[1,2], Jessica Dubois[3,4], Fumiko Hoeft[1,5,*], Kevin S Weiner[6,*]

[1]Department of Psychiatry and Behavioral Sciences & Weil Institute of Neuroscience, University of California, San Francisco, San Francisco, United States • [2]Aix-Marseille University, CNRS, CRPN, Marseille, France • [3]University Paris Cité, NeuroDiderot, INSERM, Paris, France • [4]University Paris-Saclay, NeuroSpin, UNIACT, CEA, Saclay, France • [5]Department of Psychological Sciences, University of Connecticut, Storrs, United States • [6]Department of Psychology, Department of Neuroscience, Helen Wills Neuroscience Institute, University of California, Berkeley, Berkeley, United States

eLife Assessment

This **valuable** study relating brain anatomy to reading ability in children studied longitudinally finds that a difference in sulcal anatomy is more predictive of reading ability than cognitive measures. The evidence that this anatomical difference is predictive and confers some benefit in the connectivity between brain areas is based on careful analyses and is **solid**. The findings would be of interest to researchers studying brain and cognitive development.

https://doi.org/10.7554/eLife.103007.2.sa3

Abstract

A growing body of literature indicates strong associations between indentations of the cerebral cortex (i.e., sulci) and individual differences in cognitive performance. Interruptions, or gaps, of sulci (historically known as *pli de passage*) are particularly intriguing as previous work suggests that these interruptions have a causal effect on cognitive development. Here, we tested how the presence and morphology of sulcal interruptions in the left posterior occipitotemporal sulcus (pOTS) longitudinally impact the development of a culturally-acquired skill: reading. Forty-three children were successfully followed from age 5 in kindergarten, at the onset of literacy instruction, to ages 7 and 8 with assessments of cognitive, pre-literacy, and literacy skills, as well as MRI anatomical scans at ages 5 and 8. Crucially, we demonstrate that the presence of a left pOTS gap at 5 years is a robust and relatively specific longitudinal predictor of better future reading skills in children, with large observed benefits on reading behavior ranging from pseudoword decoding to reading comprehension. The effect of left pOTS interruptions on reading acquisition accumulated through time and showed strong predictive power relative to benchmark cognitive and familial predictors of reading ability and disability. Finally, we show that increased local U-fiber white matter connectivity associated with such sulcal interruptions possibly underlie these behavioral benefits, by providing a computational advantage. To our knowledge, this is the first quantitative evidence supporting a potential integrative gray-white matter mechanism underlying the cognitive benefits of macro-anatomical differences in sulcal morphology related to longitudinal improvements in a culturally-acquired skill.

Introduction

Identifying neuroanatomical features associated with uniquely human cognitive abilities is a central goal in cognitive neuroscience. A growing number of empirical findings identify relationships between the morphology of indentations (known as sulci) in the cerebral cortex and individual differences in cognition (e.g., as reviewed in Cachia et al., 2021; Mangin et al., 2019;

Weiner, 2019, 2023). For example, sulcal morphology in various human association cortices that have expanded significantly throughout evolution has been related to individual differences in abstract reasoning, executive function, inhibitory control, memory, face processing, and reading, as well as psychopathology (Amiez et al., 2018; Borst et al., 2014; Cachia et al., 2014; Garrison et al., 2015; Im et al., 2016; Lahutsina et al., 2023; Maboudian et al., 2024; Parker et al., 2023; Tissier et al., 2018; Voorhies et al., 2021; Willbrand et al., 2022; Yao et al., 2022). Moreover, recent research suggests that individual differences in sulcal interruptions are also related to individual differences in cognitive abilities such as numerical processing (Roell et al., 2021; Schwizer Ashkenazi et al., 2024), language and memory (Santacroce et al., 2024), and reading (Borst et al., 2016; Cachia et al., 2018). Sulcal interruptions (hereby referred to as gyral gaps) are a particular case of interconnecting gyri, historically known as pli de passage as coined by Gratiolet (1854), when they emerge on the brain surface. Nevertheless, it is presently unknown whether gyral gaps that are uniquely human are longitudinally (and thus likely causally) related to changes in cognitive abilities, especially in "evolutionarily new" domains such as reading.

For instance, previous work suggests that the maturation of a functional region for reading in ventral temporal cortex (hereafter VTC) when children learn to read, commonly referred to as the visual word form area (VWFA; Cohen et al., 2000; Dehaene et al., 2015; Dehaene & Cohen, 2011), is in part dependent on the pre-existing neural architecture of this region, in terms of both connectivity and cortical morphology (Beelen et al., 2019; Bouhali et al., 2014; Li et al., 2020; Moulton et al., 2019; Saygin et al., 2016; Skeide et al., 2016; Vandermosten et al., 2016). Of note, the human OTS is laterally displaced compared to the OTS in non-human primates (Natu et al., 2021). Additionally, recent studies have linked individual differences in the presence of gyral gaps in this laterally displaced human OTS to individual differences in reading skills, although findings have not been entirely uniform (Borst et al., 2016; Cachia et al., 2018; Del Mauro et al., 2022; Roell et al., 2022).

However, no study has yet explored if this unique neuroanatomical feature in human VTC is related to individual differences in reading skills, as well as other reading precursors (and related cognitive skills such as phonological awareness) as children learn to read - and additionally, if white matter changes may mediate this complex relationship.

Specifically, recent studies that identify relationships between sulcal morphology and cognition propose the novel hypothesis that this relationship reflects underlying differences in white matter architecture. Yet, to our knowledge, this hypothesis has not yet been tested. Sulcal interruptions are particularly interesting to study as they have been linked to a higher density of underlying short range white matter fibers (Bodin et al., 2021) - in which, mechanistically, a higher density of short range white matter fibers could provide a more efficient neuroanatomical infrastructure contributing to individual differences in human cognition compared to a less dense, and therefore, inefficient, neuroanatomical infrastructure. To test these options, here, we investigated whether the presence of left OTS interruptions serve as a longitudinal predictor of reading skills for the first time. We also examine the predictive power of these interruptions relative to benchmark cognitive precursors of reading and assess the role of white matter properties that may mediate this relationship.

Utilizing longitudinal data from 43 children followed from kindergarten at age 5 to age 8 on three measurement occasions, with comprehensive behavioral assessments of literacy, pre-literacy, cognitive skills, and MRI scans, our results provide novel evidence supporting that the interruption of the left posterior OTS (pOTS) is a robust longitudinal predictor of better reading skills, independent from—and strongly predictive relative to —typical pre-literacy skills. This relationship holds and accumulates across time, suggesting that the morphology of the left pOTS is a strong and consistent indicator of reading proficiency. This anatomical marker contributes beyond established behavioral and familial predictors of reading, and this association appears relatively specific to literacy rather than reflecting a general cognitive advantage. Increased local U-fiber structural connectivity associated with left pOTS interruption may underpin the reading

benefit of this neural marker. By exploring the neuroanatomical underpinnings of reading, our results highlight the intricate connection between the brain's evolved structure and its capacity to adapt to culturally invented cognitive demands, such as reading.

Results

Due to previous findings showing (i) a link between gyral gaps of the occipito-temporal sulcus (OTS) and reading abilities (Borst et al., 2016; Cachia et al., 2018) and (ii) the importance of the neighboring mid-fusiform sulcus (MFS) as a cytoarchitectonic and functional landmark in VTC (Grill-Spector & Weiner, 2014; Weiner, 2019; Weiner et al., 2014), potentially relevant for graphemic orthographic processing (Bouhali et al., 2019), the OTS and MFS were delineated manually in each hemisphere, without awareness of the reading scores. This was done at each time point with available MRI scans (Time 1 at age 5 - T1, and Time 3 at age 8 - T3) on cortical surface reconstructions (in Freesurfer; Dale et al., 1999; Fischl et al., 1999; Reuter et al., 2012; Zaretskaya et al., 2018; see Materials and Methods) obtained by both cross-sectional and longitudinal pipelines (see Figure 1 for data overview and sulcal definitions, as well as Supplementary Figure 1 in Appendix A for a view of sulcal definitions in all participants). Because there were two T_1-weighted MRI scans per child (at T1 and T3), a within-subject template was created through the Freesurfer longitudinal pipeline to refine brain segmentations at each time point - an approach that increases the precision and discrimination of cortical surface reconstructions (Reuter et al., 2012). Brain-behavior relationships with regard to reading development described were restricted to the left pOTS, and other delineated sulci serve as controls to demonstrate its specificity.

When present, a gyral gap in the posterior OTS is a reliable longitudinal landmark

The OTS showed a very high incidence of interruption as a whole: approximately 79.3% of children showed at least one interruption in the left OTS, while 74.6% showed interruptions in the right OTS (see Supplementary Figure 2 for interruption statistics for all sulci, time points and processing pipelines). The posterior section of the OTS, whose interruption pattern has been more specifically implicated with reading skills (Cachia et al., 2018), had at least one interruption in the left hemisphere in 56.5% of children, and in 58.3% in the right.

We further evaluated the stability and reliability of left pOTS interruption measurements across T1 and T3. Acceptable internal consistency in measuring the presence of left pOTS interruptions between T1 and T3 was observed when using segmentations from the cross-sectional processing pipeline (T1-T3 Cronbach's α = .778, estimated 95% confidence interval based on the procedure by Feldt et al. (1987): CI= [.591 - .880]). However, consistency was excellent and significantly improved with the longitudinal processing pipeline (T1-T3 α = .976, CI= [.958 - .976]). Only one child changed left pOTS interruption classification between T1 and T3 under longitudinal processing, due to the measured inter-component distance crossing the predefined 1-mm threshold, consistent with high stability of this macro-anatomical feature. Similarly, other anatomical measures, including sulcal length, sulcal interruption distance, surface area and thickness, also showed higher consistency using Freesurfer's longitudinal pipeline (see Supplementary Table 1 for the left pOTS). Left pOTS interruptions therefore appeared as a stable anatomical feature in time (at least in the 5- to 8-year-old range), in agreement with the prevailing view that the sulcal morphology of secondary sulci - including interruption patterns - is set in utero (Cachia et al., 2021). Left pOTS interruption was best measured with two T_1-weighted scans using Freesurfer's longitudinal (rather than cross-sectional) processing.

Because our interest lies in how sulcal morphology prospectively predicts the emergence of reading skills over time, we hereafter report the association between reading skills and left pOTS interruption measured through the longitudinal pipeline at Time 1 (T1). We report results from all 43 children who could be processed through this longitudinal pipeline (see participants'

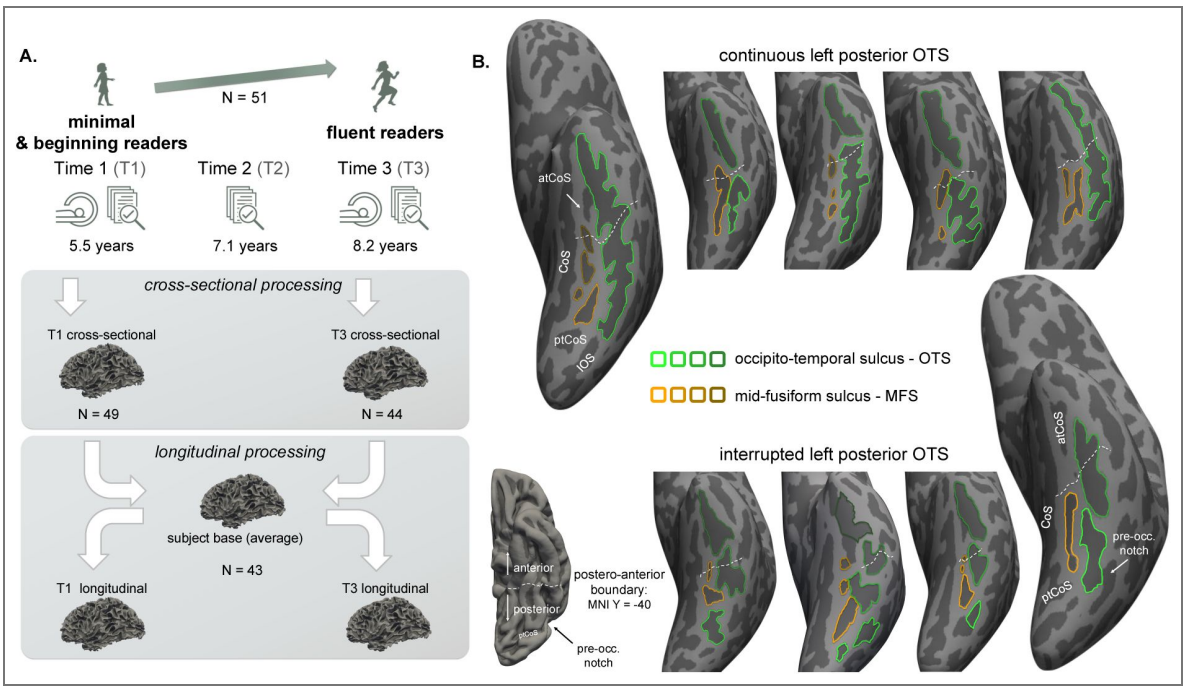


Figure 1. Overview of the experimental design, analytical pipeline, and sulcal definitions.

A. Children were followed longitudinally from approximately 5.5 years of age (Time 1, T1) to age 8.2 (Time 3, T3) with both behavioral assessments and MRI scans, in addition to an intermediate behavior-only assessment at age 7.1 (Time 2, T2). T_1-weighted images from T1 and T3 were first processed separately in Freesurfer (cross-sectional processing), and then through the longitudinal pipeline with the creation of a subject template and refinement of T1 and T3 tissue segmentations. **B.** The occipito-temporal sulcus (OTS, green) and mid-fusiform sulcus (MFS, orange) were defined manually, as shown on 9 sample individual inflated left hemispheres (T1 longitudinal processing). Posterior and anterior sections of ventral temporal cortex were defined in each individual based on a boundary (dotted white line) located at MNI Y=-40 (see projection on Freesurfer average (*fsaverage)* surface on the bottom left, which roughly aligned with the anterior tip of the MFS). Left hemispheric views from exemplar children are shown, either with a continuous left posterior OTS on the top row (i.e., no more than one connex OTS component in the space posterior to the white dotted line) or an interrupted left pOTS on the bottom row (two or more sulcal components present in posterior space, with a total Euclidean distance between components ≥1mm). Green and orange gradient colors are respectively used to highlight the different sulcal components of the OTS and MFS. CoS: collateral sulcus; ptCoS/atCoS: posterior/anterior transverse branch of the CoS; IOS: inferior occipital sulcus.

characteristics in Table 1). However, similar results are observed regardless of type of processing (longitudinal vs. cross-sectional) and time point used for the analyses (see Appendix A - Figure S3A).

Left pOTS interruption is prospectively associated with better reading skills

The presence of interruptions in the left pOTS at T1, measured with the longitudinal processing pipeline, was significantly associated with reading skills measured at T3. Compared with children with a continuous pOTS, children with an interrupted left pOTS at T1 showed higher reading TOWRE composite standard scores at T3, while controlling for demographic variables and total brain volume (see Figure 2A; F-test comparing nested models predicting TOWRE scores by age at T3, handedness and gender only, or by adding the residuals of a logistic regression explaining left pOTS interruption by total brain volume, age at T1, gender and handedness: $F(1, 38) = 11.34$, $p = 1.7*10^{-3}$, $\Delta R^2 = 20.5\%$). This corresponded to a large effect size (Cohen's d on reading residuals accounting for demographic variables: $d = 0.957$).

The association between left pOTS interruption and reading skills was robust to different measurements of reading skills and sulcal interruption in four ways. First, left pOTS interruption (measured at T1 with longitudinal processing) was predictive of reading skills as measured by all reading tests at T3, with a range of $\Delta R^2 = 13.4\%$ to 24.3% of variance explained on a variety of reading tests, ranging from accuracy and speed/fluency reading aloud single words and pseudo-words, to silent sentence reading and passage comprehension, as well as spelling scores (see Figure 2B). All of these associations were statistically significant ($p < .012$). Interruption of the left pOTS was most strongly associated with passage comprehension and sentence reading scores.

Second, associations between sulcal interruption and reading skills were specific to the left pOTS within VTC. No significant associations were observed for anterior OTS (aOTS) or MFS in either hemispheres, or for the right pOTS (see Figure 2C, all p-values > 0.58, all $\Delta R^2 < 0.7\%$).

Third, left pOTS interruption predicted T3 TOWRE composite scores regardless of whether it was measured through the longitudinal or cross-sectional processing pipeline of Freesurfer, at T1 or T3 (see Supplementary Figure 3A; T3 longitudinal: $F(1, 38) = 9.57$, $p = 3.7*10^{-3}$, $\Delta R^2 = 17.9\%$; T1 cross-sectional: $F(1, 40) = 10.96$, $p = 2.0*10^{-3}$, $\Delta R^2 = 19.6\%$; T3 cross-sectional: $F(1, 39) = 3.20$, $p = .08$, $\Delta R^2 = 6.62\%$), although the contribution was marginal at T3 with the cross-sectional pipeline. This importantly implies that left pOTS morphology at T1 was still a significant longitudinal predictor of T3 reading behavior when T3 brain images were not used to improve T1 anatomical segmentations, i.e., with T1 cross-sectional processing alone, despite noisier tissue segmentations with cross-sectional processing (see also Reuter et al., 2012).

Fourth, we tested whether interruption distance, as a continuous measure of sulcal interruption rather than the binary presence or absence of a gyral gap, was more predictive of reading skills. The total interruption distance between segments of the left pOTS was marginally predictive of future reading skills ($F(1, 38) = 3.54$, $p = .067$, $\Delta R^2 = 7.6\%$, see Supplementary Figure 3B). However, two additional analyses suggest that this association is likely due to trivial correlations between interruption distance and the binary presence of an interruption: (i) a similar analysis restricted to the subsample of children with interrupted left pOTS showed no association between reading and left pOTS total interruption distance ($F(1, 21) < 0.001$, $\Delta R^2 = 0.3\%$, $p = .98$); and (ii) interruption distance does not explain additional variance over and above the binary presence of an interruption across all children ($F(1, 37) = 0.04$, $p = .85$), while binary interruption significantly explains more variance in reading skills over and above interruption distance ($F(1, 37) = 6.99$, $p = .012$).

We further tested other anatomical features of the left pOTS at T1 and found that no other feature considered predicted reading skills, including total sulcal length, sulcal depth, depth of the sulcal pit, mean thickness, and total surface area (see Supplementary Figure 3B, all p-values > .26). In sum, the prospective relationship between left pOTS morphology at T1 and reading skills at T3 was specific to the binary presence of a gyral gap.

	Full longitudinal sample (N=43)	Initial minimal readers (N=24, subset)
Gender	18 Females / 25 Males	8 Females / 16 Males
Handedness	38 Right / 5 Left	23 Right / 1 Left
Family history of dyslexia	18 + / 25 -	8 + / 16 -
Age T1	5.49 years [5.31 - 5.86]	5.47 [5.19 - 5.71]
Age T2	7.13 years [6.87 - 7.50]	7.10 [6.79 - 7.46]
Age T3	8.18 years [8.00 - 8.57]	8.20 [8.04 - 8.54]
TOWRE reading scores T1		
SWE (raw score)	3.0 [0.0 - 9.0]	0.0 [0.0 - 2.3]
PDE (raw score)	2.0 [0.0 - 5.0]	0.0 [0.0 - 1.0]
TOWRE reading scores T2		
SWE (raw score)	49.0 [38.0 - 58.5]	44.0 [31.0 - 53.3]
PDE (raw score)	18.0 [13.0 - 24.5]	18.5 [13.0 - 24.0]
SWE (standard score)	109.0 [103.0 - 120.0]	107.5 [98.8 - 115.5]
PDE (standard score)	104.0 [99.0 - 112.0]	105.0 [99.0 - 111.5]
TOWRE composite (standard score)	108.0 [101.0 - 118.0]	108.0 [100.5 - 116.2]
TOWRE reading scores T3		
SWE (raw score)	61.0 [55 - 67.5]	58.5 [50.5 - 69.0]
PDE (raw score)	24.0 [20 - 35.5]	24.0 [17.0 - 36.5]
SWE (standard score)	110.0 [103.5 - 118.0]	106.5 [101.8 - 118.0]
PDE (standard score)	104.8 [98.0 - 115.0]	100.5 [96.0 - 116.2]
TOWRE composite (standard score)	108.9 [101.0 - 116.0]	104.5 [97.5 - 119.0]
Rapid Automatized Naming (standard score, average of color and object naming)		
T1	98.0 [89.5 - 107.5]	97.0 [88.9 - 103.1]
T2	97.0 [88.3 - 107.8]	97.3 [84.8 - 104.1]
T3	97.0 [84.5 - 105.3]	95.8 [84.5 - 104.0]
Phonological awareness (CTOPP) (standard composite score of elision & blending subtests)		
T1	112.0 [104.5 - 118.0]	106.0 [100.0 - 112.0]
T2	118.0 [107.5 - 131.5]	118.0 [111.2 - 133.8]
T3	119.8 [109.0 - 127.0]	118.0 [108.2 - 127.0]
Letter Knowledge T1 (WMRT) (standard score)	110.0 [103.5 - 114.5]	107.0 [102.8 - 110.2]
Performance IQ (WJC Concept Formation, standard score)		
T1	119.0 [106.0 - 127.0]	113.5 [103.0 - 123.5]
T3	117.0 [109.0 - 123.0]	117.0 [107.5 - 121.0]
Socio-economic status		
Average parental years of education	17 years [16 - 18]	16 years [15.5 - 17]
Yearly household income	140k$ [88.5 - 200] (*N=39)	150k$ [92.5 - 200] (*N=23)

Table 1. Sample demographics and main neuropsychological scores at the three time points T1-3

(median and inter-quartile range are indicated), for the N=43 children with longitudinal MRI processing reported in the main text.

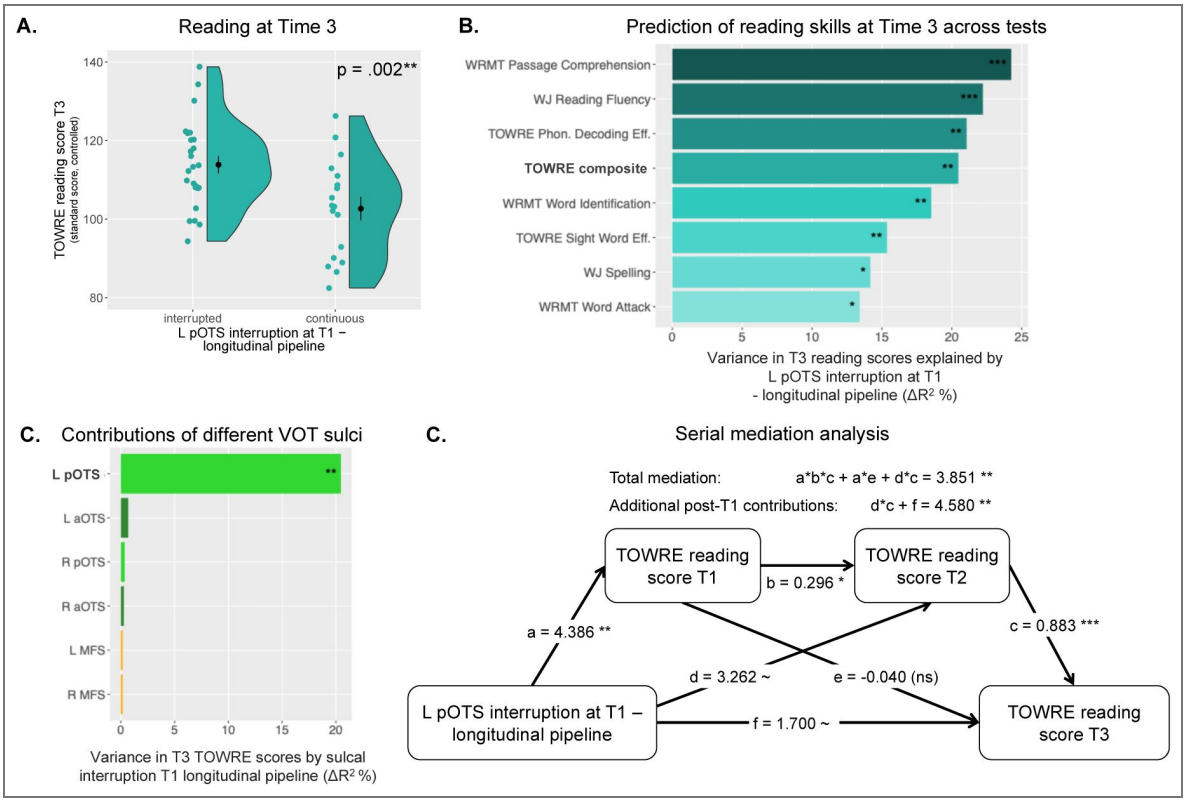


Figure 2. Cumulative impact of left pOTS interruption on reading scores longitudinally.

A. TOWRE composite standard scores at T3 are significantly higher in children with an interrupted than continuous left pOTS (assessed at T1 with longitudinal processing), after removing residual variance due to age, gender and handedness. Black dots represent mean scores in each group, ± standard error of the mean (s.e.m.). **B. Prediction of reading and spelling scores across tests:** percentage of variance in different reading and spelling standardized tests explained by left pOTS interruption (ΔR^2 in models controlling for demographic variables and total brain volume). Left pOTS interruption at T1 was significantly associated with all reading scores at T3, ranging from timed and untimed measures of single word and pseudo-word reading, as well as sentence and passage comprehension or spelling. The bolded TOWRE composite score corresponds to the effect depicted in panel a. **C. Percentages of variance in TOWRE T3 reading scores** explained by interruption by different sulci in VTC (posterior and anterior OTS (pOTS and aOTS, respectively) and MFS, in each hemisphere). **D. Serial mediation model demonstrating the cumulative impact of left pOTS interruption on TOWRE reading scores over time** from T1 to T2 to T3. Coefficients are standardized. Reading scores are controlled for age at testing, gender, and handedness, as well as left pOTS interruption by these demographic variables and total brain volume (***: p<0.001; **: p<.01, *: p<.05).

The effect of left pOTS interruption on reading accumulates over time

We next tested whether the effect of left pOTS interruption on reading scores observed at T3 was present at earlier time points and how it evolved over time. Compared with children with a continuous left pOTS, children with interrupted left pOTS also had higher TOWRE composite reading scores at T2, at around 7.1 years-old, and already at T1, at approximately 5.5 years-old (see Supplementary Figure 4B and C; T2 TOWRE composite scores: $F(1, 38) = 9.43$, $p = 3.9*10^{-3}$, $\Delta R^2 = 14.3\%$, Cohen's d = 0.872; population-normalized T1 TOWRE composite scores, in the absence of norms in this age-range: $F(1, 38) = 7.63$, $p = 8.8*10^{-3}$, $\Delta R^2 = 16.0\%$, Cohen's d = 0.919).

However, differences in emerging reading skills at T1 did not alone drive the later differences observed at T2 and T3. To test this, we conducted a supplementary analysis in the subset of 24 children with the lowest reading scores at T1 (defined as those with a TOWRE T1 population-derived composite score equivalent below 95). Hereafter, we refer to these children as *initial minimal readers*: 79% could not decode a single syllable or pseudoword in 45 s, with a maximum of four pseudowords decoded, and 63% could not recognize a single frequent sight word, with a maximum of seven words recognized. Although these children showed no difference in T1 reading scores as a function of whether their left pOTS was interrupted or continuous ($F(1, 19) = 0.21$, $p = .66$, $\Delta R^2 = 0.82\%$, Cohen's d = 0.034), there were growing statistically significant differences at T2 ($F(1, 19) = 5.20$, $p = .034$, $\Delta R^2 = 16.91\%$, Cohen's d = 0.965) and at T3 ($F(1, 19) = 6.93$, $p = .016$, $\Delta R^2 = 26.15\%$, Cohen's d = 1.107; see Supplementary Figure 4D-F).

A serial mediation analysis confirmed that the beneficial effect of left pOTS interruption on reading acquisition accumulated over time, from T1 to T2 and T3, and did not merely result from differences already present at T1 (see Figure 2C). Residuals of T1, T2 and T3 TOWRE reading scores, controlled for age, gender and handedness, and residuals of a logistic regression of left pOTS interruption, accounting for demographic variables and total brain volume, were entered into a serial mediation model. T1 and T2 reading scores were significant serial mediators of the effect of left pOTS interruption on T3 reading (total serial mediation a*b*c+a*e+d*c: estimated coefficient $\beta = 3.851$, standard error SE = 1.450, $p = .008$). Notably, left pOTS interruption at T1 predicted T1 TOWRE score (path a: $\beta = 4.386$, SE=1.560, $p = .004$), T1 TOWRE predicted T2 TOWRE (b: $\beta = 0.296$, SE = 0.149, $p = .049$) and T2 TOWRE predicted in turn T3 TOWRE (c: $\beta = 0.883$, SE = 0.104, $p < .001$, see Figure 2C). Importantly, there were also significant contributions of left pOTS interruption onto T3 reading on top of those mediated through T1, with large cumulative effects through T2 and directly onto T3 (d*c+f: $\beta = 4.580$, SE = 1.757, $p = .009$), while indirect effect through T1 TOWRE scores alone were non-significant (a*b*c+a*e: $\beta = 0.971$, SE = 0.839, $p = .25$). Left pOTS interruption had marginal direct effects both onto T2 TOWRE scores (d: $\beta = 3.262$, SE = 1.673, $p = .053$) and onto T3 TOWRE (f: $\beta = 1.700$, SE = 0.983, $p < .09$). Interestingly, the magnitude of direct contributions of left pOTS interruption onto reading scores seemed to decrease over time (a: $\beta = 4.386$, d: $\beta = 3.262$, f: $\beta = 1.700$). Overall, left pOTS interruption therefore appeared as a continued constraint on reading acquisition between 5 and 8 years of age.

Left pOTS interruption makes a strong longitudinal contribution relative to established predictors of reading

Having observed that left pOTS interruption before or at the onset of literacy acquisition is a strong predictor of future reading skills, we next asked how predictive this neuroanatomical feature was relative to benchmark cognitive predictors of reading skills, and how these variables related to one another. Specifically, we asked whether left pOTS interruption was a predictor of reading independent from known pre-literacy skills, or whether an association with a specific pre-literacy skill before learning to read may trivially explain its predictive power. We considered constructs shown to predict reading acquisition or risk of dyslexia longitudinally, as measured by classical neuropsychological tests: rapid naming (of colors and objects, as categories independent from early reading skills), phonological awareness, performance IQ, and receptive vocabulary. Letter knowledge and TOWRE scores at T1 were also considered, as emerging reading skills.

Family history of dyslexia was also taken into consideration, and indexed by mother scores in the Adult Reading History Questionnaire (ARHQ, Lefly & Pennington, 2000, see Methods for reasons for selecting mother ARHQ).

Relative to all these predictors, left pOTS interruption was the strongest longitudinal predictor of reading, with its 20.5% of variance in T3 TOWRE explained after controlling for demographic variables (ΔR^2 = 20.5%, F(1, 38) = 11.34, p = $1.7*10^{-3}$), followed by letter knowledge with 19.6% of variance (letter identification, F(1, 38) = 10.75, p = $2.2*10^{-3}$) and phonological awareness with 16.2% (CTOPP, F(1, 38) = 8.51, p = $5.9*10^{-3}$; see Figure 3A). Rapid naming (RAN objects and colors), early reading scores (TOWRE T1), family history of dyslexia (mother ARHQ) and receptive vocabulary (PPVT) were also significant but weaker predictors of T3 reading skills, with separate respective contributions to T3 TOWRE scores of 13.6%, 11.5%, 11.2% and 9.5% of variance ($p \in [0.01, 0.04]$). In the subsample of children with minimal reading skills at T1, left pOTS interruption was the second strongest longitudinal predictor of reading explaining 26.1% of variance in T3 reading (F(1,19) = 6.93, p = .02), close after rapid naming (ΔR^2 = 28.0%, F(1, 19) = 7.61, p = .01, see Figure 3D). Dyslexia family history was the third and only other significant predictor in this subsample (ΔR^2 = 21.3%, F(1, 19) = 5.29, p = .03; all other ps > .09).

Concerning the extent to which left pOTS interruption is related at T1 to other longitudinal predictors of reading, we observed significant associations between sulcal interruption and early reading skills as previously reported, with left pOTS interruption accounting for ΔR^2 = 16.0% of variance in T1 TOWRE scores already above demographic variables and total brain volume (F(1, 38) = 7.63, p = $8.8*10^{-3}$; see Figure 3B), and for 11.6% of letter knowledge, which can be considered an early literacy measure (F(1, 38) = 6.09, p = .018). We also observed significant sulcal contributions for explaining performance IQ (concept formation: ΔR^2 = 14.3%, F(1, 38) = 6.78, p = .013), but not for other variables (p>.08). The same analysis repeated in the subsample of 24 initial minimal readers at T1, i.e. removing confounds with early reading skills, still showed an association between left pOTS interruption and performance IQ (ΔR^2 = 14.2%, F(1, 19) = 3.42, p = .0797, of marginal significance likely due to the reduced sample size given the effect size similar to that observed in the full sample), but no significant association with other pre-literacy skills (see Figure 3D). Thus, we observed an association between left pOTS interruption and performance IQ at T1, apparently independent from early reading skills, while the association with letter knowledge and TOWRE scores at T1 may only reflect early effects of left pOTS interruption on emerging reading skills observable in the full sample only.

These significant associations between left pOTS interruption, performance IQ and pre-literacy skills (Figure 3B and D), as well as potential associations between pre-literacy skills themselves (cf. correlation matrices in Supplementary Figure 5D and H), make it challenging to fairly assess the separate unique contributions of each of these variables to the longitudinal prediction of reading skills. For instance, because left pOTS interruption explains a significant amount of variance at T1 in letter knowledge, early TOWRE scores and performance IQ, some of its true direct effect on future T3 reading skills is already captured in a baseline model including these cognitive predictors only, even before left pOTS interruption is explicitly entered as a predictor. Yet, even when controlling in a such a stringent way for all pre-literacy skills, early reading abilities, family history of dyslexia and their shared variance with left pOTS interruption in such nested model comparison, left pOTS interruption still had a marginal additional contribution to predicting T3 TOWRE reading scores (ΔR^2 = 5.0%, F(1, 31) = 3.74, p = .062) in spite of this important collinearity. Rapid naming and family history were the only two other predictors that outperformed left pOTS interruption in terms of unique contributions to T3 reading (ΔR^2 = 7.7% and 6.6% for RAN and mother ARHQ respectively, see Figure 3C), possibly because these predictors were among the least confounded with others (see Figure S5D for correlations among predictors) and thus among the least affected by collinearity. Letter knowledge, in contrast, had the smallest, non-significant unique contributions to T3 reading (ΔR^2 = 0.61%, p = .50). In order to reduce the collinearity confounds resulting from the associations between pOTS interruption and T1 early reading skills, we also reran the same analysis in the subsample of initial minimal readers. Left pOTS interruption in this subsample significantly contributed to predicting future T3

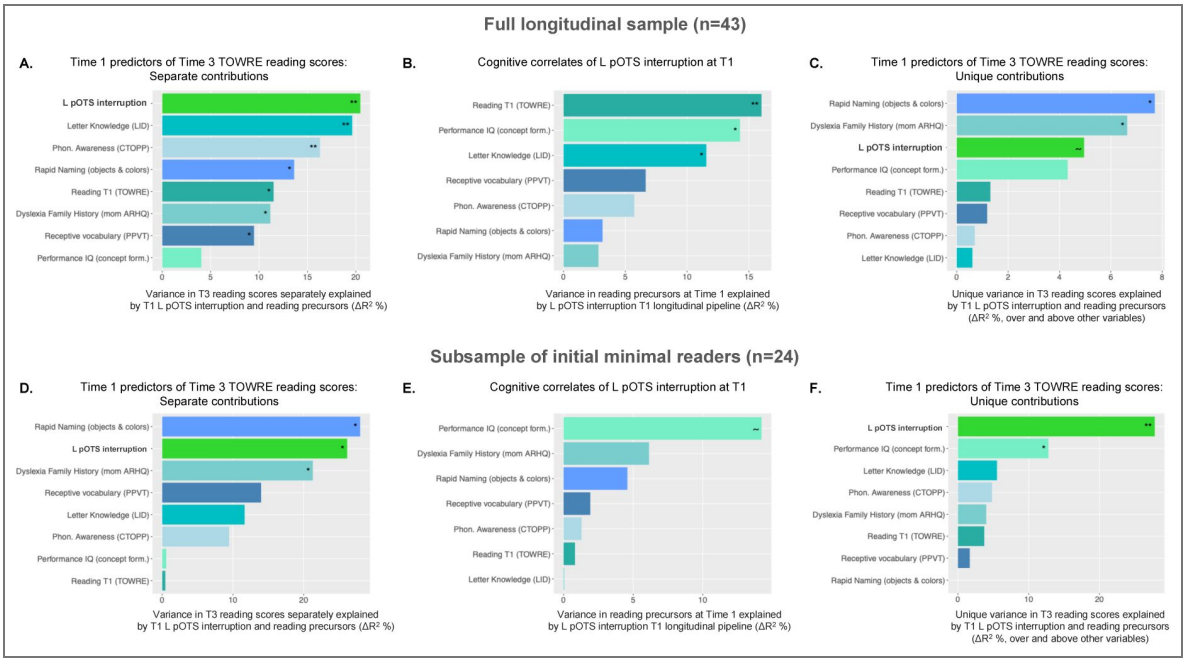


Figure 3. Longitudinal contributions of left pOTS interruption to reading relative to cognitive precursors of reading, as well as their associations,

in the full longitudinal sample (A-C, top row) and in the sub-sample of initial minimal readers (D-F, bottom row), respectively. A & D. Variance in TOWRE reading scores at Time 3 explained by left pOTS interruption and classical predictors of reading at Time 1 separately, over and above demographic variables, ranked by decreasing contributions (ΔR^2 in percentage, relative to baseline reference models with only demographic variables accounting respectively for a total R^2 of 11.0 and 2.2% of variance in T3 reading). B & E. Associations between left pOTS interruption and predictors of reading at Time 1: percentage of variance (ΔR^2) in left pOTS interruption explained by each pre-literacy skill separately, over and above demographic variables and total brain volume. C & F. Unique variance in TOWRE T3 reading scores explained by left pOTS interruption and classical predictors of reading at Time 1, over and above other variables of interest and demographic variables, ranked by decreasing contributions (ΔR^2 in percentage). For reference, the full models including all considered predictors accounted respectively for a total R^2 of 58.9 and 75.8% of variance in T3 reading. (Stars on each bar denote significant contributions as assessed by nested model comparison using an F-test. **: p<0.01, *: p<.05)

reading over and above all pre-literacy skills and family history (ΔR^2 = 27.8%, F(1, 12) = 13.80, p = .003), making it the first independent longitudinal contributor of reading (see Figure 3F). The second largest longitudinal predictor in terms of unique contributions was performance IQ (ΔR^2 = 12.8%, F(1, 12) = 6.34, p = .027). No other predictor had a significant unique contribution to the longitudinal prediction of T3 reading in the subset initial minimal readers, including rapid naming even though its separate contribution was larger than that of pOTS.

In summary, after taking into account its collinearity with early T1 literacy and reducing it in the subsample of minimal initial readers, left pOTS interruption at around 5 years of age emerged as a particularly strong independent longitudinal contributor to reading skills at around 8 years, with predictive performance comparable to or greater than that of classical cognitive and familial predictors of reading ability and disability (dyslexia) that are more proximal to reading behavior.

Left pOTS interruption has a relatively specific cognitive impact on reading abilities

The incidental observation of a significant association between left pOTS interruption and performance IQ at T1, although smaller in magnitude and not sufficient to account for the association with literacy skills, raises the possibility that left pOTS interruption may confer a broader cognitive advantage rather than a specific benefit for reading acquisition. We therefore conducted a supplementary analysis to examine the relative cognitive specificity versus domain generality of this association. Of note, reading and general cognitive skills are strongly correlated and mutually influence each other across development (Ferrer et al., 2007; Peng & Kievit, 2020; Ünal et al., 2023), making this question of cognitive specificity inherently difficult to address.

We gathered all cognitive and linguistic scores obtained at T3, including measures of mathematics, general cognition, executive functioning, language, and reading (see Table S2 for a full list and description of the included subtests), and tested how each measure was associated with left pOTS interruption, residualized for demographic variables and total brain volume. Left pOTS interruption showed strong and significant associations with several measures of language, reading, mathematics, and general cognitive ability (see Supplementary Figure 6). As expected, we also observed strong and significant correlations among the behavioral measures themselves.

However, the strongest correlations with left pOTS interruption residuals were observed for reading scores. Moreover, hierarchical clustering grouped left pOTS interruption with a sub-cluster that included most reading assessments - spanning phonological decoding, word reading, and sentence reading fluency -, as well as spelling and more surprisingly an expressive language index. This reading-related cluster was most closely related to another cluster comprising measures of written and oral language comprehension (i.e., receptive language), followed by a cluster of phonological awareness. In contrast, measures of mathematics, general cognitive ability, backward working memory and rapid naming formed a separate cluster. Thus, although left pOTS interruption is associated with several cognitive, linguistic, and academic measures, its pattern of associations is most closely aligned with children's reading outcomes, supporting a relatively specific cognitive advantage for reading.

Why are interruptions beneficial? Left pOTS interruption and local white-matter connectivity advantages

The underlying mechanisms contributing to why the interruption in the cortical folds around the left pOTS region may so strongly and reliably favor adequate reading acquisition remains unknown. Previous work investigating the relationship between sulcal morphology and behavior have proposed a mediating role of the underlying white matter (Amiez et al., 2018; Borst et al., 2016; Cachia et al., 2018; Garrison et al., 2015; Miller et al., 2021; Voorhies et al., 2021), although these potential mechanisms have not been explicitly explored.

In the case of sulcal interruptions, two explanations stand out as potential mechanisms underlying the behavioral benefits for reading of left pOTS interruptions. First, given that white-matter fibers in general, and especially long-distance projection fibers, tend to terminate more in gyri than sulci (Dannhoff et al., 2023; Ge et al., 2018), having an extra gyral crown, or superficial *pli de passage* (Mangin et al., 2019), in the left pOTS may allow to pack more connections, especially long-distance connections through e.g., potentially a combination of fibers from the (posterior) arcuate fasciculus, vertical occipital fasciculus and inferior longitudinal fasciculus (Bouhali et al., 2014; Grotheer et al., 2021; Moulton et al., 2019; Weiner et al., 2017; Yeatman et al., 2012). Such increased long-distance connections may support a more efficient reading network. Alternatively, the presence of a *pli de passage* in the left pOTS may allow for better local white-matter connectivity through short-range U-shaped fibers traveling under the cortical gray matter. Indeed, Bodin and colleagues (2021) showed that, in the superior temporal sulcus (STS), the presence of *plis de passage* was associated with a higher density of U-shaped fibers connecting adjacent gyri. Superficial *plis de passage*, i.e., complete sulcal interruptions with a gyral gap visible on the cortical surface, showed an even higher local connectivity advantage than deep *plis de passage*, the complete gyral gaps or partial interruptions with sulcal pinching which are buried within the depth of the sulcus. (Please note that in our study, we focus on superficial gyral gaps visible on the cortical surface only, and do not classify deep *plis de passage*.)

In order to test these hypotheses, we leveraged diffusion-weighted data available from Time 3 (n=29 children with data of sufficient quality, see Methods; too few children having usable diffusion data at T1). Gyral gaps between left pOTS sulcal components were manually delineated at the level of left pOTS interruptions (see Figure 4A, and Supplementary Figure 7) and we assessed white-matter properties and connectivity around left pOTS gaps vs. left pOTS sulcal components (what we refer to as pOTS proper) in the same individuals (two leftmost violin plots in Figure 4B, top) and then also between individuals (individuals with uninterrupted pOTS and pOTS proper in individuals with an interrupted pOTS; right two violin plots in Figure 4B, top).

According to hypothesis (i), left pOTS gaps should show higher myelination, both in the cortex and the white-matter in its vicinity, than surrounding left pOTS components in the same individuals. Diffusion tensor models were therefore fitted to derive indices related to myelin content. Average Mean Diffusivity (MD) was first estimated in the cortical ribbon as the most robust DTI index reflective of myelination and neurite properties in the cortex (Fukutomi et al., 2019). In children with interruptions (n=15), MD in left pOTS gaps was higher than in the cortex of the left pOTS proper (paired two-sample T-test $t(14) = 9.74$, $p = 1.3*10^{-7}$; see Figure 4B, top row), while MD in the left pOTS proper of these children was similar to MD in the continuous left pOTS of the remaining children (n=14, Welch two-sample T-test: $t(25.6) = -0.39$, $p = .70$; Figure 4B top row, right handside). These results thus suggest lower myelination around left pOTS gaps, instead of the higher myelination predicted by hypothesis (i). This led to higher MD in the overall left pOTS in children with interruption - considering the cortex of the left pOTS proper and gyral gap(s) - than in children with continuous sulci ($t(26.4) = -2.16$, $p = .040$, Figure 4B bottom row). Additional analyses suggested that such differences were not driven by partial volume effects due to the presence of more or less surrounding white matter tissue or cerebro-spinal fluid close to gyral or sulcal cortex (see Supplementary DTI analyses in Appendix A).

The pattern of results observed for mean MD in the cortex was moreover confirmed by supplementary analyses of Radial Diffusivity (RD) and Fractional Anisotropy (FA) in the cortex around the left pOTS, and of MD, RD, and FA in the superficial white-matter located above the cortex (FA and RD being relevant markers of myelin content in white matter; see Dubois et al., 2014; Lazari & Lipp, 2021), suggesting lower myelination both in and above gyral gaps (Supplementary Figure 8 and Supplementary Table 3). In short, against hypothesis (i), the present data do not support the idea that left pOTS interruptions are associated with more terminations of long-distance projection fibers.

Next, we tested hypothesis (ii) according to which left pOTS gaps would be associated with increased local connectivity surrounding the left pOTS, through the packing of more U-shaped fibers traveling superficially. In order to compare the structural connectivity around the left pOTS

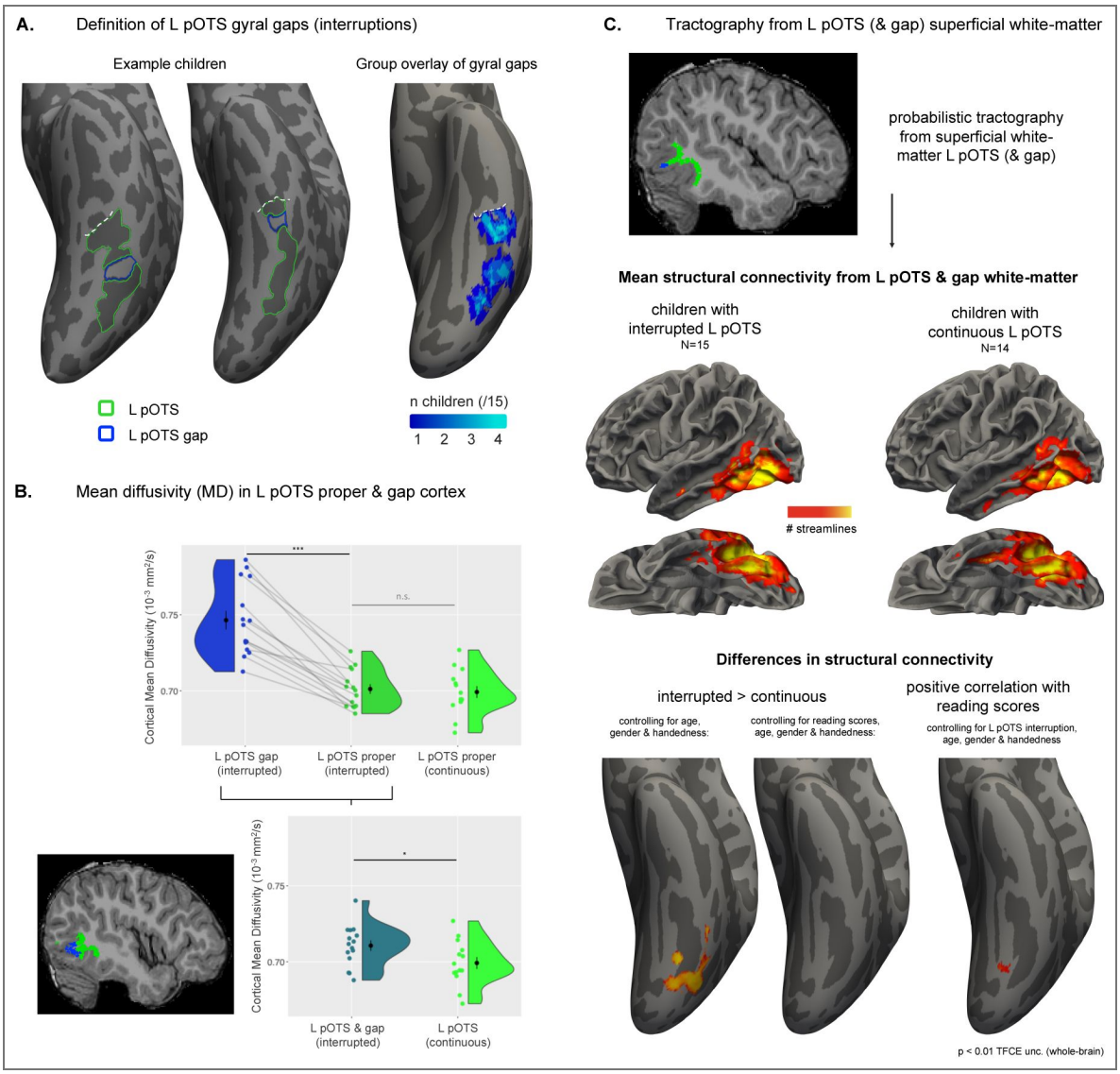


Figure 4. White-matter correlates of left pOTS interruptions.

A. Gyral gaps in the left pOTS were delineated manually (in blue) and contrasted to left pOTS proper (green). Two exemplar left hemispheres are shown in children with a left pOTS interruption. The group-level map shows overlays of left pOTS gaps (projected onto the *fsaverage* surface) of 15 children with interrupted left pOTS, out of 29 with usable diffusion data. **B.** Average MD in the cortex of the left pOTS proper and gyral gaps. Top row (from left to right): difference in cortical MD in the left pOTS gap (blue) and left pOTS proper (dark green) in children with an interrupted sulcus, and in left pOTS in children with continuous sulci for comparison (light green). Bottom row, left: example cortical regions in which MD was average in a child with a sulcal interruption. Bottom row, right: MD comparison between children with interrupted and continuous sulci over the left pOTS cortex, including left pOTS gap if applicable. **C.** Probabilistic tractography from superficial white-matter seeds above the left pOTS cortex and possible left pOTS gap. From top to bottom: example seed in a participant (green: left pOTS proper, blue: left pOTS gap); mean whole-brain connectivity in children with interrupted (n=15) and continuous left pOTS (n=14); statistical differences in connectivity (from left to right): between the two groups not taking into account reading scores, between the two groups when modeling reading scores (no association observed), between connectivity and T3 TOWRE reading scores ($p<0.01$ TFCE uncorrected threshold).

proper and gaps in children with a continuous vs. an interrupted sulcus, we performed whole-brain probabilistic tractography from a superficial white-matter layer located under the left pOTS and gap cortex (see Figure 4C for a seed region in an example participant). Mean structural connectivity from these seeds from all children showed strong local connectivity to fusiform and inferior temporal gyri adjacent to the superficial white-matter seeds above the left pOTS, and extending to the middle temporal gyrus, parahippocampal cortex, and anterior OTS (Figure 4C middle panel), consistent with the presence of short, neighborhood association fibers traveling under the sulcal cortex (Dannhoff et al., 2023; Reveley et al., 2015). When comparing structural connectivity between children with interrupted and continuous left pOTS through permutation testing, children with sulcal interruption showed higher local connectivity, as indexed by density maps reflecting the number of reconstructed streamlines, in a cluster located around the posterior section of the left pOTS, after controlling for demographic variables (age, gender and handedness, $p<0.01$ threshold-free cluster enhancement -TFCE-uncorrected, Figure 4C bottom panel, leftmost map), suggesting a higher development of short-distance connections, in agreement with hypothesis (ii).

In order to assess whether this higher local connectivity in children with left pOTS gaps may also directly participate in their reading benefit, a second permutation analysis was run modeling the presence of left pOTS interruption, T3 TOWRE reading scores, and demographic covariates. In this analysis, no statistical difference was observed between children with and without interruptions (at $p<0.01$ TFCE), but connectivity and reading scores positively correlated in a white matter cluster that partially overlapped with the connectivity cluster from the previous analysis (Figure 4C bottom panel, middle and rightmost maps). Despite the reduced statistical power, this suggests that advantages in local structural connectivity around the posterior left pOTS in children with sulcal interruption may be associated with better reading scores and constitutes a candidate pathway subtending the behavioral reading advantage.

Discussion

To the best of our knowledge, the current study provides empirical evidence for the first time supporting that the presence of a gyral gap of the left pOTS before, or at, the onset of literacy instruction robustly and uniquely predicts better future reading skills throughout the first three years of elementary school. Importantly, this effect was observed prospectively and was shown to accumulate over time. It was specific to the left posterior section of the OTS and to the presence of interruptions rather than any other anatomical feature of sulcus. This marker of individual sulcal anatomy proved to be a large and independent contributor to shaping children's reading abilities, with predictive performance comparable to or greater than that of well-established familial and pre-literacy cognitive measures. Cognitive benefits seemed relatively specific to reading, although significant associations were also observed for language, math and general cognitive skills intercorrelated with reading. Increased local U-fiber connectivity around the left pOTS associated with the presence of a gyral gap appeared as a likely mechanism underlying the behavioral benefits of such gaps on reading development. Below, we discuss these results in the context of (i) robust and specific longitudinal associations between left pOTS interruptions and reading behavior, (ii) precision approaches to the study of the individual differences in cognition, (iii) developmental origins of the impact of left pOTS interruption on reading, and (iv) common processes by which cortical constraints may influence cultural acquisitions.

Robust and relatively specific longitudinal associations between left pOTS interruptions and reading behavior

Previous findings demonstrated a cross-sectional association between left pOTS interruptions and reading ability in children (Borst et al., 2016) and adults (Cachia et al., 2018). Building on this work, the present findings showed a large positive contribution of left pOTS interruption to reading abilities in a population of beginning readers, and crucially, demonstrate its longitudinal validity: sulcal interruptions at the onset of literacy instruction in 5-6 year-olds predicted children's future reading skills at age 8.

Importantly, left pOTS interruptions were robustly associated with all aspects of reading tested at 8 years, ranging from single-word and pseudo-word naming (in both accuracy and speed) to sentence reading fluency and silent passage comprehension, as well as even spelling skills. Interestingly, the association between sulcal anatomy and reading was largest for passage comprehension - the most demanding and advanced aspect of reading.

Additionally, brain-behavior associations also extended to the most basic aspects of reading, e.g., letter recognition and orthographic processing, as demonstrated by significant associations between left pOTS interruption and letter knowledge already at age 5, on top of associations with timed word and pseudo-word reading (Figure 3B). This *general* effect on reading skills thus extends beyond previous reports, established on reading measures mostly tapping into phonological decoding skills (Borst et al., 2016; Cachia et al., 2018). Moreover, left pOTS morphology may have an effect on reading development across alphabets throughout a wide range of orthographic depths, as we replicate in English prior results in French and Portuguese (Borst et al., 2016; Cachia et al., 2018) - two scripts much more transparent in terms of grapheme-to-phoneme conversion.

Our study further extends previous results by Borst, Cachia and colleagues (2016; 2018) to a population of children with above-average cognitive and reading scores, from relatively high-SES families, and enriched for family history of dyslexia (cf. Table 1) - although no child met criteria for dyslexia by T3. While our sample is not representative of the general population, thereby limiting the direct generalizability of our results, half of the adult sample from Cachia et al. (2018) came from low socio-economic and rural backgrounds (matched to a sample of illiterate adults), suggesting that the association between left pOTS interruption and reading skills is robust across socio-economic contexts. Since our reviewed preprint came out, the association between left pOTS interruption and reading was further replicated (Cachia et al., 2025; Yao et al., 2026), notably in a larger child sample along the full reading continuum and oversampling children with dyslexia (Yao et al., 2026). (Of note, the finding that pOTS morphology correlated in adults with reading fluency rather than accuracy (Cachia et al., 2025) may also explain one of two prior negative findings concerning the association between pOTS and reading (Del Mauro et al., 2022)). Thus, our results together with the extant literature overall underline a common effect of left pOTS interruption observed across the continuum of reading abilities, and possibly of cognitive abilities.

Interestingly, left pOTS interruption also showed associations with other cognitive skills: it was associated with performance IQ at age 5 (see Figure 3B and E), and with performance IQ, math and language skills more broadly at age 8 (Supplementary Figure 6). However, left pOTS interruption was most strongly associated and clustered with assessments of reading at age 8 (as well as spelling and expressive language). In addition, the prospective association between left pOTS interruption and reading skills was not merely driven by an association with general cognitive abilities at age 5 (Figure 3C and F) and no association with any other pre-literacy skill was found (Figure 3B and E). Math, reading, language and cognitive skills are intrinsically highly intercorrelated, as we also observed at age 8, and are known to mutually and positively influence each other over time (Ünal et al., 2023; Peng & Kievit, 2020; Ferrer et al., 2007; Huettig & Hulstijn, 2025; Dehaene et al., 2015). The weaker pOTS cognitive associations may therefore result from the mere combination of a specific pOTS influence reading and of

associations between reading and other cognitive domains, although more research is needed to better understand these associations and how they unfold during development. Together, these findings highlight a relative degree of specificity of left pOTS influences to reading abilities, in line with the well-documented functional specificity of the left pOTS for orthographic processing and reading (e.g., Dehaene & Cohen, 2011; Gaillard et al., 2006; Pillet et al., 2024; Wilson et al., 2013; Woolnough & Tandon, 2024; Zemmoura et al., 2015; Zhan et al., 2023).

Precision approaches to the study of individual differences in cognition

Results of the present study add to a growing body of literature indicating that inter-individual differences in the morphology of gyral gaps in association cortices are associated with various aspects of cognition such as numerical processing and reading skills (Borst et al., 2016; Cachia et al., 2018, 2021; Roell et al., 2021; Santacroce et al., 2024; Schwizer Ashkenazi et al., 2024; Tissier et al., 2018). The present and previous findings relating gyral gaps to individual differences in cognitive performance require manual definitions of gyral gaps in a relatively small sample of individuals, consistent with a "precision imaging" approach (Cachia et al., 2021; Gratton et al., 2022; Rosenberg & Finn, 2022). While brain-behavior relationships, typically studied through automatically-extracted estimates such as cortical thickness or surface area, have been found to have small effect sizes and to require large N approaches (Marek et al., 2022), macro-anatomical features of neuroanatomy may be expected to have a larger impact and require smaller samples. Moving forward, however, future studies can integrate these two approaches by building tools to precisely measure gyral gaps in an automated fashion, which would result in both a large N, as well as precisely identified gyral gaps, in each individual using deep learning (Lyu et al., 2021; Rivière et al., 2022). Such large-N approaches are needed to answer a number of outstanding questions, such as, in our case, whether there is an additional behavioral reading benefit to having two or more left pOTS interruptions - which were too infrequent in our sample -, or whether gyral gaps placed in specific locations along the left pOTS may be optimal to facilitate reading acquisition.

Developmental origins of the impact of left pOTS interruption on reading

Developmentally, the OTS starts to appear around 25-30 weeks of gestation (Chi et al., 1977; Garel et al., 2001; Hansen et al., 1993). Although the relative developmental timeline between the appearance of a sulcus and that of its interruptions remains unclear, interruption(s) of the OTS are thus thought to be stable features of neuroanatomy determined early on. In our study, we observed high reliability of the pOTS interruption from age 5 (T1) to age 8 (T3), which is consistent with the longitudinal stability of interruptions of (i) the intraparietal sulcus and postcentral sulcus (Schwizer Ashkenazi et al., 2024), (ii) sulci in inferior frontal cortex (Tissier et al., 2018), and (iii) patterns of the anterior cingulate sulcus (Cachia et al., 2016) observed in late childhood (approximately 8 years onwards) to early adulthood. Of course, genetic and environmental factors also likely contribute to the brain-behavior relationship identified in the present study. For example, a recent report has reported a marginal association in the interruption of the left OTS among mother-child dyads (Fehlbaum et al., 2022), suggesting potential contributions of genetic factors or of the familial environment in shaping left OTS morphology. Further, gyrification is significantly impacted by the prenatal environment (Amiez et al., 2018) and gestational age at birth (Dubois et al., 2019), notably in the case of regions surrounding the left pOTS (Papini et al., 2020; Schmitt et al., 2023). Multiple (e.g., twin) vs. single pregnancy and intrauterine growth restriction may also influence such gyrification patterns (Amiez et al., 2018; Dubois et al., 2008), each of which can be considered in future studies.

Moreover, in the present study, effects of left pOTS interruptions on reading skills accumulated and grew, as shown by significant serial mediation effects from T1 to T2 to T3 (Figure 2D). As such, left pOTS interruptions seem to have a continued beneficial influence on developing reading skills throughout the first few elementary grades. Whether such benefits to reading development

continue to accumulate beyond 8 or 9 years of age remains to be explored, but these benefits seem to at least be retained later on, as suggested by reports in older children and adults (Borst et al., 2016; Cachia et al., 2018, 2025; Yao et al., 2026).

Local connectivity differences at gyral gaps: a computational advantage?

The present findings empirically support that, despite variability in the location of individual gaps along the pOTS, children with at least one gyral gap have increased local connectivity in a posterior cluster of the left pOTS, compared to those with a continuous pOTS. Such increased connectivity could have a computational advantage. For instance, previous work has highlighted that junctions in gyral ridges, known as 3-hinge gyri, such as those which surround gyral gaps, serve as hubs in cortico-cortical connections with higher connection degrees, stronger connections, and high betweenness centrality in graphs of structural connectivity (Zhang et al., 2020). The rich connectivity of such regions has also been documented through fine-grained post-mortem dissections (Shinohara et al., 2020; as also reviewed in Dannhoff et al., 2023), revealing that '*pyramid shape crossings*' are characterized by, and seem to result from, the convergence of both intra- and inter-gyral U fibers. Crucially, we showed that such increased local U-shaped fibers also relate to reading skills.

Building on these results, mechanistically, computational simulations of white-matter development and gyrification suggest that gyral regions with a high concentration of growing axonal fibers tend to form 3-hinge gyri (Ge et al., 2018; see also Chavoshnejad et al., 2021). Therefore, pOTS gyral gaps may be intrinsically related to better U-fiber connectivity at the level where such gaps meet the adjacent fusiform and inferior temporal gyral crests. Such junctions may serve as connectivity hubs for processing written words, and such hubs may result in a computational advantage related to skilled reading, which can be tested in future research.

Common cortical constraints on cultural learning

Interestingly, associations between sulcal interruptions and behavior have also been reported for symbolic, but not non-symbolic, number processing (Roell et al., 2021; Schwizer Ashkenazi et al., 2024), highlighting common processes by which cultural acquisitions, too recent on the scale of human evolution to rely on dedicated neural circuits (Dehaene & Cohen, 2007), may be boosted by the presence pre-existing gyral gaps. In the case of the left pOTS, various other aspects of early brain structure have been proposed to constrain the development of reading skills and of functional activations for reading in the VWFA and broader neural network involved in reading such as anatomical connectivity (Bouhali et al., 2014; Grotheer et al., 2021; Li et al., 2020; Moulton et al., 2019; Saygin et al., 2016) and cortical morphology. In terms of the latter, longitudinal studies following pre-reading children at familial risk of dyslexia showed for instance that the surface area of the left fusiform gyrus (Beelen et al., 2019) and patterns of gray-matter volume around the VWFA (Skeide et al., 2016) could be predictive of future dyslexia diagnosis.

Conclusion

Altogether, reading is a cultural invention too evolutionarily recent to have shaped brain structure. Instead, learning to read tunes pre-existing ventral occipito-temporal territories for written language (Dehaene & Cohen, 2007; Dehaene-Lambertz et al., 2018; Kubota et al., 2023). Thus, stable properties of children's macro-anatomical architecture, such as the presence of left pOTS gyral gaps and the associated increase in white-matter connectivity, may guide the neural plasticity during such a critical cultural learning experience, impacting learning outcomes. Our present findings support the importance of cortical constraints on literacy learning and link longitudinal measures of local cortical folding, white matter properties, and reading ability with a precision imaging approach for the first time.

Materials and Methods

Participants

51 children were followed longitudinally from kindergarten to third grade, as part of an NIH-funded project (K23HD054720) focusing on children's reading development. Specifically, children were assessed at three time points: T1 at a median age of 5.5 years, T2 at a median age of 7.1 years, and T3 around 8.2 years. Neuropsychological assessments, demographic and family information were collected at all time points, while only T1 and T3 included MRI acquisitions (see Figure 1A). Eight children were excluded from the main analyses assessing longitudinal associations between OTS and MFS morphology and reading acquisition: 4 children due to drop-out (did not come back for T3 visit), 1 due to neuroanatomical findings (large bilateral cysts in the anterior temporal lobe) and 3 due to excessive motion in the MRI at either T1 (2 children) or T3 (1 child). The resulting main sample consisted of N=43 children (58% males; 88% right-handed; see Table 1 for specific demographics). Children came from high socio-economic backgrounds (e.g., median average parental years of formal education: 17 years, equivalent to a master's degree) and had above-average cognitive abilities and reading skills, despite 42% of children having family history of dyslexia (e.g., a parent or sibling diagnosed with dyslexia, see Table 1). No child had received a diagnosis of dyslexia by T3, nor did any child qualify for such a diagnosis based on the in-depth neuropsychological assessments conducted in the study.

The Institutional Review Boards of Stanford University, where data were collected, and of the University of California San Francisco, where data were analyzed due to transition of the principal investigator FH, approved the present study. Informed assent and consent were obtained from children and their guardians, respectively. The dataset included in this study has resulted in several prior publications on other aspects of the data (Black et al., 2012; Hosseini et al., 2013; Myers et al., 2014; Vandermosten et al., 2020; Xia et al., 2021; Yamagata et al., 2016).

Behavioral tests

Children were tested at T1, T2, and T3 with an extensive battery of standardized neuropsychological assessments concerning literacy and pre-literacy skills. Reading skills were assessed at all time points with the Sight Word Efficiency (SWE) and Phonemic Decoding Efficiency (PDE) subtests from the Test of Word Reading Efficiency (TOWRE first edition, Torgesen et al., 1999). In these tests, children had to correctly read aloud as many items as possible in 45 seconds from lists of increasing difficulty of high frequency words, often with irregular grapheme-phoneme mappings (sight words), and of pseudo-words (phonemic decoding). We used TOWRE scores as the primary measure of reading scores in the current study as this speeded reading score was most similar to speeded reading tests that had previously shown associations with the sulcation of the OTS (Borst et al., 2016; Cachia et al., 2018). At T1, letter knowledge was assessed through the Letter Identification subtest of the Woodcock Reading Mastery Test R/NU (WRMT-R/NU, Woodcock, 1998). At T2 and T3, reading skills were also assessed with the Word Identification and Word Attack subtests from WRMT-R/NU, which respectively consist in untimed reading of low frequency words and of pseudo-words in isolation, with increasing difficulty. Additionally, the WRMT Passage Comprehension subtest required children to appropriately complete a sentence or a short passage choosing a word among multiple options (pictures are included for a third of the easiest items). Reading Fluency and Spelling subtests from the Woodcock–Johnson III Tests of Achievement (Woodcock, McGrew, et al., 2001) were also administered. In the Reading Fluency subtest, participants had to read a series of sentences and indicate for each whether it was true or false, within a 3-min time constraint. In the spelling subtest, children had to write down and correctly spell dictated words of increasing difficulty.

Children's cognitive and pre-literacy skills were also assessed, notably with the following tests: The Concept Formation subtest of the WJ-III Tests of Cognitive Abilities (Woodcock, Mather, et al., 2001) was used at T1 and T3 as an estimate of general cognitive abilities, and more specifically

performance intelligence quotient (pIQ). This learning task captures fluid reasoning by measuring children's ability to identify and state the rule for why some colored geometric figures are in a box whereas others are not (corrective feedback and reinforcement of correct answers are provided).

Receptive vocabulary was measured with the Peabody Picture Vocabulary Test (fourth edition, Dunn & Dunn, 2007). Blending and Elision subtests from the Comprehensive Test of Phonological Processing (CTOPP first edition, Wagner et al., 1999) were used to derive a composite measure of phonological awareness. Rapid Automatized Naming (RAN) abilities were measured by averaging scores on the object and color subtests (Wolf & Denckla, 2005), to provide an estimate of rapid naming abilities independent from letter or symbol reading. Many other cognitive, academic and linguistic abilities were further assessed at T3 and served for a supplementary analysis - please refer to Supplementary Table 2 for a full list of tests and their description.

Information regarding family history of dyslexia was obtained using the Adult Reading History Questionnaire (ARHQ, Lefly & Pennington, 2000), for both parents (except for one father), and considering possible diagnosis of older siblings. Several measures of dyslexia family history were derived: a binary dyslexia family history variable (at least one parent or sibling with a dyslexia diagnosis, vs. none), mother and father ARHQ scores, as well as averaged parental ARHQ scores weighted by either total general or educational time spent with the child by each caregiver.

MRI Data acquisition

MRI data were acquired on a 3T GE Healthcare scanner at the Richard M. Lucas Center for Imaging at Stanford University using an eight-channel phased-array head coil. T_1-weighted (T_1w) images were acquired at T1 and T3 with a fast spoiled gradient echo sequence (repetition time: TR = 8.52 ms; echo time TE = 3.42 ms; inversion time TI = 400 ms; flip angle FA = 15 °; NEX = 1; 128 slices; thickness = 1.2 mm; field of view FOV = 22 cm; in-plane resolution = 256 × 256; voxel size = 0.859 × 0.859 × 1.2 mm).

Diffusion-weighted images (DWI) were collected with a high-angular resolution diffusion-imaging (HARDI) protocol at Time 3 with a single-shot spin-echo, echo-planar imaging sequence (46 axial 3mm-thick slices; repetition time TR = 5000 ms; echo time TE = 81.7 ms; in-plane resolution = 128 × 128; voxel size = 2.0 × 2.0 × 3.0 mm^3; 150 directions with b = 2500 s/mm^2; 6 volumes with b = 0 s/mm^2, posterior-to-anterior phase encoding). DWI data were successfully collected in 35 participants. Among these, one was excluded due to a neurological finding, 2 due to acquisition issues related to the bounding box (missing parts of the brain) and 3 for excessive motion (see DWI processing), resulting in a final DWI sample at T3 of 29 children.

T_1-weighted data processing

T_1-weighted images were processed in Freesurfer (Dale et al., 1999; Fischl et al., 1999) version 7.1.1, using the high-resolution sub-millimeter pipeline (Zaretskaya et al., 2018), in order to (1) segment gray and white matter, (2) reconstruct gray and white matter surfaces, and (3) dissociate sulcal from gyral components. Images were first processed cross-sectionally for each scan. When images were available at both Time 1 and Time 3 for a child (N=43), the longitudinal Freesurfer pipeline was used, which first creates an unbiased subject average across time points ("base") and then restarts multiple processing steps at each time point separately from the base template in order to refine tissue segmentations on each image (respectively T1 and T3 longitudinally processed images, Reuter et al., 2012; see Figure 1A). The application of such longitudinal processing steps has been demonstrated to result in more accurate assessments of anatomy and evaluation of longitudinal change, as we also confirmed. Segmentations were inspected visually for quality of the acquisition and of the Freesurfer segmentation bilaterally in ventral occipito-temporal cortex susceptible to affect the identification of the OTS and MFS. This led to the rejection of 1 participant at T1 and 2 subjects at T3 due to excessive motion artifacts. An additional participant was excluded due to large bilateral arachnoid cysts in the anterior temporal lobe.

Sulcal definitions & criteria

Definition of the OTS and MFS

The OTS and MFS were identified manually, for each participant at each time point (T1 and T3), and processing pipeline (Freesurfer cross-sectional vs. longitudinal). Sulci were labeled by FB and verified or corrected by KW. Sulcal labeling was conducted without awareness of the children's reading scores, and relied on the identification and the anatomy of nearby sulci based on previous criteria from our group (Weiner, 2019; Weiner et al., 2014) and others (Cachia et al., 2018). Surrounding sulci considered when defining the OTS and MFS included the collateral sulcus (CoS) medially, the posterior transverse collateral sulcus (ptCoS) posteriorly, the anterior transverse CoS (atCoS) anteriorly, and the inferior and superior temporal sulci (ITS and STS) laterally. Additionally, the pre-occipital notch (see Figure 1B) was referenced as a useful landmark identifying the posterior extent of the OTS on the pial surface. Sulci were defined on Freesurfer's inflated surface as components with negative curvature, taking into consideration folding patterns as seen in the volume or white and pial surfaces, and potentially exclusive of connected branches corresponding to other sulci.

The ptCoS was first identified as the sulcus extending from the CoS-proper (Petrides, 2018) and forming the posterior boundary of the fusiform gyrus.

The OTS was then defined as the long and deep sulcus separating the fusiform gyrus from the inferior temporal gyrus. As such, its posterior boundary was defined by a virtual line extending from the ptCoS to the pre-occipital notch (as best identified on the pial surface), which delineates the boundary between temporal and occipital cortex (Duvernoy, 1999; Ono et al., 1990; Petrides, 2018). Note that, as a consequence, our OTS definition is exclusive of occipital components and branches, considered as the inferior occipital sulcus, although we retained the classical denomination of the *occipito*-temporal sulcus. In cases when the OTS extended into the occipital cortex, the OTS-proper was thus interrupted at the level of a *pli de passage* most consistent with the virtual line between the ptCoS and pre-occipital notch. The pre-occipital notch was not considered as part of the OTS and excluded from OTS delineations in the frequent case where it was connected to the OTS. Just as the ptCoS can be considered as the posterior hook identifying the posterior extent of the fusiform gyrus, the atCoS can be considered as an anterior hook identifying the anterior extent of the fusiform gyrus (Duvernoy, 1999). Laterally, the identification of the ITS, and of the STS and its posterior branches (Huntgeburth & Petrides, 2012) were used to disambiguate OTS components in complex cases.

The MFS was defined as the longitudinal sulcus that divides the fusiform gyrus into lateral and medial components following guidelines by Weiner and colleagues (Weiner et al., 2014; as also reviewed in Weiner, 2019). One characteristic feature of the MFS, due to its location in between the CoS and OTS on the cortical surface, is that it forms a distinct ω shape on coronal slices (corresponding to the middle bump of the ω). Indeed, as a putative tertiary sulcus, the MFS is about half as deep as the surrounding OTS and CoS (Weiner et al., 2014). Additionally, the anterior tip of the MFS consistently aligns with the posterior tip of the hippocampus (Grill-Spector & Weiner, 2014).

Posterior / anterior OTS

A fixed boundary in MNI template space at coordinate Y=-40 was used to distinguish posterior sections of the OTS (corresponding to typical coordinates of the VTC) from anterior sections. Posterior and anterior volumes in MNI space were first transformed to Freesurfer average subject space (*fsaverage*), then projected onto the *fsaverage* surface, and finally projected onto each individual surface, to split the OTS into posterior and anterior sections (pOTS and aOTS respectively, see Figure 1B).

Definition of gyral gaps / plis de passage for T3 DWI analyses

In order to investigate white-matter properties of the sulcal interruptions that may be associated with a reading benefit, the gyral gaps, including gyral crowns, located between chunks of the left pOTS were delineated manually by outlining the nearest points between the sulcal components (as

shown in Figure 4).

Sulcal measurements

From the defined OTS, pOTS, aOTS, and MFS, seven anatomical features were extracted as possibly relevant features: (i) sulcal interruption, (ii) interruption distance between sulcal components, (iii) sulcal length, (iv) sulcal depth, (v) depth of the sulcal pit, (vi) sulcal thickness, and (vii) sulcal surface area. Interruption distance between sulcal components was considered as Cachia and colleagues (2018) observed an association between the size of left pOTS interruptions along the postero-anterior axis and reading skills. In the present work, interruption distance was measured as the Euclidean distance between two sulcal components, or the sum of distances in instances where there were three or more components of a sulcus. Specifically, sulcal interruptions were counted as such when sulcal components were separated by at least 1mm, or when the total interruption distance across all components of a sulcus was larger than 1mm, as very small interruptions may be attributable to noisy scans, partial volume effects or poor brain segmentations. Sulcal length was considered as a complementary feature of sulcal interruption and interruption distance, known to be behaviorally relevant (see e.g., Parker et al., 2023, for associations between MFS length and facial recognition abilities). Sulcal length was measured by custom software as the largest geodesic distance between any two points within a given sulcal component, or the sum of lengths of all sulcal components in cases of interrupted sulci. Sulcal depth was measured using the *calcSulc* toolbox (Madan, 2019), by extracting the median Euclidean distance of the 8% deepest vertices of a given sulcus (as identified by Freesurfer *sulc* values) to the enclosing brain surface. We chose to select the 8% deepest vertices, rather than the initially set number of vertices (100 in the toolbox), in order to prevent the biasing of sulcal depth across sulci of different sizes (notably the much smaller MFS). Results were unchanged when using a percentage of deepest vertices ranging from 4 to 12%. The depth of the sulcal pit was also considered as sulcal pits often constitute stable functional and behavioral landmarks (e.g., Natu et al., 2021).

DWI processing

Each diffusion-weighted image was first denoised using the Marchenko-Pastur principal component analysis technique implemented in MRTrix3 (Veraart et al., 2016). Because no diffusion data with reverse phase encoding was available, distortions due to susceptibility artifacts were corrected using the *Synb0-DISCO* approach (Schilling et al., 2020). Accordingly, an undistorted b0 image was estimated using deep learning (3D U-nets) in order to match the geometry of structural T_1w images and intensity contrasts from diffusion images. The distortions were then estimated and applied using the *topup* tool in FSL (Andersson et al., 2003). Both distortion and motion corrections were applied using the *eddy* function in FSL (Andersson & Sotiropoulos, 2016). Due to relatively high motion in this pediatric sample, *eddy* was run with conservative options consisting in 8 iterations with decreasing amounts of smoothing at 10, 6, 4, 2, 0, 0, 0, 0 mm full width at half maximum (FWHM), with additional outlier detection and replacement, and estimation of susceptibility-induced field changes by subject movement (Andersson et al., 2018). Volumes with more than 10% of outlier slices as identified by *eddy*, or with more than 1 mm of displacement relative to the previous volume, were excluded. Three participants with more than 10% of discarded volumes were excluded from further analyses, resulting in a final sample of 29 children with DWI data at Time 3.

In order to most accurately align individual T_1w and DWI data, non-linear transformations were estimated using Advanced Normalization Tools (ANTs, Avants et al., 2011) between the participant's bias-corrected T_1w image and the individual average of all distortion-corrected b0 images (which have a relatively high gray/white contrast compared to other DWI-derived images). Specifically, non-linear transformations were used as an additional measure to correct for distortions induced in VTC by susceptibility artifacts due to the proximity of air/tissue interfaces in the nearby ear canals.

DWI data were first used to derive diffusion tensor indices in order to compare microstructural properties of the cortex and white matter around the left pOTS between sulcal components per se and the gyral gaps, and between children with and without interruptions. Second, probabilistic tractography from this region was performed to investigate the correlates of pOTS interruptions in terms of structural connectivity, between groups of children and according to reading scores.

The tensor model was estimated in FSL using *dtifit* in order to derive diffusion-tensor imaging (DTI) maps of mean and radial diffusivities (MD and RD) and fractional anisotropy (FA), as indices of myelin content. While MD is the most robust DTI index reflective of myelination and neurite properties in the cortex (Fukutomi et al., 2019), FA and RD are relevant markers of myelin content in white matter (Dubois et al., 2014; Lazari & Lipp, 2021).

To estimate MD, RD, and FA values most accurately in the cortical ribbon and superficial white matter above the left pOTS and its interruptions, individual MD, RD, and FA maps were transformed into native T_1w space and averaged within local masks of the cortical ribbon and superficial white matter. These masks were obtained by propagating Freesurfer labels of the left pOTS and its interruptions into the volumes using the *mri_label2vol* function, either into the whole cortical ribbon or into a 2-voxel-thick layer of the superficial white matter (see Figure 4B and C). Additionally, a white matter mask was used to prevent the white-matter layer from extending into gray matter voxels of neighboring sulci.

To identify differences in structural connectivity associated with interruption of the left pOTS, probabilistic tractography was performed from a mask of superficial white matter located above the left pOTS labels, inclusive of the gyral crowns of the interruption(s) / *pli(s) de passage* in children with left pOTS interruption. Masks of superficial white matter above the left pOTS and interruptions were transformed from T1-w images into diffusion native spaces through ANTs using nearest-neighbor interpolation, to serve as tractography seeds. FSL's *bedpostx* function (Bayesian Estimation of Diffusion Parameters Obtained using Sampling Technique) was used to estimate distributions of diffusion parameters and to model crossing fibers in each voxel (3 fibers per voxel). *Probtrackx* was used to perform probabilistic tractography from individual seeds, using the Euler integration for the precise computing of probabilistic streamlines, distance correction to compensate for the fact that connectivity distribution drops with distance from the seed mask, and with segmentations of gray matter and ventricles as termination masks. The minimum allowable angle between two tractography steps was set to 30 degrees, to allow for the reconstruction of U-shaped fibers, which pass by the white-matter underlying sulcal fundi connecting adjacent gyri - in this case, the fusiform and inferior temporal gyri. Because seed volumes significantly differed between children with interrupted and continuous left pOTS, the number of streamlines started per seed voxel in the tractography algorithm was scaled individually based on seed volume to correspond to an average of 5,000 seeds per voxel (default value in *probtrackx*) and to yield similar total numbers of streamlines. The latter was implemented in order to accurately assess group differences in structural connectivity, unbiased by seed volume. Individual masks of gray matter and ventricles were used as stop masks for the tractography. Distance correction and modified Euler streamlining options were used to improve the accuracy of tractography. The obtained individual probabilistic fiber density maps were then transformed into template MNI space for statistical analysis using non-linear registration in ANTs, from individual diffusion space (individual FA map) to the HCP1065 MNI FA template, and smoothed with a 4-mm-FWHM Gaussian kernel. Voxel-wise permutation (non-parametric) testing was performed with FSL's *randomise* function (Winkler et al., 2014), to assess the effect of left pOTS interruption on structural connections of its underlying white-matter, controlling for age, gender and handedness. Results of such permutation tests were compared to those of a similar model with reading scores (controlled for demographic variables) as an additional predictor, so as to assess whether differences in structural connectivity were more attributable to interruption of the left pOTS or differences in reading skills.

Statistical analyses

Statistical analyses were performed in R version 4.2.2 (R Core Team, 2022). The impact of sulcal anatomical features on reading skills, or other test scores, across time points was assessed while controlling for the possible impact of demographic variables (e.g., age, gender and handedness, on top of total brain volume (TBV) in the case of sulcal features). Specifically, for each sulcal feature and each assessment score of interest, an F-test was performed to compare two nested models predicting scores at time Ti (i=1, 2 or 3) from demographic variables only (baseline model), or taking into additional consideration sulcal features measured at time Tj (j= 1 or 3), controlled for nuisance variables, (test model) as follows:

Baseline model: $Score_{Ti} \sim Age_{Ti}$ + Gender + Handedness

Test model: $Score_{Ti} \sim Age_{Ti}$ + Gender + Handedness + residuals (logit(sulcal $interruption_{Tj} \sim Age_{Tj}$ + TBV_{Tj} + Gender + Handedness))

where *logit* indicates a logistic regression.

Or, in the case of continuous sulcal measurements:

Test model: $Score_{Ti} \sim Age_{Ti}$ + Gender + Handedness + residuals(sulcal $measure_{Tj} \sim Age_{Tj}$ + TBV_{Tj} + Gender + Handedness)

F-statistics and associated p-values are reported for these nested model comparisons, as well as the difference in variance explained between models (ΔR^2 in %). For more coherence with statistical tests, plotted scores are controlled for demographic variables (age, gender and handedness; i.e., residuals are plotted, after adding back population mean to retain the estimation of average scores, see e.g., Fig. 2a or Fig. S4). Plots were created using the R *ggplot2* package (Wickham, 2016). Early publicly available code corresponding to the R *ggrain* package was adjusted to create *raincloud plots* (Allen et al., 2019). To investigate the unique contributions of several predictors of reading (see for instance Figure S5b and f) while controlling for other variables, such predictors were added to both the baseline and test models, following the same logic. Finally, to investigate possible associations at T1 between sulcal interruption and precursors of reading, we compared the model fits between nested logistic regressions predicting left pOTS interruption based on demographic variables and TBV (baseline), including or not each reading precursor separately as follows:

Baseline model: logit(sulcal $interruption_{T1} \sim Age_{T1}$ + TBV_{T1} + Gender + Handedness)

Test model: logit(sulcal $interruption_{T1} \sim Precursor_score_{T1}$ + Age_{T1} + TBV_{T1} + Gender + Handedness)

Statistics were conducted on standard scores from the standardized neuropsychological tests as they are more normally distributed than raw scores (a requirement for parametric tests). More importantly, the rationale of the study is to assess whether sulcal morphology will predict whether a child is going to be a relatively good reader or not, independently from age. Note that despite standardization of raw scores within age bins in some tests, age still had a residual significant effect on test scores (TOWRE T3 standard score notably) and was therefore included in baseline models. Because the TOWRE is not normed below 6 years of age, an equivalent of a composite TOWRE standard score was obtained for T1 by scaling the principal component of T1 SWE and PDE raw scores (to a mean of 100 with standard deviation of 15).

To reduce degrees of freedom concerning measures of dyslexia family history, a stepwise backward selection was applied to identify family history measure(s) that best predicted T3 TOWRE composite scores (accounting for age, gender and handedness). Binary family history factor (i.e. *does a direct relative have a formal diagnosis?*), mother and father ARHQ scores and their weighted averages based on educational and total general parental time spent with the child were entered into the model. Among these, mom ARHQ scores were selected as the most reliable measure predicting children TOWRE T3 reading scores, without any significant additional contribution from other variables.

Because some children already knew how to read at Time 1 (median of 3 words read in 45s in the SWE subtest, ranging from 0 to 58; median of 2 syllables or pseudowords read in 45s in the PDE subtest, ranging from 0 to 36), complementary analyses were performed to (i) investigate whether the predictive power of left pOTS interruption held in a population of children with minimal reading skills at T1 (Fig. S4), and (ii) to better assess the effect size of independent contributions of left pOTS morphology on future reading skills (Fig. S5) over initial reading scores at T1 (to which it was already correlated). We therefore selected a subsample of children whose TOWRE standard score equivalent at Time 1 was below 95, resulting in a subset of 24 *initial minimal readers* (median of 0 words read in 45s in SWE, range: 0-7; median of 0 syllables in 45s in PDE, range: 0-4; see Table 1).

The serial mediation analysis was performed using the R *lavaan* package (Rosseel, 2012). Parameter estimates and confidence intervals were estimated using 10,000 bootstrap samples.

Data availability

Public archiving of anonymised study data is not permitted by the ethics approval. Readers seeking access to the data should contact the corresponding author via email (florence.bouhali@univ-amu.fr). Data will be released to researchers on the following conditions: approval from the local research ethics committee and with appropriate safeguards to protect from the identfication of individuals.

Acknowledgements

Authors would like to thank Olivier Coulon for useful discussion, as well as Jacob Miller for sharing his script to extract sulcal length. Data collection was funded by the National Institutes of Health (NIH) under grant K23HD054720 to FH. FB was supported by NIH grant R01HD094834, the Institute of Convergence for Language, Communication and the Brain (ILCB, grants from France 2030 ANR-16-CONV-0002 and the Excellence Initiative of Aix-Marseille University A*MIDEX), and the French Foundation for Medical Research (Fondation pour la Recherche Médicale, FRM ARF202209015734). JD was supported by the Fondation Médisite (under the aegis of the Fondation de France, grant FdF-18-00092867) and the IdEx Université de Paris (ANR-18-IDEX-0001). FH was supported by the National Science Foundation (NRT-2152202, MCA-2120888), NIH (R01DC013064, U24AT011281, R01HD094834, R01HD096261), U.S. Department of Agriculture's (USDA) / National Institute of Food and Agriculture (NIFA) (2022-11791) and the Oak Foundation (OCAY-19-215). KSW was supported by the National Science Foundation (CAREER 2042251) and the Brain and Behavior Research Foundation (NARSAD 30738).

Additional files

Appendix A.

Additional information

Funding

Funder	Grant reference number	Author
HHS | National Institutes of Health (NIH)	K23HD054720	Fumiko Hoeft
HHS | National Institutes of Health (NIH)	R01HD094834	Fumiko Hoeft
Agence Nationale de la Recherche (ANR)	https://doi.org/10.67599/anr-16-conv-0002	Florence Bouhali
Fondation pour la Recherche Médicale (FRM)	ARF202209015734	Florence Bouhali
Fondation de France (Fondationfrance)	FdF-18-00092867	Jessica Dubois
Agence Nationale de la Recherche (ANR)	ANR-18-IDEX-0001	Jessica Dubois
National Science Foundation (NSF)	NRT-2152202	Fumiko Hoeft

National Science Foundation (NSF)	MCA-2120888	Fumiko Hoeft
HHS | National Institutes of Health (NIH)	R01DC013064	Fumiko Hoeft
HHS | National Institutes of Health (NIH)	U24AT011281	Fumiko Hoeft
HHS | National Institutes of Health (NIH)	R01HD096261	Fumiko Hoeft
U.S. Department of Agriculture (USDA)	2022-11791	Fumiko Hoeft
Oak Foundation (Oak Foundation USA)	OCAY-19-215	Fumiko Hoeft
National Science Foundation (NSF)	CAREER 2042251	Kevin S Weiner
Brain and Behavior Research Foundation (BBRF)	NARSAD 30738	Kevin S Weiner

Author ORCID iDs

Florence Bouhali: https://orcid.org/0000-0001-7357-1531
Jessica Dubois: https://orcid.org/0000-0003-4865-8111
Kevin S Weiner: https://orcid.org/0000-0002-8734-5049

References

1. **Allen M,** Poggiali D, Whitaker K, Marshall TR, van Langen J, Kievit RA (2019) Raincloud plots: A multi-platform tool for robust data visualization. *Wellcome Open Research* **4** https://doi.org/10.12688/wellcomeopenres.15191.2 | PubMed
2. **Amiez C,** Wilson CRE, Procyk E (2018) Variations of cingulate sulcal organization and link with cognitive performance. *Scientific Reports* **8**:13988 https://doi.org/10.1038/s41598-018-32088-9 | PubMed
3. **Andersson JLR,** Graham MS, Drobnjak I, Zhang H, Campbell J (2018) Susceptibility-induced distortion that varies due to motion: Correction in diffusion MR without acquiring additional data. *NeuroImage* **171**:277-295 https://doi.org/10.1016/j.neuroimage.2017.12.040 | PubMed
4. **Andersson JLR,** Skare S, Ashburner J (2003) How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging. *NeuroImage* **20**:870-888 https://doi.org/10.1016/S1053-8119(03)00336-7 | PubMed
5. **Andersson JLR,** Sotiropoulos SN (2016) An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *NeuroImage* **125**:1063-1078 https://doi.org/10.1016/j.neuroimage.2015.10.019 | PubMed
6. **Avants BB,** Tustison NJ, Song G, Cook PA, Klein A, Gee JC (2011) A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage* **54**:2033-2044 https://doi.org/10.1016/j.neuroimage.2010.09.025 | PubMed
7. **Beelen C,** Vanderauwera J, Wouters J, Vandermosten M, Ghesquière P (2019) Atypical gray matter in children with dyslexia before the onset of reading instruction. *Cortex* **121**:399-413 https://doi.org/10.1016/j.cortex.2019.09.010 | PubMed
8. **Black JM,** Tanaka H, Stanley L, Nagamine M, Zakerani N, Thurston A, Kesler S, Hulme C, Lyytinen H, Glover GH, *et al.* (2012) Maternal history of reading difficulty is associated with reduced language-related gray matter in beginning readers. *NeuroImage* **59**:3021-3032 https://doi.org/10.1016/j.neuroimage.2011.10.024 | PubMed
9. **Bodin C,** Pron A, Le Mao M, Régis J, Belin P, Coulon O (2021) Plis de passage in the superior temporal sulcus: Morphology and local connectivity. *NeuroImage* **225** https://doi.org/10.1016/j.neuroimage.2020.117513 | PubMed
10. **Borst G,** Cachia A, Tissier C, Ahr E, Simon G, Houdé O (2016) Early Cerebral Constraints on Reading Skills in School-Age Children: An MRI Study. *Mind. Brain, and Education* **10**:47-54 https://doi.org/10.1111/mbe.12098

11. **Borst G,** Cachia A, Vidal J, Simon G, Fischer C, Pineau A, Poirel N, Mangin J.-F, Houdé O (2014) Folding of the anterior cingulate cortex partially explains inhibitory control during childhood: A longitudinal study. *Developmental Cognitive Neuroscience* **9**:126-135 https://doi.org/10.1016/j.dcn.2014.02.006 | PubMed

12. **Bouhali F,** Bézagu Z, Dehaene S, Cohen L (2019) A mesial-to-lateral dissociation for orthographic processing in the visual cortex. *Proceedings of the National Academy of Sciences* **116**:201904184 https://doi.org/10.1073/pnas.1904184116 | PubMed

13. **Bouhali F,** de Schotten MT, Pinel P, Poupon C, Mangin J.-F, Dehaene S, Cohen L (2014) Anatomical connections of the visual word form area. *Journal of Neuroscience* **34** https://doi.org/10.1523/JNEUROSCI.4918-13.2014 | PubMed

14. **Cachia A,** Borst G, Jardri R, Raznahan A, Murray GK, Mangin J.-F, Plaze M (2021) Towards Deciphering the Fetal Foundation of Normal Cognition and Cognitive Symptoms From Sulcation of the Cortex. *Frontiers in Neuroanatomy* **15** https://doi.org/10.3389/fnana.2021.712862 | PubMed

15. **Cachia A,** Borst G, Tissier C, Fisher C, Plaze M, Gay O, Rivière D, Gogtay N, Giedd J, Mangin J.-F, *et al.* (2016) Longitudinal stability of the folding pattern of the anterior cingulate cortex during development. *Developmental Cognitive Neuroscience* **19**:122-127 https://doi.org/10.1016/j.dcn.2016.02.011 | PubMed

16. **Cachia A,** Borst G, Vidal J, Fischer C, Pineau A, Mangin J.-F, Houdé O (2014) The Shape of the ACC Contributes to Cognitive Control Efficiency in Preschoolers. *Journal of Cognitive Neuroscience* **26**:96-106 https://doi.org/10.1162/jocn_a_00459 | PubMed

17. **Cachia A,** Dupont C, Palmis S, Menu I, Auzias G, Sein J, Coulon O, Longcamp M (2025) The sulcal patterns of the occipito-temporal and anterior cingulate cortices influence reading and writing in children and adults. *Cerebral Cortex* **35**:bhaf145 https://doi.org/10.1093/cercor/bhaf145 | PubMed

18. **Cachia A,** Roell M, Mangin J.-F, Sun ZY, Jobert A, Braga L, Houde O, Dehaene S, Borst G (2018) How interindividual differences in brain anatomy shape reading accuracy. *Brain Structure and Function* **223**:701-712 https://doi.org/10.1007/s00429-017-1516-x | PubMed

19. **Chavoshnejad P,** Li X, Zhang S, Dai W, Vasung L, Liu T, Zhang T, Wang X, Razavi MJ (2021) Role of axonal fibers in the cortical folding patterns: A tale of variability and regularity. *Brain Multiphysics* **2**:100029 https://doi.org/10.1016/j.brain.2021.100029

20. **Chi JG,** Dooling EC, Gilles FH (1977) Gyral development of the human brain. *Annals of Neurology* **1**:86-93 https://doi.org/10.1002/ana.410010109 | PubMed

21. **Cohen L,** Dehaene S, Naccache L, Lehéricy S, Dehaene-Lambertz G, Hénaff M.-A, Michel F (2000) The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain* **123**:291-307 https://doi.org/10.1093/brain/123.2.291 | PubMed

22. **Dale AM,** Fischl B, Sereno MI (1999) Cortical Surface-Based Analysis. *NeuroImage* **9**:179-194 https://doi.org/10.1006/nimg.1998.0395 | PubMed

23. **Dannhoff G,** Poudel PP, Bhattarai C, Kalthur SG, Maldonado IL (2023) Depicting the anatomy of the gyral white matter: Ubi sumus? quo vadimus?. *Brain Communications* **5**:fcad265 https://doi.org/10.1093/braincomms/fcad265 | PubMed

24. **Dehaene S,** Cohen L (2007) Cultural Recycling of Cortical Maps. *Neuron* **56**:384-398 https://doi.org/10.1016/j.neuron.2007.10.004 | PubMed

25. **Dehaene S,** Cohen L (2011) The unique role of the visual word form area in reading. *Trends in Cognitive Sciences* **15**:254-262 https://doi.org/10.1016/j.tics.2011.04.003 | PubMed

26. **Dehaene S,** Cohen L, Morais J, Kolinsky R (2015) Illiterate to literate: Behavioural and cerebral changes induced by reading acquisition. *Nature Reviews Neuroscience* **16**:234-244 https://doi.org/10.1038/nrn3924 | PubMed

27. **Dehaene-Lambertz G,** Monzalvo K, Dehaene S (2018) The emergence of the visual word form: Longitudinal evolution of category-specific ventral visual areas during reading acquisition. *PLoS Biology* **16** https://doi.org/10.1371/journal.pbio.2004103 | PubMed

28. **Del Mauro G,** Del Maschio N, Abutalebi J (2022) The relationship between reading abilities and the left occipitotemporal sulcus: A dual perspective study. *Brain and Language* **235**:105189 https://doi.org/10.1016/j.bandl.2022.105189 | PubMed

29. **Dubois J,** Benders M, Borradori-Tolsa C, Cachia A, Lazeyras F, Ha-Vinh Leuchter R, Sizonenko SV, Warfield SK, Mangin JF, Hüppi PS (2008) Primary cortical folding in the human newborn: An early marker of later functional development. *Brain* **131**:2028-2041 https://doi.org/10.1093/brain/awn137 | PubMed

30. **Dubois J,** Dehaene-Lambertz G, Kulikova S, Poupon C, Hüppi PS, Hertz-Pannier L (2014) The early development of brain white matter: A review of imaging studies in fetuses, newborns and infants. *Neuroscience* **276**:48-71 https://doi.org/10.1016/j.neuroscience.2013.12.044 | PubMed

31. **Dubois J,** Lefèvre J, Angleys H, Leroy F, Fischer C, Lebenberg J, Dehaene-Lambertz G, Borradori-Tolsa C, Lazeyras F, Hertz-Pannier L, *et al.* (2019) The dynamics of cortical folding waves and prematurity-related deviations revealed by spatial and spectral analysis of gyrification. *NeuroImage* **185**:934-946 https://doi.org/10.1016/j.neuroimage.2018.03.005 | PubMed

32. **Dunn LM,** Dunn DM (2007) *Peabody Picture Vocabulary Test—Fourth Edition. Minneapolis* MN: Pearson. https://doi.org/10.1037/t15144-000

33. **Duvernoy HM** (1999) The human brain: Surface, three-dimensional sectional anatomy with MRI, and blood supply.

34. **Fehlbaum LV,** Peters L, Dimanova P, Roell M, Borbás R, Ansari D, Raschle NM (2022) Mother-child similarity in brain morphology: A comparison of structural characteristics of the brain's reading network. *Developmental Cognitive Neuroscience* **53**:101058 https://doi.org/10.1016/j.dcn.2022.101058 | PubMed

35. **Feldt LS,** Woodruff DJ, Salih FA (1987) Statistical Inference for Coefficient Alpha. *Applied Psychological Measurement* **11**:93-103 https://doi.org/10.1177/014662168701100107

36. **Ferrer E,** McArdle JJ, Shaywitz BA, Holahan JM, Marchione K, Shaywitz SE (2007) Longitudinal Models of Developmental Dynamics Between Reading and Cognition From Childhood to Adolescence. *Developmental Psychology* **43**:1460-1473 https://doi.org/10.1037/0012-1649.43.6.1460 | PubMed

37. **Fischl B,** Sereno MI, Dale AM (1999) Cortical Surface-Based Analysis: II: Inflation, Flattening, and a Surface-Based Coordinate System. *NeuroImage* **9**:195-207 https://doi.org/10.1006/nimg.1998.0396 | PubMed

38. **Fukutomi H,** Glasser MF, Murata K, Akasaka T, Fujimoto K, Yamamoto T, Autio JA, Okada T, Togashi K, Zhang H, *et al.* (2019) Diffusion Tensor Model links to Neurite Orientation Dispersion and Density Imaging at high b-value in Cerebral Cortical Gray Matter. *Scientific Reports* **9**:12246 https://doi.org/10.1038/s41598-019-48671-7 | PubMed

39. **Gaillard R,** Naccache L, Pinel P, Clémenceau S, Volle E, Hasboun D, Dupont S, Baulac M, Dehaene S, Adam C, *et al.* (2006) Direct Intracranial, fMRI, and Lesion Evidence for the Causal Role of Left Inferotemporal Cortex in Reading. *Neuron* **50**:191-204 https://doi.org/10.1016/j.neuron.2006.03.031 | PubMed

40. **Garel C,** Chantrel E, Brisse H, Elmaleh M, Luton D, Oury J.-F, Sebag G, Hassan M (2001) Fetal Cerebral Cortex: Normal Gestational Landmarks Identified Using Prenatal MR Imaging. *American Journal of Neuroradiology* **22**:184-189 PubMed

41. **Garrison JR,** Fernyhough C, McCarthy-Jones S, Haggard M, Simons JS (2015) Paracingulate sulcus morphology is associated with hallucinations in the human brain. *Nature Communications* **6**:8956 https://doi.org/10.1038/ncomms9956 | PubMed

42. **Ge F,** Li X, Razavi MJ, Chen H, Zhang T, Zhang S, Guo L, Hu X, Wang X, Liu T (2018) Denser Growing Fiber Connections Induce 3-hinge Gyral Folding. *Cerebral Cortex* **1991** https://doi.org/10.1093/cercor/bhx227 | PubMed

43. **Gratiolet P** (1854) *Mémoire sur les plis cérébraux de l'homme et des primates (Vol. 1)* Arthus Bertrand.
44. **Gratton C,** Nelson SM, Gordon EM (2022) Brain-behavior correlations: Two paths toward reliability. *Neuron* **110**:1446-1449 https://doi.org/10.1016/j.neuron.2022.04.018 | PubMed
45. **Grill-Spector K,** Weiner KS (2014) The functional architecture of the ventral temporal cortex and its role in categorization. *Nature Reviews Neuroscience* **15**:536-548 https://doi.org/10.1038/nrn3747 | PubMed
46. **Grotheer M,** Yeatman J, Grill-Spector K (2021) White matter fascicles and cortical microstructure predict reading-related responses in human ventral temporal cortex. *NeuroImage* **227** https://doi.org/10.1016/j.neuroimage.2020.117669 | PubMed
47. **Hansen PE,** Ballesteros MC, Soila K, Garcia L, Howard JM (1993) MR imaging of the developing human brain. Part 1. Prenatal development. *RadioGraphics* **13**:21-36 https://doi.org/10.1148/radiographics.13.1.8426929 | PubMed
48. **Hosseini SMH,** Black JM, Soriano T, Bugescu N, Martinez R, Raman MM, Kesler SR, Hoeft F (2013) Topological properties of large-scale structural brain networks in children with familial risk for reading difficulties. *NeuroImage* **71**:260-274 https://doi.org/10.1016/j.neuroimage.2013.01.013 | PubMed
49. **Huettig F,** Hulstijn J (2025) The Enhanced Literate Mind Hypothesis. *Topics in Cognitive Science* **17**:909-918 https://doi.org/10.1111/tops.12731 | PubMed
50. **Huntgeburth SC,** Petrides M (2012) Morphological patterns of the collateral sulcus in the human brain. *European Journal of Neuroscience* **35**:1295-1311 https://doi.org/10.1111/j.1460-9568.2012.08031.x | PubMed
51. **Im K,** Raschle NM, Smith SA, Ellen Grant P, Gaab N (2016) Atypical Sulcal Pattern in Children with Developmental Dyslexia and At-Risk Kindergarteners. *Cerebral Cortex* **26**:1138-1148 https://doi.org/10.1093/cercor/bhu305 | PubMed
52. **Kubota E,** Grill-Spector K, Nordt M (2023) Rethinking cortical recycling in ventral temporal cortex. *Trends in Cognitive Sciences* https://doi.org/10.1016/j.tics.2023.09.006 | PubMed
53. **Lahutsina A,** Spaniel F, Mrzilkova J, Morozova A, Brabec M, Musil V, Zach P (2023) Morphology of Anterior Cingulate Cortex and Its Relation to Schizophrenia. *Journal of Clinical Medicine* **12**:33 https://doi.org/10.3390/jcm12010033 | PubMed
54. **Lazari A,** Lipp I (2021) Can MRI measure myelin? Systematic review, qualitative assessment, and meta-analysis of studies validating microstructural imaging with myelin histology. *NeuroImage* **230**:117744 https://doi.org/10.1016/j.neuroimage.2021.117744 | PubMed
55. **Lefly DL,** Pennington BF (2000) Reliability and validity of the adult reading history questionnaire. *Journal of Learning Disabilities* https://doi.org/10.1177/002221940003300306 | PubMed
56. **Li J,** Osher DE, Hansen HA, Saygin ZM (2020) Innate connectivity patterns drive the development of the visual word form area. *Scientific Reports* **10**:1-13 https://doi.org/10.1038/s41598-020-75015-7 | PubMed
57. **Lyu I,** Bao S, Hao L, Yao J, Miller JA, Voorhies W, Taylor WD, Bunge SA, Weiner KS, Landman BA (2021) Labeling lateral prefrontal sulci using spherical data augmentation and context-aware training. *NeuroImage* **229**:117758 https://doi.org/10.1016/j.neuroimage.2021.117758 | PubMed
58. **Maboudian SA,** Willbrand EH, Kelly JP, Jagust WJ, Weiner KS, Alzheimer's Disease Neuroimaging Initiative (2024) Defining Overlooked Structures Reveals New Associations between the Cortex and Cognition in Aging and Alzheimer's Disease. *Journal of Neuroscience* **44** https://doi.org/10.1523/JNEUROSCI.1714-23.2024 | PubMed
59. **Madan CR** (2019) Robust estimation of sulcal morphology. *Brain Informatics* **6**:1-11 https://doi.org/10.1186/s40708-019-0098-1 | PubMed
60. **Mangin J.-F,** Le Guen Y, Labra N, Grigis A, Frouin V, Guevara M, Fischer C, Rivière D, Hopkins WD, Régis J, *et al.* (2019) "Plis de passage" Deserve a Role in Models of the Cortical Folding Process. *Brain Topography* **32**:1035-1048 https://doi.org/10.1007/s10548-019-00734-8 | PubMed

61. **Marek S,** Tervo-Clemmens B, Calabro FJ, Montez DF, Kay BP, Hatoum AS, Donohue MR, Foran W, Miller RL, Hendrickson TJ, *et al.* (2022) Reproducible brain-wide association studies require thousands of individuals. *Nature* **603**:654-660 https://doi.org/10.1038/s41586-022-04492-9 | PubMed

62. **Miller JA,** Voorhies WI, Lurie DJ, D'Esposito M, Weiner KS (2021) Overlooked tertiary sulci serve as a meso-scale link between microstructural and functional properties of human lateral prefrontal cortex. *Journal of Neuroscience* **41**:2229-2244 https://doi.org/10.1523/JNEUROSCI.2362-20.2021 | PubMed

63. **Moulton E,** Bouhali F, Monzalvo K, Poupon C, Zhang H, Dehaene S, Dehaene-Lambertz G, Dubois J (2019) Connectivity between the visual word form area and the parietal lobe improves after the first year of reading instruction: A longitudinal MRI study in children. *Brain Structure and Function* https://doi.org/10.1007/s00429-019-01855-3 | PubMed

64. **Myers CA,** Vandermosten M, Farris EA, Hancock R, Gimenez P, Black JM, Casto B, Drahos M, Tumber M, Hendren RL, *et al.* (2014) White Matter Morphometric Changes Uniquely Predict Children's Reading Acquisition. *Psychological Science* **25**:1870-1883 https://doi.org/10.1177/0956797614544511 | PubMed

65. **Natu VS,** Arcaro MJ, Barnett MA, Gomez J, Livingstone M, Grill-Spector K, Weiner KS (2021) Sulcal Depth in the Medial Ventral Temporal Cortex Predicts the Location of a Place-Selective Region in Macaques, Children, and Adults. *Cerebral Cortex* **31**:48-61 https://doi.org/10.1093/cercor/bhaa203 | PubMed

66. **Ono M,** Kubik S, Abernathey CD (1990) *Atlas of the cerebral sulci* Thieme.

67. **Papini C,** Palaniyappan L, Kroll J, Froudist-Walsh S, Murray RM, Nosarti C (2020) Altered Cortical Gyrification in Adults Who Were Born Very Preterm and Its Associations With Cognition and Mental Health. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* **5**:640-650 https://doi.org/10.1016/j.bpsc.2020.01.006 | PubMed

68. **Parker BJ,** Voorhies WI, Jiahui G, Miller JA, Willbrand E, Hallock T, Furl N, Garrido L, Duchaine B, Weiner KS (2023) Hominoid-specific sulcal variability is related to face perception ability. *Brain Structure and Function* **228**:677-685 https://doi.org/10.1007/s00429-023-02611-4 | PubMed

69. **Peng P,** Kievit RA (2020) The Development of Academic Achievement and Cognitive Abilities: A Bidirectional Perspective. *Child Development Perspectives* **14**:15-20 https://doi.org/10.1111/cdep.12352 | PubMed

70. **Petrides M** (2018) *Atlas of the morphology of the human cerebral cortex on the average MNI brain* Academic Press.

71. **Pillet I,** Cerrahoğlu B, Philips RV, Dumoulin S, de Beeck HO (2024) The position of visual word forms in the anatomical and representational space of visual categories in occipitotemporal cortex. *Imaging Neuroscience* **2**:1-28 https://doi.org/10.1162/imag_a_00196 | PubMed

72. **Reuter M,** Schmansky NJ, Rosas HD, Fischl B (2012) Within-subject template estimation for unbiased longitudinal image analysis. *NeuroImage* **61**:1402-1418 https://doi.org/10.1016/J.NEUROIMAGE.2012.02.084 | PubMed

73. **Reveley C,** Seth AK, Pierpaoli C, Silva AC, Yu D, Saunders RC, Leopold DA, Ye FQ (2015) Superficial white matter fiber systems impede detection of long-range cortical connections in diffusion MR tractography. *Proceedings of the National Academy of Sciences of the United States of America* **112**:E2820-E2828 https://doi.org/10.1073/pnas.1418198112 | PubMed

74. **Rivière D,** Leprince Y, Labra N, Vindas N, Foubet O, Cagna B, Loh KK, Hopkins W, Balzeau A, Mancip M, *et al.* (2022) Browsing Multiple Subjects When the Atlas Adaptation Cannot Be Achieved via a Warping Strategy. *Frontiers in Neuroinformatics* **16**:803934 https://doi.org/10.3389/fninf.2022.803934 | PubMed

75. **Roell M,** Bellon E, Polspoel B, Declercq M, De Smedt B (2022) Investigating the association between variability in sulcal pattern and academic achievement. *Scientific Reports* **12**:12323 https://doi.org/10.1038/s41598-022-15335-y | PubMed

76. **Roell M,** Cachia A, Matejko AA, Houdé O, Ansari D, Borst G (2021) Sulcation of the intraparietal sulcus is related to symbolic but not non-symbolic number skills. *Developmental Cognitive Neuroscience* **51**:100998 https://doi.org/10.1016/j.dcn.2021.100998 | PubMed

77. **Rosenberg MD,** Finn ES (2022) How to establish robust brain–behavior relationships without thousands of individuals. *Nature Neuroscience* https://doi.org/10.1038/s41593-022-01110-9 | PubMed

78. **Rosseel Y** (2012) lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software* **48** https://doi.org/10.18637/jss.v048.i02

79. **Santacroce F,** Cachia A, Fragueiro A, Grande E, Roell M, Baldassarre A, Sestieri C, Committeri G (2024) Human intraparietal sulcal morphology relates to individual differences in language and memory performance. *Communications Biology* **7**:1-8 https://doi.org/10.1038/s42003-024-06175-9 | PubMed

80. **Saygin ZM,** Osher DE, Norton ES, Youssoufian DA, Beach SD, Feather J, Gaab N, Gabrieli JDEE, Kanwisher N (2016) Connectivity precedes function in the development of the visual word form area. *Nature Neuroscience* **19**:1250-1255 https://doi.org/10.1038/nn.4354 | PubMed

81. **Schilling KG,** Blaber J, Hansen C, Cai L, Rogers B, Anderson AW, Smith S, Kanakaraj P, Rex T, Resnick SM, *et al.* (2020) Distortion correction of diffusion weighted MRI without reverse phase-encoding scans or field-maps. *PLoS ONE* **15**:e0236418 https://doi.org/10.1371/journal.pone.0236418 | PubMed

82. **Schmitt S,** Ringwald KG, Meller T, Stein F, Brosch K, Pfarr J.-K, Hahn T, Lemke H, Meinert S, Repple J, *et al.* (2023) Associations of gestational age with gyrification and neurocognition in healthy adults. *European Archives of Psychiatry and Clinical Neuroscience* **273**:467-479 https://doi.org/10.1007/s00406-022-01454-0 | PubMed

83. **Schwizer Ashkenazi S,** Roell M, McCaskey U, Cachia A, Borst G, O'Gorman Tuura R, Kucian K (2024) Are numerical abilities determined at early age? A brain morphology study in children and adolescents with and without developmental dyscalculia. *Developmental Cognitive Neuroscience* **67**:101369 https://doi.org/10.1016/j.dcn.2024.101369 | PubMed

84. **Shinohara H,** Liu X, Nakajima R, Kinoshita M, Ozaki N, Hori O, Nakada M (2020) Pyramid-shape crossings and intercrossing fibers are key elements for construction of the neural network in the superficial white matter of the human cerebrum. *Cerebral Cortex* **30**:5218-5228 https://doi.org/10.1093/cercor/bhaa080 | PubMed

85. **Skeide MA,** Kraft I, Müller B, Schaadt G, Neef NE, Brauer J, Wilcke A, Kirsten H, Boltze J, Friederici AD (2016) NRSN1 associated grey matter volume of the visual word form area reveals dyslexia before school. *Brain* **139**:2792-2803 https://doi.org/10.1093/brain/aww153 | PubMed

86. **Tissier C,** Linzarini A, Allaire-Duquette G, Mevel K, Poirel N, Dollfus S, Etard O, Orliac F, Peyrin C, Charron S, *et al.* (2018) Sulcal polymorphisms of the IFC and ACC contribute to inhibitory control variability in children and adults. *eNeuro* **5** https://doi.org/10.1523/ENEURO.0197-17.2018 | PubMed

87. **Torgesen JK,** Wagner RK, Rashotte CA (1999) *TOWRE: Test of word reading reading efficiency* The Psychological Corporation.

87a. **Ünal ZE,** Greene NR, Lin X, Geary DC (2023) What Is the Source of the Correlation Between Reading and Mathematics Achievement? Two Meta-analytic Studies. *Educational Psychology Review* **35**:4 https://doi.org/10.1007/s10648-023-09717-5

88. **Vandermosten M,** Hoeft F, Norton ES (2016) Integrating MRI brain imaging studies of pre-reading children with current theories of developmental dyslexia: A review and quantitative meta-analysis. *Current Opinion in Behavioral Sciences* **10**:155-161 https://doi.org/10.1016/j.cobeha.2016.06.007 | PubMed

89. **Vandermosten M,** Schevenels K, Economou M, Hoeft F (2020) The influence of intergenerational transfer of white matter tracts on early reading development. *bioRxiv* https://doi.org/10.1101/2020.10.09.333096

90. **Veraart J,** Novikov DS, Christiaens D, Ades-aron B, Sijbers J, Fieremans E (2016) Denoising of diffusion MRI using random matrix theory. *NeuroImage* **142**:394-406 https://doi.org/10.1016/j.neuroimage.2016.08.016 | PubMed

91. **Voorhies WI,** Miller JA, Yao JK, Bunge SA, Weiner KS (2021) Cognitive insights from tertiary sulci in prefrontal cortex. *Nature Communications* **12**:1-14 https://doi.org/10.1038/s41467-021-25162-w | PubMed

92. **Wagner R,** Torgesen J, Rashotte C, Pearson N (1999) *Comprehensive test of phonological processing: CTOPP* Pearson Assessment.

93. **Weiner KS** (2019) The Mid-Fusiform Sulcus (sulcus sagittalis gyri fusiformis). *Anatomical Record* **302**:1491-1503 https://doi.org/10.1002/ar.24041 | PubMed

94. **Weiner KS** (2023) The hypothesis of fundal cognition. *Nature Reviews Neuroscience* **24**:521-521 https://doi.org/10.1038/s41583-023-00725-6 | PubMed

95. **Weiner KS,** Golarai G, Caspers J, Chuapoco MR, Mohlberg H, Zilles K, Amunts K, Grill-Spector K (2014) The mid-fusiform sulcus: A landmark identifying both cytoarchitectonic and functional divisions of human ventral temporal cortex. *NeuroImage* **84**:453-465 https://doi.org/10.1016/j.neuroimage.2013.08.068 | PubMed

96. **Weiner KS,** Yeatman JD, Wandell BA (2017) The posterior arcuate fasciculus and the vertical occipital fasciculus. *Cortex* **97**:274-276 https://doi.org/10.1016/j.cortex.2016.03.012 | PubMed

97. **Wickham H** (2016) ggplot2: Elegant Graphics for Data Analysis. https://ggplot2.tidyverse.org

98. **Willbrand EH,** Voorhies WI, Yao JK, Weiner KS, Bunge SA (2022) Presence or absence of a prefrontal sulcus is linked to reasoning performance during child development. *Brain Structure and Function* **227**:2543-2551 https://doi.org/10.1007/s00429-022-02539-1 | PubMed

99. **Wilson SM,** Rising K, Stib MT, Rapcsak SZ, Beeson PM (2013) Dysfunctional visual word form processing in progressive alexia. *Brain* **136**:1260-1273 https://doi.org/10.1093/brain/awt034 | PubMed

100. **Winkler AM,** Ridgway GR, Webster MA, Smith SM, Nichols TE (2014) Permutation inference for the general linear model. *NeuroImage* **92**:381-397 https://doi.org/10.1016/j.neuroimage.2014.01.060 | PubMed

101. **Wolf M,** Denckla M (2005) RAN/RAS: Rapid automatized naming and rapid alternating stimulus tests. https://www.proedinc.com/Products/10435/ranras-rapid-automatized-naming-and-rapid-altern.aspx

102. **Woodcock RW** (1998) *Woodcock Reading Mastery Tests-Revised Normative Update: Examiner's manual.* Circle Pines, MN: American Guidance Service.

103. **Woodcock RW,** Mather N, McGrew KS, Wendling BJ (2001) Woodcock-Johnson III Tests of Cognitive Abilities.

104. **Woodcock RW,** McGrew KS, Mather N (2001) Woodcock-Johnson III Tests of Achievement.

105. **Woolnough O,** Tandon N (2024) Dissociation of reading and naming in ventral occipitotemporal cortex. *Brain* **147**:2522-2529 https://doi.org/10.1093/brain/awae027 | PubMed

106. **Xia Z,** Wang C, Hancock R, Vandermosten M, Hoeft F (2021) Development of thalamus mediates paternal age effect on offspring reading: A preliminary investigation. *Human Brain Mapping* **42**:4580-4596 https://doi.org/10.1002/hbm.25567 | PubMed

107. **Yamagata B,** Murayama K, Black JM, Hancock R, Mimura M, Yang TT, Reiss AL, Hoeft F (2016) Female-specific intergenerational transmission patterns of the human corticolimbic circuitry. *Journal of Neuroscience* **36**:1254-1260 https://doi.org/10.1523/JNEUROSCI.4974-14.2016 | PubMed

108. **Yao JK,** Mitchell JL, Davison A, Yeatman JD (2026) Sulcal anatomy of ventral temporal cortex and reading development. *bioRxiv* https://doi.org/10.64898/2026.04.06.716640

109. **Yao JK,** Voorhies WI, Miller JA, Bunge SA, Weiner KS (2022) Sulcal depth in prefrontal cortex: A novel predictor of working memory performance. *Cerebral Cortex* https://doi.org/10.1093/cercor/bhac173 | PubMed

110. **Yeatman JD,** Dougherty RF, Ben-Shachar M, Wandell BA (2012) Development of white matter and reading skills. *Proceedings of the National Academy of Sciences* **109**:E3045-E3053 https://doi.org/10.1073/pnas.1206792109 | PubMed

111. **Zaretskaya N,** Fischl B, Reuter M, Renvall V, Polimeni JR (2018) Advantages of cortical surface reconstruction using submillimeter 7 T MEMPRAGE. *NeuroImage* **165** https://doi.org/10.1016/j.neuroimage.2017.09.060 | PubMed

112. **Zemmoura I,** Herbet G, Moritz-Gasser S, Duffau H (2015) New insights into the neural network mediating reading processes provided by cortico-subcortical electrical mapping. *Human Brain Mapping* **36**:2215-2230 https://doi.org/10.1002/hbm.22766 | PubMed

113. **Zhan M,** Pallier C, Agrawal A, Dehaene S, Cohen L (2023) Does the visual word form area split in bilingual readers? A millimeter-scale 7-T fMRI study. *Science Advances* **9**:eadf6140 https://doi.org/10.1126/sciadv.adf6140 | PubMed

114. **Zhang T,** Li X, Jiang X, Ge F, Zhang S, Zhao L, Liu H, Huang Y, Wang X, Yang J, *et al.* (2020) Cortical 3-hinges could serve as hubs in cortico-cortical connective network. *Brain Imaging and Behavior* **14**:2512-2529 https://doi.org/10.1007/s11682-019-00204-6 | PubMed

Peer reviews

Reviewer #1 (Public review):

[Editors' note: this version has been assessed by the Reviewing Editor without further input from the original reviewers. The authors have addressed the comments raised in the previous round of review.]

Summary:

There is prior literature showing a robust relationship between sulcal interruptions in the posterior occipital temporal sulcus (pOTS) and reading ability. The goals of this study were to extend these findings to children examined longitudinally as they become better readers, and to examine the underlying white matter properties in individuals with and without pOTS sulcal interruptions. To do this, the authors collected longitudinal structural, diffusion, and behavioral data in 51 children (TP1 age 5.5, TP3 age 8.2 years).

First, the authors found that the gyral gap was consistent across time within the subject. This is expected, as they state in the introduction that sulcal patterns are typically established in utero. Next, they found that children with an interrupted pOTS have higher reading scores (across a variety of measures) at timepoint (TP) 3 than children with continuous pOTS, and this was specific to the pOTS, as no associations emerged for the anterior OTS or MFS; this is again expected from prior literature. They then found that the binary presence of this gap, but not anterior OTS or MFS predicted T3 reading performance. Further, they found that a subsample of the lowest readers at TP1 did not have differences in reading score by gyral gap, but that this difference emerged at TP3. Additionally, the gyral gap at TP1 is similar to variance TOWRE 3 reading skills as some behavioral measures at TP1. Examining underlying white matter in a smaller subset of children, the authors found higher MD in children with an interrupted pOTS vs. those with a continuous pOTS, which was contrary to their hypothesis, and higher local connectivity for interrupted, aligning with their hypothesis, but this difference was no longer present when accounting for TP3 reading scores. The authors conclude that structural properties, in this case, the gyral gap, may guide neural plasticity for reading.

Strengths:

This paper has an interesting set of longitudinal data to examine the perhaps changing relationship between sulcal interruptions in the pOTS with reading scores. I commend the

authors on data collection and attention to detail in the anatomical analyses.

https://doi.org/10.7554/eLife.103007.2.sa2

Reviewer #2 (Public review):

Summary:

This manuscript examined the impact of sulcal morphology on reading development. A very specific feature on the ventral surface of the brain was identified, namely the presence of an interruption in the posterior portion of the left occipitotemporal sulcus (pOTS). Compared to children with a continuous pOTS, children with an interruption at age 5 years had better reading ability at age 8. This was a large effect measured in 43 children. Surprisingly, this morphological feature was a better predictor of reading ability than measures of pre-literacy cognitive skills, such as phonological awareness. The effect was tested and reproduced across several different measures of reading ability. The authors hypothesised that the mechanism underlying this benefit related to greater local connectivity, which confers a computational advantage. This was demonstrated using analysis of diffusion-weighted imaging data in 29 of the children obtained at age 8.

Strengths:

The novelty of the manuscript is threefold: (i) the measure was made in children who were pre-literate (previous work was in older children and adults); (ii) longitudinal brain imaging and behavioural data were analysed; and (iii) diffusion data were analysed to test a hypothesis about the underlying mechanism.

The manuscript is exceptionally well written. The methods are detailed and easily reproduced. The approach is thoughtful and meticulous. All possible alternatives appear to have been considered. Where possible, further analyses have been done to address these alternatives. For example, the testing of the specificity of the sulcal interruption to left pOTS was an important addition. None predicted reading skills.

https://doi.org/10.7554/eLife.103007.2.sa1

Author response:

The following is the authors' response to the original reviews.

> ***Public Reviews:***
>
> ***Reviewer #1 (Public review):***
>
> *Summary:*
>
> *There is prior literature showing a robust relationship between sulcal interruptions in the posterior occipital temporal sulcus (pOTS) and reading ability. The goals of this study were to extend these findings to children examined longitudinally as they become better readers, and to examine the underlying white matter properties in individuals with and without pOTS sulcal interruptions. To do this, the authors collected longitudinal structural, diffusion, and behavioral data in 51 children (TP1 age 5.5, TP3 age 8.2 years).*
>
> *First, the authors found that the gyral gap was consistent across time within the subject. This is expected, as they state in the introduction that sulcal patterns are typically established in utero. Next, they found that children with an interrupted pOTS have higher reading scores (across a variety of measures) at timepoint (TP) 3 than children with continuous pOTS, and this was specific to the pOTS, as no associations emerged for the*

anterior OTS or MFS; this is again expected from prior literature. They then found that the binary presence of this gap, but not anterior OTS or MFS predicted T3 reading performance. Further, they found that a subsample of the lowest readers at TP1 did not have differences in reading score by gyral gap, but that this difference emerged at TP3. Additionally, the gyral gap at TP1 is similar to variance TOWRE 3 reading skills as some behavioral measures at TP1. Examining underlying white matter in a smaller subset of children, the authors found higher MD in children with an interrupted pOTS vs. those with a continuous pOTS, which was contrary to their hypothesis, and higher local connectivity for interrupted, aligning with their hypothesis, but this difference was no longer present when accounting for TP3 reading scores. The authors conclude that structural properties, in this case, the gyral gap, may guide neural plasticity for reading.

Strengths:

This paper has an interesting set of longitudinal data to examine the perhaps changing relationship between sulcal interruptions in the pOTS with reading scores. I commend the authors on data collection and attention to detail in the anatomical analyses.

Weaknesses:

However, my enthusiasm was somewhat dampened after finding numerous prior publications on this very topic and I'm unclear as to how much more this paper adds to the current literature. Would we expect the existence of sulcal interruptions to be aligned with reading skills in older kids but not younger kids? Is the point to see if the interruptions exist prior to reading (but these children are not really prereaders)? What is the alternative- why would these interruptions not exist? After all, this anatomy is determined prenatally. Children who have pOTS interruptions at T1 should also have these interruptions at T3 (and indeed that is what the authors find). So how can this be the mechanism that drives plasticity? The authors also talk about the neuronal recycling hypothesis but their data cannot speak to this because they do not have fMRI data nor does their sample include only prereaders with no reading experience. The conclusions are overall overstated and not supported by the results. I think this paper could add interesting knowledge for the specific subfield of reading and the brain. However, the current state of the results, especially with the inclusion of so many trending results and the comparison of so many different processing pipelines and models, in addition to a conclusion that is not motivated by the work makes it difficult to appreciate the paper.

We thank Reviewer 1 for their detailed feedback. The reviewer raises several related concerns regarding the novelty and framing of the study, the interpretation of stable sulcal anatomy, early reading abilities, statistical robustness, and the proposed link between pOTS morphology, white matter, and reading development. In response, we have revised the manuscript to make clearer that the main contribution is not the stability of pOTS morphology itself, nor simply the replication of previously reported cross-sectional associations with reading, but rather how this stable anatomical feature relates prospectively to reading acquisition from the onset of literacy instruction. In particular, we now make more explicit that left pOTS interruption contributes to later reading beyond associations already present at age 5, including through an updated serial mediation analysis, and we better highlight how this relationship develops over time.

The reviewer's comments also prompted us to explain more clearly how the complementary analyses in the full sample and in the subgroup with very limited initial reading skills converge to support our conclusion that left pOTS interruption is a particularly strong longitudinal predictor of later reading relative to the benchmark behavioral and familial predictors considered, despite the collinearity between pOTS interruption and emerging literacy skills at T1. We have renamed this subgroup more accurately from "pre-readers" to "initial minimal readers" and clarified its characteristics and range of reading abilities.

We also moved key analyses from the Supplement to Figure 3 to make the distinction between separate and unique longitudinal contributions easier to follow. The initial-minimal-reader analyses provide a complementary assessment in which confounding with early reading is substantially reduced. At the same time, as the reviewer accurately pointed out, some of our previous wording implied a more categorical superiority than the analyses warrant; we have therefore softened these claims throughout the manuscript while preserving the conclusion that pOTS interruption is among the strongest longitudinal contributors considered. We also added, in response to Reviewer 2, a new systematic analysis supporting the relative cognitive specificity of the pOTS effect.

Finally, we made several interpretations more precise, notably by presenting the white-matter findings as a candidate pathway rather than a definitive mechanism and by removing implications that our data directly test the neuronal recycling hypothesis. We believe these revisions make the longitudinal contribution, scope, and significance of the findings considerably clearer, and we are grateful to the reviewer for prompting these improvements.

> ***Reviewer #2 (Public review):***
>
> *Summary:*
>
> *This manuscript examined the impact of sulcal morphology on reading development. A very specific feature on the ventral surface of the brain was identified, namely the presence of an interruption in the posterior portion of the left occipitotemporal sulcus (pOTS). Compared to children with a continuous pOTS, children with an interruption at age 5 years had better reading ability at age 8. This was a large effect measured in 43 children. Surprisingly, this morphological feature was a better predictor of reading ability than measures of pre-literacy cognitive skills, such as phonological awareness. The effect was tested and reproduced across several different measures of reading ability. The authors hypothesised that the mechanism underlying this benefit related to greater local connectivity, which confers a computational advantage. This was demonstrated using analysis of diffusion-weighted imaging data in 29 of the children obtained at age 8.*
>
> *Strengths:*
>
> *The novelty of the manuscript is threefold: (i) the measure was made in children who were pre-literate (previous work was in older children and adults); (ii) longitudinal brain imaging and behavioural data were analysed; and (iii) diffusion data were analysed to test a hypothesis about the underlying mechanism.*
>
> *The manuscript is exceptionally well written. The methods are detailed and easily reproduced. The approach is thoughtful and meticulous. All possible alternatives appear to have been considered. Where possible, further analyses have been done to address these alternatives. For example, the testing of the specificity of the sulcal interruption to left pOTS was an important addition. None predicted reading skills.*
>
> *Weaknesses:*
>
> *The correlation of the interruption with all kinds of literacy measures and in particular reading comprehension and then PIQ suggest this interruption might confer a more general cognitive advantage rather than specifically a reading one.*
>
> *It would be interesting to know if the anatomical difference predicts any other cognitive ability or if there might be any cognitive cost (a negative correlation) of this sulcal interruption.*

> *The location of the interruption in the sulcus is quite variable and in some cases, there is more than one interruption. The sample size is probably not big enough to compare these different patterns or to evaluate the influence of the location of the sulcal interruption.*
>
> *The sample is quite high-functioning and the generalisability of the findings outside of this specific population is inevitably limited.*

We would like to thank Reviewer 2 for their positive feedback and for their insightful suggestions.

Cognitive specificity: We agree with the reviewer that the association between left pOTS interruption and performance IQ, as measured in the concept formation learning task, on top of its association with multiple measures of literacy, raises the possibility that this sulcal feature relates to general cognition more broadly. The fact that reading and general cognitive abilities are highly correlated and mutually influence one another over development complicates any attempt to interpret the effect of left pOTS interruption as purely reading-specific. Our results however support a relative specificity of this sulcal feature to reading: left pOTS interruption remains a strong longitudinal predictor of later reading, even when controlling for the moderate association with performance IQ, and its impact on reading skills is larger than on pIQ. Importantly, a new systematic analysis of associations with all the cognitive and academic skills obtained at T3 further supports a relative cognitive specificity to reading: We show that while left pOTS interruption is also significantly associated with several assessments of math, language and general cognition at age 8 (possibly due to the inherent correlations across these domains), hierarchical clustering indicates that pOTS interruption aligns most closely with literacy-related outcomes (cf. See rev 2 recommendation comment (1)). We observe no cognitive costs of left pOTS interruption across the many linguistic, pre-literacy and cognitive domains tested.

Neuroanatomical precision: We also fully agree that our sulcal marker of interest, i.e., whether the posterior portion of the left pOTS is interrupted or not, is relatively coarse and may lack in neuroanatomical precision (though this robustness may alternatively be seen as a strength). Beyond this first approximation, several questions remain unanswered, including whether additional interruptions confer an additional reading advantage, or whether interruptions in more specific locations of the left posterior OTS would be optimal. Unfortunately, multiple interruptions were too rare and gap locations too unevenly distributed in our cohort to support adequately powered comparisons. These questions can only be answered on larger sample sizes, which we now mention in our revised Discussion. Please refer to See rev 2 recommendation comment (3) for more details and for the results of tentative explorations which did not support clear trends.

Sample limitations: The characteristics of our sample (being both from a high socio-economic background, as often seen in pediatric research, and high-functioning) indeed limit the generalizability of our results, as we now acknowledge in the Discussion. However, we replicate results by Cachia and colleagues (2018) which were obtained in a sample where half of the participants were from a rural and (very) low socio-economic background, in order to match a sample of illiterate adults, and this association has been now found in multiple samples with various SES and cognitive characteristics. This suggests that overall such associations may be generalizable across extremely varied socio-economic backgrounds and also possibly along the spectrum of general cognitive skills (cf. R2 comment-toaddress #5).

Additional correction made during revision:

During the revision process, we identified one subject for whom an anterior OTS gap had been incorrectly included in the diffusion analyses (participant 6007 on previous

Supplementary Figure S6). We removed this gap, reran all diffusion analyses, and updated the relevant results and figures. The results and conclusions remained essentially unchanged.

> ***Recommendations for the authors:***
>
> ***Reviewer #1 (Recommendations for the authors):***
>
> *Comment (1) The results are currently framed in a manner that makes it difficult to figure out what the main findings of the paper are, and to see if those findings support their conclusions.*
>
> *The beginning of the results is focused on showing that the longitudinal pipeline has a better consistency across time points, yet the authors still report follow-up analyses with the cross-sectional pipeline, even though those results are only trending. I think it would make more sense to only focus on the longitudinal pipeline because that is the novelty of this manuscript.*

We fully agree with Reviewer 1 that the longitudinal pipeline shows better consistency and likely constitutes a more accurate, less noisy measure of sulcal interruptions, as we report. For this reason, all the results reported in the main figures concern the longitudinally processed data only.

Yet, we choose to also include and report supplementary cross-sectional analyses since these analyses, while less reliable, are important for several conceptual reasons:

(1) Since we wanted to demonstrate that T1 brain anatomy predicted T3 behavior longitudinally, it seemed important as a proof of principle to show that this prediction held with T3 brain data completely unseen, i.e., with T1 cross-sectional processing only. We have now clarified the text referring to this analysis (lines 192-196):

“This importantly implies that left pOTS morphology at T1 was still a significant longitudinal predictor of T3 reading behavior when T3 brain images were not used to improve T1 anatomical segmentations, i.e., with T1 cross-sectional processing alone, despite noisier tissue segmentations with cross-sectional processing (see also Reuter et al., 2012).”

(2) Since most studies are cross-sectional, reporting cross-sectional effects may constitute a more relevant reference for other studies, and for the generalizability of our results.

(3) Larger samples were available for the cross-sectional processing, allowing to test a maximum number of children, despite the higher noise in the sulcal measurements.

Please also note that across the 4 types of processing (2 timepoints x cross-sectional & longitudinal pipelines), only pOTS interruptions measured cross-sectionally at T3 showed a marginal association with reading scores, while the other 3 processings showed a large significant effect. Having established that cross-sectional measures are less reliable, and since our main goal was to demonstrate that T1 morphology predicted T3 reading, such trending results are found in the least reliable and least relevant condition, with otherwise large and significant effects, which seems quite acceptable. If anything, we believe that transparently reporting these consistent results altogether support, rather than challenge, the robustness of the association.

Our main findings are summarized in the last paragraph of the Introduction as well as in the first paragraph of our Discussion.

> *Comment (2) Also, as the authors note, it is known that sulcal patterns develop in utero; therefore, we would expect a perfect alignment between TP1 and TP3- there should be a more in-depth analysis of the lack of alignment in the main results. And further, this is more reason that the results should focus on the longitudinal pipeline which is closer to*

> *that expectation. However, it is unclear whether T1 data processed through the longitudinal pipeline preserves anatomical nuances of that time point (since the images are averaged across the timepoints). Maybe the cross-sectional pipeline is a better depiction of how anatomy changes as a child develops.*

We thank Reviewer 1 for raising these relevant questions.

First, we would like to clarify that T1 longitudinal processing does preserve anatomical nuances of this time point. In fact, the averaging of T1 and T3 MRIs into a subject base is only an intermediate step of the longitudinal pipeline, which then crucially allows to re-initialize T1 MRI and T3 MRI tissue segmentations separately from their respective scans, but using this time the common tissue priors provided by the subject base (hence the reference to distinct "T1 longitudinal" and "T3 longitudinal" processing). We describe this in the Methods (lines 764-769):

"When images were available at both Time 1 and Time 3 for a child (N=43), the longitudinal Freesurfer pipeline was used, which first creates an unbiased subject average across time points ("base") and then restarts multiple processing steps at each time point separately from the base template in order to refine tissue segmentations on each image (respectively T1 and T3 longitudinally processed images, Reuter et al., 2012; see Figure 1A)."

Therefore, a perfect match between T1 and T3 anatomy is not expected, even if we could expect (and indeed find) pOTS morphology, and interruption patterns in particular, to be extremely stable. Actually, our left pOTS interruption classification only changes between the T1 and T3 longitudinal processing for one child, with two left pOTS sulcal components separated by a distance estimated as 1.34mm at T1 and 0.99mm at T3, leading to a different classification given our 1 mm minimum-distance interruption criterion. We added the following sentence accordingly (lines 146-149):

"Only one child changed left pOTS interruption classification between T1 and T3 under longitudinal processing, due to the measured inter-component distance crossing the predefined 1-mm threshold, consistent with high stability of this macro-anatomical feature."

Our results are therefore consistent with the view that sulcal interruption patterns are set in utero and stable after birth. Given that there actually is limited direct evidence to this effect, we chose to confirm this stability as a first step, which we refer to in the Discussion (lines 606-610):

"In our study, we observed high reliability of the pOTS interruption from age 5 (T1) to age 8 (T3), which is consistent with the longitudinal stability of interruptions of (i) the intraparietal sulcus and postcentral sulcus (Schwizer Ashkenazi et al., 2024), (ii) sulci in inferior frontal cortex (Tissier et al., 2018), and (iii) patterns of the anterior cingulate sulcus (Cachia et al., 2016) observed in late childhood (approximately 8 years onwards) to early adulthood."

As discussed above, the longitudinal processing pipeline leads to better precision in assessing neuroanatomy, as well as its change across time (see the seminal paper by Reuter et al., 2012), which is why we mostly use measures derived from the longitudinal processing.

> *Comment (3) The analyses starting on p.4 ("Left pOTS interruption is prospectively associated with better reading skills") seem overly fixated on cross-sectional vs. longitudinal processing pipelines. The key comparisons should be on anatomical specificity (posterior or anterior OTS, etc) as well as including T1 reading skills as a regressor to see if the interruptions at T1 explain additional variance in T3 skills.*

We thank Reviewer 1 for highlighting that we had started our result paragraph by a verification that was less essential. We have now reordered our results to highlight those concerning the anatomical specificity to the left pOTS, in the analyses reported under "The

association between left pOTS interruption and reading skills was robust to different measurements of reading skills and sulcal interruption in four ways" (line 175 onward).

Testing whether the interruptions at T1 explain additional variance in T3 skills over and above T1 reading skills is so essential that it is actually the object of the entire next subsection "The effect of left pOTS interruption on reading accumulates over time" (lines 235-274). This claim is supported by the following findings:

(1) In the subsample of minimal readers with little to no reading experience at T1, we see that despite the absence of associations between left pOTS interruptions and reading at T1, the association is present and significant at T2 and T3 and appears to grow (lines 243-253 and Supplementary Figure 4);

(2) In the serial mediation analysis across the whole sample that follows, we observe important and significant additional contributions of left pOTS on later reading skills beyond the contributions at T1. We have now clarified this essential point and rephrased these results, notably adding this particular sentence:

"Importantly, there were also significant contributions of left pOTS interruption onto T3 reading on top of those mediated through T1, with large cumulative effects through T2 and directly onto T3 (d*c+f: β = 4.580, SE = 1.757, p = .009), while indirect effect through T1 TOWRE scores alone were non-significant (a*b*c+a*e: β = 0.971, SE = 0.839, p = .25)." '

We thank the reviewer for bringing to our attention that this was lacking in the previous version and needed to be made explicit. (Please note that p-values for the serial mediation analysis slightly changed, as the bootstrap was rerun to estimate this additional path.)

The analyses in the following section "Left pOTS interruption makes a strong longitudinal contribution relative to established predictors of reading" also support the same conclusion that left pOTS interruptions explain variance in T3 reading over and above T1 reading, despite the concern of collinearity with early literacy skills in our entire sample, as we explain below in details.

> *Comment (4) The authors do a version of this on p. 6 lines 269-273 and show that indeed, pOTS binary classification contributes a very similar proportion of the variance in T3 reading scores as the next best predictor (letter knowledge). They do not seem to be significantly different- are they?*

When longitudinal contributions are tested separately, left pOTS interruption accounts for 20.5% of T3 reading variance while T1 letter knowledge accounts for 19.6% of this variance. While these values appear numerically close, one should first bear in mind that, as an emerging literacy measure rather than a pre-literacy predictor from a different domain, letter knowledge is expected to be a strong predictor of future literacy skills. The fact that left pOTS interruption, a much less proximal neuroanatomical marker, explains slightly more variance in reading stands out as a more interesting observation.

More importantly, while proportions of explained variance cannot be compared statistically in a straightforward fashion, the collinearity between these two predictors across the full sample anyway completely biases this comparison: because left pOTS interruption already accounts for significant variance in letter knowledge (cf. Figure 3B and analyses reported in lines 320-326), the estimation of the longitudinal contributions of letter knowledge to reading skills is inflated by this collinearity, picking up in addition some of the true effects of sulcal interruption on T3 reading (and vice versa).

For that reason, our claim that left pOTS interruption is a strong longitudinal predictor of reading relative to pre-literacy skills actually rests on the integration of all of the results reported in the corresponding section (lines 275-367), including: the contributions of each

predictor separately under discussion, in addition also their unique contributions above all other predictors, and the analysis of the underlying collinearity and how it reduces the power in testing the true effects of pOTS interruption, as well as the same analyses repeated in the subsample of minimal initial readers where the collinearity between pOTS and emerging literacy is removed. (Please note that in this subsample, letter knowledge has minimal nonsignificant contributions to T3 reading, contrarily to sulcal interruptions - cf. Figure 3D and F and Comment-to-address #6).

To clarify this entire section and how all these results converge to support our conclusion, we have now moved several supplementary figure panels from Supplementary Figure 5 to Figure 3 (also addressing See rev 2 recommendation comment (9)) and largely edited the text, including:

"These significant associations between left pOTS interruption, performance IQ and pre-literacy skills (Figure 3B and D), as well as potential associations between pre-literacy skills themselves (cf. correlation matrices in Supplementary Figure 5D and H), make it challenging to fairly assess the separate unique contributions of each of these variables to the longitudinal prediction of reading skills. For instance, because left pOTS interruption explains a significant amount of variance at T1 in letter knowledge, early TOWRE scores and performance IQ, some of its true direct effect on future T3 reading skills is already captured in a baseline model including these cognitive predictors only, even before left pOTS interruption is explicitly entered as a predictor. Yet, even when controlling in a such a stringent way for all pre-literacy skills, early reading abilities, family history of dyslexia and their shared variance with left pOTS interruption in such nested model comparison, left pOTS interruption still had a marginal additional contribution to predicting T3 TOWRE reading scores (ΔR^2 = 5.0%, F(1, 31) = 3.74, p = .062) inspite of this important collinearity. Rapid naming and family history were the only two other predictors that outperformed left pOTS interruption in terms of unique contributions to T3 reading (ΔR^2 = 7.7% and 6.6% for RAN and mother ARHQ respectively, see Figure 3C), possibly because these predictors were among the least confounded with others (see Figure S5D for correlations among predictors) and thus among the least affected by collinearity. Letter knowledge, in contrast, had the smallest, non-significant unique contributions to T3 reading (ΔR^2 = 0.61%, p = .50)."

Thus, while the separate contributions of letter knowledge to T3 reading was almost as large as that of left pOTS interruption, their unique contributions above each other and other predictors strikingly differed: ΔR^2 = 5.0% for left pOTS interruption vs. 0.61% for letter knowledge.

Please also note that our statistical approach is particularly stringent: our nested model comparisons evaluated the unique contributions of each variable above all other predictors, i.e. after accounting for the variance in the outcome jointly explained by the covariance between predictors. In such cases of highly collinear variables, statistical power is importantly reduced.

> *Comment (5) And again even if the T1 presence of pOTS interruption explains this additional variance, what does that tell us? The brain (as measured by this pOTS measure) does not change from T1 to T3, as evidenced by this result: Lines 178-180: "left pOTS morphology at T1 remained a significant longitudinal predictor of T3 reading without taking into account brain information at T3, i.e., with T1 cross-sectional processing alone." T3 anatomical data is therefore redundant to this fit. This is further confirmed by the results in lines 220-226, showing differences in T1 TOWRE scores between children with continuous vs. interrupted pOTS. Additilly, the finding that "the effect of pOTS interruption accumulates over time" may instead suggest that TOWRE skills are more accurately measured at T3, especially for the lowest scores. It had alreaonady been shown that pOTS interruptions correlate with reading skills in readers-*

> *paired with assessment issues/lack of norms at T1 of these reading measures and the lack of change in pOTS over time, these results are not surprising.*

We fully agree with the reviewer that given the previously observed associations between left OTS interruptions and reading skills in 10-year-olds and adults (Borst et al., 2016; Cachia et al., 2018) and the general understanding that sulcal patterns are stable after birth, it was natural to expect that left OTS patterns observed before or at the onset of literacy acquisition would predict future reading skills. Yet, validating hypotheses -however natural they may seem- is essential for moving science forward.

Beyond this expected longitudinal association, the fact that left pOTS interruption at T1 predicts additional variance in T3 reading above early literacy skills and all other cognitive, linguistic and familial predictors of literacy also tells us new and important things. Notably:

(1) It shows that this sulcal marker is a large and important predictor of future reading skills, with a predictive performance comparable to or exceeding the conservative benchmark of pre-literacy skills and familial predictors used for the identification of children at risk of dyslexia. This highlights the relevance of this neuroanatomical marker at a theoretical level for reading acquisition, with possible practical implications for early identification and intervention further down the line.

(2) We learn that the association between left pOTS morphology and reading is not merely a byproduct of pre-existing associations between sulcal morphology and cognitive or familial factors before literacy learning. Instead, our results clarify the neurocognitive origins of the relationship between reading and OTS morphology by showing a prospective and sustained association with reading acquisition between sulcal morphology and reading acquisition during literacy learning.

Concerning the relevance of T1 vs. T3 brain data, please refer to our response to Comment-to-address #2, in which we clarify that the longitudinal pipeline serves to improve the reliability of the assessment of sulcal interruptions at T1, but that our results are indeed also confirmed using only T1 brain data.

Concerning the matter of measurement accuracy of the TOWRE at T1, we agree that reading is harder to measure at the very beginning of literacy acquisition, and that the TOWRE is not normed below age 6. For this reason, our analyses controlled for early literacy skills at T1 using both scores of letter knowledge, as measured by the age-appropriate and normed WMRT letter identification subtest, and the principal component of the two TOWRE raw scores (scaled). This ensured a strict control procedure for all measurable emerging literacy skills: letter knowledge, but also word reading and phonological decoding (if already acquired in precocious/early readers, or to confirm this was not acquired yet).

In the absence of TOWRE norms for all children at T1, we used this TOWRE principal component to derive a T1 TOWRE population-scaled index. This scaling mostly served to ensure that this early reading score would be of a similar nature and scale as the T2 and T3 TOWRE standard scores, notably for statistical and plotting purposes (Supplementary Figure 4A and D).

Interestingly, we observed in the full longitudinal sample that left pOTS interruption was already associated around age 5 with this T1 TOWRE score, in addition to the letter identification scores. In the context of stronger associations observed at T2 and T3, this suggests that the T1 TOWRE score, even if not normed nor very accurate, does still capture relevant variance in early literacy skills.

> *Comment (6) Further, while one main conclusion is that the gyral gap predicts T3 reading ability better than typical behavioral TP1 measures (but it really does not appear to), this result is mixed in with a series of nonsignificant results (labeled as marginal or trending,*

> *but p>0.05) and difficult to parse sections. Some, but not all, analyses are performed in a set of 'pre-readers' (defined as scoring <95th percentile on a reading test which is not particularly stringent), and at some points, the authors get a significant result in the whole sample, but not the prereaders and interpret these results as only trending because of a smaller sample size (line 287), but at other times they get a trending result in the main group, and a significant result in the pre-reader group (line 303), and choose to focus on the result of the prereaders. This inconsistency makes me skeptical of why some analyses were done with different samples. The definition of a pre-reader is not really a pre-reader but rather just below an average reader. This also leads back to my question above on the underlying hypotheses, which critically needs to focus on prereaders. Can you further explain the decision to run some analysis in pre-readers (example line 285/286) and further explain what it means to be a pre-reader in this sample? Perhaps try to think of a more direct way to examine the impact of the gyral gap at TP1 predicting behavioral performance at TP3, as a trending result in the whole sample followed by significance in just a subset is not convincing (line 298, 303).*

We thank Reviewer 1 for this opportunity to clarify our results and some misunderstandings.

We acknowledge, and are sorry for, the lack of clarity of our previous explanations, especially as some analyses are complex and some conclusions require integrating results across multiple analyses to understand the overall implications. We believe that our statistical approach was justified, though perhaps not well explained in the original version. We have now clarified and addressed these issues in the revision by renaming pre-readers to 'minimal initial readers' - which is indeed more accurate (thank you for raising this concern) -, better explaining the composition of this subsample and what analyses restricted to this subsample bring (notably lines 243-253, and in the descriptive demographic Table 1), rewriting the entire section "Left pOTS interruption Left pOTS interruption makes a strong longitudinal contribution relative to established predictors of reading" (lines 275-367), and adding a number of figure panels to Figure 3 (previously in supplementary figure 5), to aid the comparison and integration of results across both samples to support our conclusion. To answer directly the reviewer's questions and remarks:

Why two groups? We carry out analyses both in our entire available sample and in a subsample with little to no early literacy skills, since these samples are both relevant for our hypotheses and offer different advantages:

The entire sample allows to maximize sample size, and is more representative of the general population. This naturally implies that some kindergarteners have emerging literacy skills, including letter knowledge, knowledge of a few sight words and minimal decoding. Interestingly, we found these early literacy measures to already be significantly associated with our sulcal marker, but two problems also derive from this: (1) These emerging reading skills preclude from directly testing in this sample our strong hypothesis that pOTS interruption before literary learning predicts future reading scores. (2) The collinearity between the T1 behavioral predictors and left pOTS interruption is a statistical obstacle to testing reliably how each predictor truly contributes to predicting future reading skills separately (as explained in our answer to Reviewer 1's Comment-to-address #4 above, taking the comparison with letter knowledge as a key example).

To remove the association between pOTS and T1 reading and bypass both these problems, we perform additional analyses in a decently-sized subsample of children with little to no reading abilities at T1 - that we have now renamed more accurately as 'initial minimal readers'.

Definition and composition of our subsample of "initial minimal readers":

Kindergarteners are expected to know letters and how to write their names, they may have understood the basics of the alphabetic principle and phonological decoding, and may have learned to recognize a few highly frequent sight words through repeated exposure in their books and environments (e.g., “the”, “and”). From our entire sample, including a number of children accordingly able to read some sight words and decode some syllables and pseudowords (cf. Table 1), we defined a subsample of lowest readers by using a cut-off of a TOWRE population-scaled index at T1 under 95 (due to the lack of norms, cf. see rev 1 recommendation comment (5)). For reference, the equivalent standard score of 95 would correspond to a 37th percentile in a normal distribution.

This cut-off was chosen as it translated to very little actual reading abilities overall and crucially removed the statistical association between pOTS and T1 reading needed to address our two problems, while being still decently sized to allow for our statistics (n = 24, i.e. 56% of the sample, cf. distributions on Supplementary Figure 4A). We now specify more clearly notably in the results lines 244-253:

“We conducted a supplementary analysis in the subset of 24 children with the lowest reading scores at T1 (defined as those with a TOWRE T1 population-derived composite score equivalent below 95). Hereafter, we refer to these children as initial minimal readers: 79% could not decode a single syllable or pseudoword in 45 s, with a maximum of four pseudowords decoded, and 63% could not recognize a single frequent sight word, with a maximum of seven words recognized. Although these children showed no difference in T1 reading scores as a function of whether their left pOTS was interrupted or continuous ($F(1, 19) = 0.21$, $p = .66$, $\Delta R^2 = 0.82\%$, Cohen’s $d = 0.034$), [...] “

and methods - lines 988-991:

“We therefore selected a subsample of children whose TOWRE standard score equivalent at Time 1 was below 95, resulting in a subset of 24 initial pre-readers (median of 0 words read in 45s in SWE, range: 0-7; median of 0 syllables in 45s in PDE, range: 0-4; see Table 1).” as well as Table 1.”

Expecting a larger proportion of children of this age with no word reading or decoding at all would be both unnecessary for our analyses and developmentally unrealistic, especially in this high-SES sample, selecting an atypical population at especially high risk of dyslexia.

Discrepancies between statistics in the full sample and the subsample:

Because this subsample was chosen to remove the associations at T1 between pOTS and early literacy which affected some analyses (cf. Comment #4), some discrepancies in the results across the two samples are actually expected. Conducting statistics on 56% of the full sample may also affect the statistical power for other analyses in the subsample.

First, concerning issue #1 of the simple longitudinal prediction of T2 and T3 reading, we find that left pOTS interruption also reliably predicts future reading in the subsample of initial minimal readers, even if the association is absent at age 5, confirming analyses in the main sample.

Concerning the issue #2 of the comparison with other benchmark predictors of literacy, our results in the entire sample show that despite a large amount of collinearity between our T1 predictors (documented, explained and quantified in lines 320-334, Figure 3B and Supplementary Figure 5D, including strong significant associations of left pOTS morphology with T1 reading, T1 letter knowledge and T1 performance IQ) and despite using the most stringent statistical approach to control for it (nested model comparison), left pOTS interruption still has a marginal contribution to predicting T3 reading on top of all other early literacy, pre-literacy, cognitive and familial predictors of reading (Figure 3C), in line

with its large separate contributions (Figure 3A). In the subsample of initial minimal readers, where there is only partial collinearity with performance IQ but not early literacy, we find the specific longitudinal contributions of left pOTS morphology to be strong and significant over and above all of the other T1 behavioral and familial predictors (see Figure 3F and corresponding text). In other words, the collinearity present in the full sample weakens statistical power when controlling for other collinear predictors, with only marginal unique longitudinal pOTS contributions in our stringent analyses, while removing this collinearity in the subsample reveals a large and highly significant unique contribution, in line with the large contributions of pOTS observed overall (Figure 3A, D, and F).We believe that detailing these analyses and finding coherent results on both samples, with potential differences in statistical strength explained by specific factors, overall brings reliability to our results and conclusions, rather than challenges them.

The reviewer refers also to one interesting incidental finding when examining the T1 cognitive associations of the pOTS interruption in the full sample, consisting in a significant association with the concept formation subtest, measuring a form of performance IQ (though a weaker association than with reading skills, see Figure 3B). When restricting to the subsample of initial minimal readers, this association was then indeed found marginal, though of a similar amplitude as in the full sample. We therefore speculated that the combination of a significant effect in the full sample, and of a marginal finding in the subsample, revealed a possibly true underlying effect, while likely affected by a decrease of statistical power associated with using only 56% of the sample. This effect and speculation are in no way fundamental to our study and claims. The new additional analysis revealing a significant association with performance IQ at T3 again in the full sample (Supplementary Figure 6) would however further support this idea.

> *Comment (7) Given that there are these potentially major differences in the pOTS, how are the results of structural connectivity mapped to a common template for Figure 4? There is likely an error in registration at the exact location of the observed differences.*

Individual structural connectivity results were mapped to an MNI template space using standard procedures (lines 919-922 of our Methods:):

“The obtained individual probabilistic fiber density maps were then transformed into template MNI space for statistical analysis using non-linear registration in ANTs, from individual diffusion space (individual FA map) to the HCP1065 MNI FA template, and smoothed with a 4-mm-FWHM Gaussian kernel.”

As the reviewer points out, group analyses performed in atlas space are inherently affected by variations in neuroanatomy across individuals and by the quality of registrations. Mapping individual connectivity results to template space was however the only approach that made sense to us to evaluate which connectivity benefits would be associated with the presence of such gaps, given the variability in the number and location of gyral gaps in children with interrupted sulci, and the need to find a common comparison ground with children who had a continuous sulcus. High neuroanatomical variability around the pOTS in particular likely plays against us in investigating such connectivity differences. In spite of these difficulties, we do find differences that seem to explain the reading advantage associated with the presence of gyral gaps.

> *Comment (8) While I appreciate the control analyses for the white matter connectivity, I think more robust analyses are more appropriate. Are there MD differences between white matter underneath sulci and gyri of neighboring or even distant regions? I'm not sure what the white matter analyses are really adding to the story. The white matter differences do not explain the differences in T3 reading ability.*

Our rationale for choosing and conducting the reported white matter analyses was the following: Having provided strong evidence that left pOTS sulcal interruptions facilitate reading acquisition, a naturally following question was that of its underlying mechanisms. It had been hypothesized - though never tested directly - that white-matter differences associated with the presence of gyral gaps may improve the connectivity around sulcal interruptions and constitute a potential mechanism linking interruptions to behavioral advantages (cf. lines 399-424). We therefore set out to investigate the underlying white matter, how it differed between children with and without interruptions, and its associations with reading skills.

Since long-distance projection connections tend to terminate more in gyri than sulci, one potential hypothesis was that gyral gaps in pOTS could be a hot-spot for such connections, facilitating the information flow between regions of the OTS and other functional hubs of the reading network. Because a higher density of such projection fiber endings is expected to result increased intra-cortical myelination, we test whether MD values in the gaps vs. sulcus are consistent with a higher myelination in the gaps than in sulcus, but we instead find the opposite pattern of lower myelination (cf. lines 435-457). This invalidated our first hypothesis.

In this context, it is thus unclear which questions would warrant additional analyses on MD values in other anatomical regions, or what these would bring in our specific search for mechanisms underlying the reading advantage of pOTS interruptions.

A second potential mechanism was that left pOTS gaps could be associated with increased local connectivity surrounding the left pOTS, through the packing of more U-shaped fibers traveling superficially, resulting somehow in a computational advantage for reading. This time, we found evidence in favor of this hypothesis: children with an interrupted left pOTS had more local white-matter connectivity around the pOTS travelling superficially under the left pOTS and it gaps (cf. lines 477-493). Moreover, when accounting in the same model for T3 reading on top of the presence of interruptions, this increase in local U-shape fibers was instead captured by T3 reading scores rather than the presence of a pOTS interruption. As such, these white matter differences did indeed explain differences in T3 reading, suggesting a candidate local-connectivity pathway underlying the reading advantage.

> *Comment (9) Were any of the analyses performed on the right pOTS? What are the differences in the laterality of gaps? Connectivity?*

The right pOTS had an interruption rate of 58.3%, which was similar to that of the left pOTS (56.5%, Fisher test: p = .75, cf. Supplementary Figure 1 and lines 134-139).

The right pOTS does not show associations with reading skills, which we report in Figure 2C and lines 184-186:

“[A]ssociations between sulcal interruption and reading skills were specific to the left pOTS within VTC. No significant associations were observed for anterior OTS (aOTS) or MFS in either hemispheres, or for the right pOTS (see Figure 2c, all p-values > 0.58, all $\Delta R^2 < 0.7\%$).”

Due to the absence of an association between the right pOTS and reading abilities, additional connectivity analyses in the right hemisphere did not seem warranted.

> *Comment (10) Finally, the authors' conclusion, and particularly the last section of their discussion on the neuronal recycling hypothesis, does not seem related to their data, analyses, or questions. The authors do not motivate the link between a hypothesis focused on changes in functional activation during skill acquisition, to their current work which only has structural and diffusion connectivity data. Either remove or further motivate this discussion.*

While there has been a heated debate on the specific nature and underlying mechanisms of "neuronal recycling", our intention was merely to frame our results within the general well-established and widely accepted premises that (i) cultural skills such as reading and symbolic number processing are too evolutionarily recent to rely on dedicated neural circuits and therefore must recruit pre-existing cortical territories, and (ii) reading acquisition is accompanied by the development of functional responses to written words in left posterior vOT, around the pOTS, free from dedicated evolutionary constraints.

Keeping in mind that the left pOTS and surrounding cortex see a major functional development of functional responses to words literacy acquisition is indeed fundamental in order to understand the relevance and implications of our structural results in this particular patch of cortex that hosts the VWFA, and their relative cognitive specificity to reading. Specifically, we claim that strong structure-function associations around the left pOTS for reading may mediate the behavioral benefits of left pOTS gyral gaps.

We therefore refer to these questions and hypotheses throughout our text. We first mention the existence and importance of structure-function relationships in the case of the VWFA early on in our introduction (lines 57-66). Throughout the paper, we state our hypotheses on how sulcal morphology and the associated structure of white-matter fiber tracts may result in functional and behavioral reading benefits, motivating our white-matter connectivity analyses. The discussion paragraph "Local connectivity differences at gyral gaps: a computational advantage?" (lines 629-648) then explicits in depth our final interpretation linking our sulcal marker and the computational efficiency of the functional reading network, as supported by an extensive literature.

To avoid any implication that our findings speak directly to the neuronal recycling framework, we have revised the relevant sentence (lines 651-659) and conclusion to focus solely on its broad, non-specific premises and have removed explicit mention of neuronal recycling.

"Interestingly, associations between sulcal interruptions and behavior have also been reported for symbolic, but not non-symbolic, number processing (Roell et al., 2021; Schwizer Ashkenazi et al., 2024), highlighting common processes by which cultural acquisitions, too recent on the scale of human evolution to rely on dedicated neural circuits (Dehaene & Cohen, 2007), may be boosted by the presence pre-existing gyral gaps. In the case of the left pOTS, various other aspects of early brain structure have been proposed to constrain the development of reading skills and of functional activations for reading in the VWFA and broader neural network involved in reading such as anatomical connectivity (Bouhali et al., 2014; Grotheer et al., 2021; Li et al., 2020; Moulton et al., 2019; Saygin et al., 2016) and cortical morphology. In terms of the latter, longitudinal studies following pre-reading children at familial risk of dyslexia showed for instance that the surface area of the left fusiform gyrus (Beelen et al., 2019) and patterns of grey-matter volume around the VWFA (Skeide et al., 2016) could be predictive of future dyslexia diagnosis."

We have also updated the subsection header to "Common cortical constraints on cultural learning" to reflect this more neutral framing, and adjusted the first discussion paragraph and the conclusion to avoid confusion with the neuronal recycling hypothesis. These changes clarify that our discussion situates the findings within a general conceptual background, rather than interpreting them as evidence for a specific theoretical model.

> *Comment (11) I think this paper could add interesting knowledge for the specific subfield of reading and the brain. However, the current state of the results, especially with the inclusion of so many trending results and the comparison of so many different processing pipelines and models, in addition to a conclusion that is not motivated by the work makes it difficult to appreciate the paper.*

We thank Reviewer 1 for their suggestions to improve the manuscript. We have revised the text to focus more clearly on the main longitudinal and anatomically specific findings, and have clarified the role of the initial minimal-reader subgroup, the significant longitudinal contribution of pOTS interruption beyond associations with emerging literacy at age 5, and its unique contribution relative to other established predictors of reading. We hope these changes make the central contributions of the study clearer.

Although the revision adds some content to an already detailed article, we believe that, taken together, these analyses provide a thorough and integrated account of the association between OTS morphology and reading. In particular, they help clarify how macro-anatomical differences in sulcal morphology may relate to cognition, how these associations unfold across development, and how a candidate white matter pathway may contribute to this relationship.

> *Minor points:*
>
> *Comment (12) Figure 1 needs more explanation for those who have not hand-labeled these regions. Can you clarify what differentiates the continuous left pOTS in the top row and then interrupted pOTS in the bottom row? The top row all has a gap between dark green and light green. Also, what does the color scale mean? Is it a posterior to anterior gradient?*

We thank Reviewer 1 for highlighting the need for clarification. Our legend now stipulates:

"Left hemispheric views from exemplar children are shown, either with a continuous left posterior OTS on the top row (i.e., no more than one connex OTS component in the space posterior to the white dotted line) or an interrupted left pOTS on the bottom row (two or more sulcal components present in posterior space, with a total Euclidean distance between components ≥1mm). Green and orange gradient colors are respectively used to highlight the different sulcal components of the OTS and MFS."

> *Comment (13) The discussion on sulcal interruption associations with symbolic but not non-symbolic number processing is interesting! Are these sulcal interruptions present in non-human primates?*

To our knowledge, pOTS interruptions directly analogous to the human left pOTS gyral gaps studied here have not been systematically characterized in non-human primates. Macaques do have an occipitotemporal sulcus, and comparative work shows that sulcal landmarks in macaque and human VTC can relate to functional organization. However, the human OTS is laterally displaced relative to the macaque OTS, and the specific left pOTS interruptions examined here are defined in relation to human VTC and literacy. We therefore view the comparative anatomy of these interruptions as an open question.

> *Comment (14) Your T1 voxel size (line 629 0.859 x 0.859 x 1.2mm) does not match the recommendation for the FreeSurfer pipeline you stated, being both anisotropic and having a dimension with > 1mm. "The method works well for voxel sizes 0.75 mm3. It should work with voxel between 1mm3 and 0.75mm3 ... We haven't tested the stream for inputs with anisotropic voxels." (https://surfer.nmr.mgh.harvard.edu/fswiki/SubmillimeterRecon). Further, you note in the methods section that small interruptions can be a result of bad segmentation (line 718/719). Given that segmentation is affected by voxel size, and potential partial volume effects. How do the voxel parameters impact the results and reliability of the analyses?*

Our T1 voxel volume was 0.885 mm^3, within the voxel volume range for which the FreeSurfer submillimeter stream is expected to work, although we acknowledge the caveat that the stream was not specifically validated for anisotropic voxels. Yet, Freesurfer pipelines

resample anisotropic voxels to isotropic ones by default, resulting here in a final voxel size of 0.859mm isotropic on which the submillimeter pipeline was run. To address possible segmentation concerns, we visually inspected all segmentations in bilateral ventral occipitotemporal cortex thoroughly, excluded scans with excessive motion, imposed a conservative> 1 mm criterion for classifying interruptions, and verified high T1–T3 reliability. Thus, while voxel anisotropy may affect very fine morphometric estimates, our macroanatomical feature of interest - the binary pOTS interruption classification - appears robust.

> *Comment (15) On line 417, you should refer to the panel of the figure (I think 4c bottom panel), like you did for many other results. It was helpful when you linked a particular result to its matching figure.*

We thank Reviewer 1 for pointing out this omission, which has now been corrected.

> *Comment (16) The percentage values on the bars should be larger for Figure 2, they are hard to read.*

We had chosen to represent the p-value significance on the bars, rather than the percentages themselves, which are indicated by the x-axis and reported numerically in the main text (except from the least relevant ones). We have now increased the font size of these significance stars. Thank you.

> ***Reviewer #2 (Recommendations for the authors):***
>
> *The authors may want to consider the points below.*
>
> *Comment (1) The relationship with Performance IQ is interesting. It does not predict much variance in reading scores at T3 but it is strongly predicted by the pOTS interruption at T3. There is quite a significant correlation among all the measures. The authors acknowledge the difficulty in assessing the unique contribution of each measure. I was unclear how they "over" controlled for pre-literacy skills etc. (line 296) and would like more elaboration on that in the main text.*
>
> *It would be interesting to know whether there were any cognitive measures where better performance was not predicted by the presence of the interruption? Were there any measures where worse performance was predicted? (i.e. a potential cost of this feature).*

A relative specificity of left pOTS interruption on reading:

We agree that the association between left pOTS interruption and performance IQ at T1 raised the question of the cognitive specificity of a reading advantage associated with left pOTS interruption, as opposed to a possibly more general cognitive advantage as suggested by Reviewer #2 in their public review above.

This question is of course complicated by the fact that reading skills are by nature highly correlated with general cognitive abilities, as they are also with math abilities (see e.g., the meta-analyses by Ünal et al., 2023) and language. General cognitive skills and reading are in fact so intertwined that they positively and symbiotically influence each other over development (Ferrer et al., 2007; Peng and Kievit, 2020). Yet, bearing these strong associations in mind, our results suggested a certain degree of cognitive specificity of left pOTS interruption benefits on reading. Indeed, while left pOTS interruption predicted about 14.3% of variance in T1 concept formation (measuring performance IQ), it predicts about 20.5% of T3 TOWRE reading scores. In addition, left pOTS interruption significantly and strongly predicts T3 reading scores, while T1 performance IQ did not (cf. Figure 3A, D and C, except in the subsample of initial minimal readers where pIQ has a significant contribution but about twice smaller than that of pOTS interruption - Figure 3F). Therefore, the prospective

association between left pOTS interruption and reading skills could not merely be driven by an association with general cognitive abilities at age 5. Finally, left pOTS interruption showed no other significant association with the pre-literacy skills considered (other than early literacy itself), or family history of dyslexia. This points towards a strong and relatively specific contribution of left pOTS on reading acquisition beyond its moderate association with T1 general cognitive skills.

We also conducted a new analysis to directly address this question, by systematically testing the associations between left pOTS interruption and all the cognitive and linguistic scores that were obtained at T3. Our new results also support the conclusion of a relatively specific cognitive advantage on reading, and are shown in Supplementary Figure 6 (with Supplementary Table 2 providing a description of the different subtests). The results are exposed in the main text as follows, in a new result subsection "Left pOTS interruption has a relatively specific cognitive impact on reading abilities" (lines 369-397):

"The incidental observation of a significant association between left pOTS interruption and performance IQ at T1, although smaller in magnitude and not sufficient to account for the association with literacy skills, raises the possibility that left pOTS interruption may confer a broader cognitive advantage rather than a specific benefit for reading acquisition. We therefore conducted a supplementary analysis to examine the relative domain-generality versus cognitive specificity of this association. This question is inherently difficult to address because reading and general cognitive skills are strongly correlated and mutually influence each other across development (Ferrer et al., 2007; Peng & Kievit, 2020; Ünal et al., 2023).

"We gathered all cognitive and linguistic scores obtained at T3, including measures of mathematics, general cognition, executive functioning, language, and reading (see Table S2 for a full list and description of the included subtests), and tested how each measure was associated with left pOTS interruption, residualized for demographic variables and total brain volume. Left pOTS interruption showed strong and significant associations with several measures of language, reading, mathematics, and general cognitive ability (see Supplementary Figure 6). As expected, we also observed strong and significant correlations among the behavioral measures themselves.

"However, the strongest correlations with left pOTS interruption residuals were observed for reading scores. Moreover, hierarchical clustering grouped left pOTS interruption with a subcluster that included most reading assessments - spanning phonological decoding, word reading, and sentence reading fluency -, as well as spelling and more surprisingly an expressive language index. This reading-related cluster was most closely related to another cluster comprising measures of written and oral language comprehension (i.e., receptive language), followed by a cluster of phonological awareness. In contrast, measures of mathematics, general cognitive ability, backward working memory and rapid naming formed a separate cluster. Thus, although left pOTS interruption is associated with several cognitive, linguistic, and academic measures, its pattern of associations is most closely aligned with children's reading outcomes, supporting a relatively specific cognitive advantage for reading."

We have therefore modified the Discussion paragraph that refers to the question of cognitive specificity vs. domain-generality as follows (lines 564-580):

"Interestingly, left pOTS interruption also showed associations with other cognitive skills: it was associated with performance IQ at age 5 (see Figure 3B and E), and with performance IQ, math and language skills more broadly at age 8 (Supplementary Figure 6). However, left pOTS interruption was most strongly associated and clustered with assessments of reading at age 8 (as well as spelling and expressive language). In addition, the prospective association between left pOTS interruption and reading skills was not merely driven by an association with general cognitive abilities at age 5 (Figure 3C and F) and no association with any other

pre-literacy skill was found (Figure 3B and E). Math, reading, language and cognitive skills are intrinsically highly intercorrelated, as we also observed at age 8, and are known to mutually and positively influence each other over time (Ünal et al., 2023; Peng & Kievit, 2020; Ferrer et al., 2007; Huettig & Hulstijn, 2025; Dehaene et al., 2015). The weaker pOTS cognitive associations may therefore result from the mere combination of a specific pOTS influence reading and of associations between reading and other cognitive domains, although more research is needed to better understand these associations and how they unfold during development. Together, these findings highlight a relative degree of specificity of left pOTS influences to reading abilities, in line with the welldocumented functional specificity of the left pOTS for orthographic processing and reading (e.g., Dehaene & Cohen, 2011; Gaillard et al., 2006; Pillet et al., 2024; Wilson et al., 2013; Woolnough & Tandon, 2024; Zemmoura et al., 2015; Zhan et al., 2023)."

No observed cognitive cost to left pOTS interruption:

These different analyses show that there is no negative cost to the presence of left pOTS interruption, or at least not across the wide variety of cognitive skills we considered. While left pOTS interruption was significantly and positively associated with many assessments of general cognition, math and language (perhaps due to inherent associations and interactions between these domains), several domains were not significantly associated with the morphology of the left pOTS: rapid naming, backward working memory, as well as receptive vocabulary and phonological awareness (at least at T1). Except for a nonsignificant correlation of -0.01 with CTOPP Blending Words at T3, no subtest showed negative associations with left pOTS interruption.

"Over-controlling" and assessing unique contributions:

We had used the term "over-controlling" to refer to the fact that testing the longitudinal effects of left pOTS interruption on T3 reading over and above a number of strongly collinear covariates (such as letter knowledge, T1 reading scores and performance IQ) by nested model comparison necessarily decreased the estimation of the true effect of left pOTS interruption on reading. Indeed, the baseline model with collinear variables then already accounts for a part of the variability in pOTS morphology and of its effect reading, through its joint effects with other co-variates, even before the sulcal variable is explicitly added in the model. The statistical power to test the true sulcal longitudinal effect by nested model comparison over such a strong baseline is therefore decreased, and the effect is underestimated.

Since this non-canonical term and our previous explanations appeared confusing, we clarified the corresponding paragraph as follows (lines 336-361):

"These significant associations between left pOTS interruption, performance IQ and pre-literacy skills (Figure 3B and D), as well as potential associations between pre-literacy skills themselves (cf. correlation matrices in Supplementary Figure 5D and H), make it challenging to fairly assess the separate unique contributions of each of these variables to the longitudinal prediction of reading skills. For instance, because left pOTS interruption explains a significant amount of variance at T1 in letter knowledge, early TOWRE scores and performance IQ, some of its true direct effect on future T3 reading skills is already captured in a baseline model including these cognitive predictors only, even before left pOTS interruption is explicitly entered as a predictor. Yet, even when controlling in a such a stringent way for all pre-literacy skills, early reading abilities, family history of dyslexia and their shared variance with left pOTS interruption in such nested model comparison, left pOTS interruption still had a marginal additional contribution to predicting T3 TOWRE reading scores (ΔR^2 = 5.0%, F(1, 31) = 3.74, p = .062) in spite of this important collinearity. Rapid naming and family history were the only two other predictors that outperformed left pOTS interruption in terms of unique contributions to T3 reading (ΔR^2 = 7.7% and 6.6% for RAN and mother ARHQ respectively, see Figure 3C), possibly because these predictors were among

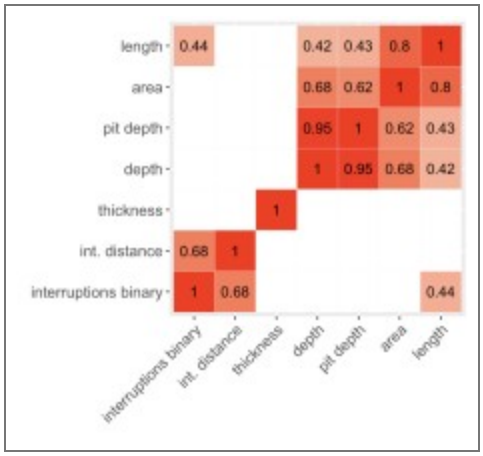
the least confounded with others (see Figure S5D for correlations among predictors) and thus among the least affected by collinearity. Letter knowledge, in contrast, had the smallest, non-significant unique contributions to T3 reading (ΔR^2 = 0.61%, p = .50). In order to reduce the collinearity confounds resulting from the associations between pOTS interruption and T1 early reading skills, we also reran the same analysis in the subsample of initial minimal readers. Left pOTS interruption in this subsample significantly contributed to predicting future T3 reading over and above all pre-literacy skills and family history (ΔR^2 = 27.8%, F(1, 12) = 13.80, p = .003), making it the first independent longitudinal contributor of reading (see Figure 3F). The second largest longitudinal predictor in terms of unique contributions was performance IQ (ΔR^2 = 12.8%, F(1, 12) = 6.34, p = .027). No other predictor had a significant unique contribution to the longitudinal prediction of T3 reading in the subset initial minimal readers, including rapid naming even though its separate contribution was larger than that of pOTS."

We also moved several figure panels from Supplementary Figure 5 to Figure 3 so that readers would have the necessary material in the main text to document this collinearity and see how it impacted the separate vs. unique contributions of each predictor to T3 reading.

> *Comment (2) It was surprising that no other anatomical feature of the pOTS (like sulcal depth) predicted reading skills. I would have expected these to be quite closely correlated with interruption distance that did predict reading. Do the authors have any thoughts as to why this might be?*

We thank Reviewer 2 for this question. While it seems intuitive that gyral gaps would affect sulcal depth, this mostly depends on whether we consider these gaps to be part of the sulcus - which here we do not- and on how the depth measure is actually calculated. With a definition of the sulcus as exclusive of gyral gaps, sulcal interruptions affect sulcal depth profiles only locally around the interruptions, where the fundus has to emerge to the convex hull surface. Because we measure sulcal depth based on the depth of the 8% deepest vertices (using an approach adjusted from that of Madan, 2019, see Methods lines 845-852), our measure of sulcal depth is essentially unaffected by interruptions.

As a confirmation and for more details, we provide Author response image 1 for the correlation matrix between the various morphological measures under consideration. (Using the same conventions as other correlation matrices displayed in the article and supplements: only significant correlations are displayed, according to FDR-corrected p-values above the diagonal.)



Author response image 1.

> *Comment (3) I hadn't quite appreciated on first reading the quite large variability in the location of the interruption (see Figure 4a, where three clusters seem to emerge). Nor did I appreciate that more than one interruption was often present (perhaps Figure 1 could be updated to show examples with more than one interruption). Some measures of distance along the OTS could be helpful and a possible comparison of the "types" based on location. I would have thought that proximity to the location of the VWFA might have meant a stronger effect. I presume this could be tested or discussed.*

We entirely agree with Reviewer 2 that these are important questions. Unfortunately, it also appears that these cannot be reliably tested in our sample.

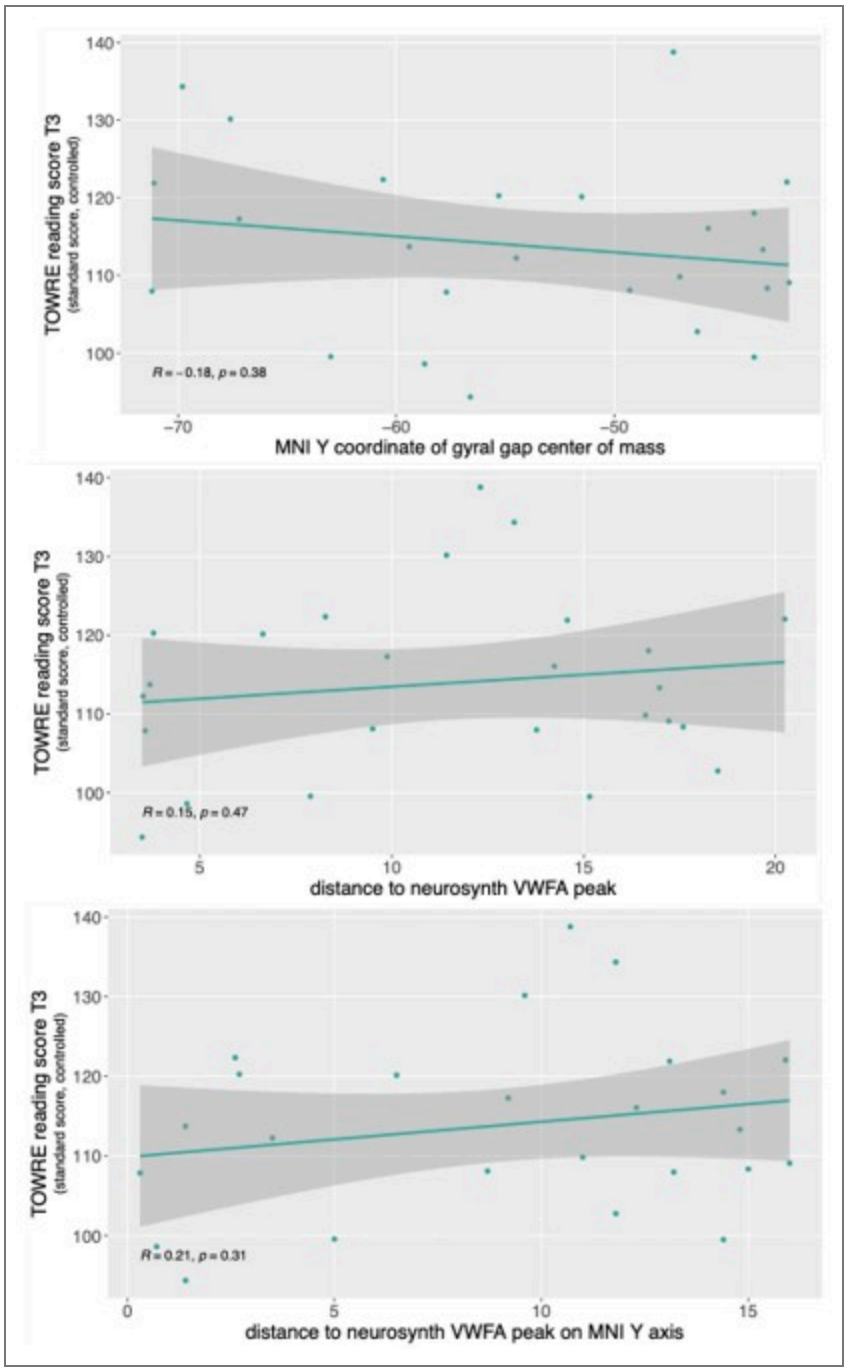
Specifically, while 25 children out of 43 had at least one interruption in the left pOTS, only 3 of those had two or more interruptions larger than 1mm each (or 6 children with two interruptions of any size). These incidence rates, combined with the fact that the locations of pOTS gaps do not distribute uniformly along the postero-anterior axis (Figure 4A), do not allow us to reliably assess in this sample either whether having more interruptions may confer an additional reading behavioral advantage, or whether interruptions in particular locations of the pOTS could be more beneficial than others.

We explored - for speculative purposes - whether a relationship may exist between reading scores and the location of left pOTS interruptions, taking the center of mass of the cortical gyral gap volumes as a coarse reference. We found no evidence for any relationship between reading and the locations of the left pOTS gap along the antero-posterior axis, nor between reading and the distance of the gap centroids to the VWFA (taking the Neurosynth reading meta-analysis peak coordinates, MNI -44 -58 -12), or between reading and the distance to the VWFA standard coordinates along the Y axis only (please see Author response image 2).

We choose not to report these analyses in the main text as they seem largely underpowered.

However, we added an example participant with two left pOTS interruptions in Figure 1 as suggested, and now added reference to these important questions in the Discussion, for future studies to address (lines 596-600):

“Such large-N approaches are needed to answer a number of outstanding questions, such as, in our case, whether there is an additional behavioral reading benefit to having two or more left pOTS interruptions - since these were too infrequent in our sample -, or whether gyral gaps placed in specific locations along the left pOTS may be optimal to for facilitating reading acquisition.”



Author response image 2.

> *Comment (4) It wasn't clear to me how the analysis of the diffusion data proceeded in the case of more than one interruption - it appears that each one was measured but this could be clarified. Is the location of relevance to the findings?*

Regarding diffusion analyses, the diffusion tensor analysis averaged values across all pOTS interruptions together in the few children who had several. The tractography analysis used the full pOTS and all its interruptions, if present, together as a seed volume.

The precise location of the gyral gaps would indeed be very relevant to the tractography findings as one expects the gyral gaps to relate in a local fashion to the neuroanatomy of the underlying white matter. However, using the whole pOTS with its gaps as a seed seemed to us the only reasonable way to investigate the white-matter correlates of the effect of gyral gaps

on reading at a group level, given (i) the large variability in gap location among the 15 of children with interruptions and diffusion data, (ii) the fact some children had multiple gaps, and most crucially (iii) the need to find a common ground for comparison with the group with no left pOTS interruption (continuous sulcus). The fact that our analyses average along the whole pOTS and across children with different pOTS gap locations likely washes away some of this neuroanatomical specificity, thus reducing the strength of any expected effect.

> *Comment (5) The authors could include a discussion of the sample's characteristics. They are not necessarily representative given the high SES and the average PIQ being more than 1SD above the mean.*

We thank Reviewer 2 for pointing out this limitation we had overseen.

High-SES and high-functioning population samples indeed constitute a common bias in pediatric research samples, since study participation requires parents to have the time, interest and access to necessary information. This bias was likely increased here by the fact the data was acquired at Stanford University, in a very privileged area (Palo Alto, CA). The median performance IQ was indeed about one standard deviation above the population mean. Moreover, our sample is enriched for family history of dyslexia. We agree with the reviewer that this limits the generalizability of our results.

Yet, we replicate the association between left pOTS interruption and reading skills observed by Cachia and colleagues (2018) in a sample where half of the participants were from a rural and (very) low socioeconomic background (so that they would match a sample of illiterate adults), and with limited higher education. This suggests that, overall, such associations may generalize across extremely varied socioeconomic backgrounds and also possibly along the spectrum of general cognitive skills.

Since our reviewed preprint came out, two other studies further replicated associations between the left pOTS and reading skills (Cachia et al., 2025; Yao et al., 2026), in populations of different ages, countries and across the reading spectrum, including many children with dyslexia (Yao et al., 2026). While SES and cognitive skills are not systematically reported in those studies, one could assume that these studies and ours pooled together would be more representative.

Accordingly, we added this paragraph to the discussion (lines 548-562):

“Our study further extends previous results by Borst, Cachia and colleagues (2016; 2018) to a population of children with above-average cognitive and reading scores, from relatively high-SES families, and enriched for family history of dyslexia (cf. Table 1) - although no child met criteria for dyslexia by T3. While our sample is not representative of the general population, thereby limiting the direct generalizability of our results, half of the adult sample from Cachia et al. (2018) came from low socio-economic and rural backgrounds (matched to a sample of illiterate adults), suggesting that the association between left pOTS interruption and reading skills is robust across socio-economic contexts. Since our reviewed preprint came out, the association between left pOTS interruption and reading was further replicated (Cachia et al., 2025; Yao et al., 2026), notably in a larger child sample along the full reading continuum and oversampling children with dyslexia (Yao et al., 2026). (Of note, the finding that pOTS morphology correlated in adults with reading fluency rather than accuracy (Cachia et al., 2025) may also explain one of two prior negative findings concerning the association between pOTS and reading (Del Mauro et al., 2022)). Thus, our results together with the extant literature overall underline a common effect of left pOTS interruption observed across the continuum of reading abilities, and possibly of cognitive abilities.”

> *Minor*

> *Comment (6) Line 342-44: A visualisation of what is meant by superficial vs deep plis de passage would help.*

While an illustration of deep vs. superficial *plis de* passage would greatly help in clarifying this distinction for the readers, we feel that providing such a visualization in our figures would be equally confusing, if not more, as it may prime readers to wrongfully assume that we differentiate between superficial and deep plis de passage in our own analyses. Instead, we have tried to further clarify the distinction between deep and superficial plis de passage in the text as follows (lines 418-424):

“Indeed, Bodin and colleagues (2021) showed that, in the superior temporal sulcus (STS), the presence of plis de passage was associated with a higher density of U-shaped fibers connecting adjacent gyri. Superficial plis de passage, i.e., complete sulcal interruptions with a gyral gap visible on the cortical surface, showed an even higher local connectivity advantage than deep plis de passage, the complete gyral gaps or partial interruptions with sulcal pinching which are buried within the depth of the sulcus. (Please note that in our study, we focus on superficial gyral gaps visible on the cortical surface only, and do not classify deep plis de passage.)”

> *Comment (7) Line 96 - replace "blindly from" with "without awareness of" or "while .... were concealed".*

The text was edited accordingly (lines 102-103 and 781) - thank you.

“[...] the OTS and MFS were delineated manually in each hemisphere, without awareness of the reading scores [...] Sulcal labeling was conducted without awareness of the children’s reading scores [...]”

> *(8) Line 547 - it is my impression that the findings of morphological differences in dyslexia are quite mixed. This gives the impression that the studies reported are representative of quite a vast number of studies and I don't think that is the case. Perhaps a sentence here with some more scepticism or alluding to the diverse findings in dyslexia?*

It was indeed not our intention here to simplify or summarize a broad literature on morphological differences in dyslexia, but rather to cite a few example results in support of the preceding assertion “various other aspects of early brain structure have been proposed to constrain the development of reading skills”.

We clarified this sentence as follows (lines 659-662):

“[L]ongitudinal studies following pre-reading children at familial risk of dyslexia showed for instance that the surface area of the left fusiform gyrus (Beelen et al., 2019) and patterns of grey matter volume around the VWFA (Skeide et al., 2016) could be predictive of future dyslexia diagnosis.”

> *Comment (9) Finally, there is quite a lot of supplemental information, some of which is very interesting. I would recommend that the authors check what could be usefully transferred to the main body of the article. eLife uses figures and figure supplements followed by Appendices and the authors might want to check this before deciding on the final version of record.*

We thank Reviewer 2 again for this suggestion to improve our manuscript. We have chosen to move several figure panels for Supplementary Figure 5 to Figure 3, following this suggestion, and given that our previous explanations in the paragraph concerning reading predictors appeared confusing to both reviewers (cf. see Rev 2 recommendation comment (1) and rev 1 recommendation comment (4)).

While figure supplements seem handy for reading the article directly on the eLife website, we chose to stick with the more standard practice of one appendix or supplementary file gathering all the additional information, which seems more accommodating of different reading practices overall.

References for Response to Reviews:

Borst G, Cachia A, Tissier C, Ahr E, Simon G, Houdé O (2016) Early Cerebral Constraints on Reading Skills in School-Age Children: An MRI Study. Mind, Brain, and Education 10:47–54.

Cachia A, Dupont C, Palmis S, Menu I, Auzias G, Sein J, Coulon O, Longcamp M (2025) The sulcal patterns of the occipito-temporal and anterior cingulate cortices influence reading and writing in children and adults. Cereb Cortex 35:bhaf145.

Cachia A, Roell M, Mangin J-F, Sun ZY, Jobert A, Braga L, Houde O, Dehaene S, Borst G (2018) How interindividual differences in brain anatomy shape reading accuracy. Brain Structure and Function 223:701–712.

Del Mauro G, Del Maschio N, Abutalebi J (2022) The relationship between reading abilities and the left occipitotemporal sulcus: A dual perspective study. Brain and Language 235:105189.

Ferrer E, McArdle JJ, Shaywitz BA, Holahan JM, Marchione K, Shaywitz SE (2007) Longitudinal Models of Developmental Dynamics Between Reading and Cognition From Childhood to Adolescence. Developmental Psychology 43:1460–1473.

Peng P, Kievit RA (2020) The Development of Academic Achievement and Cognitive Abilities: A Bidirectional Perspective. Child Development Perspectives 14:15–20.

Reuter M, Schmansky NJ, Rosas HD, Fischl B (2012) Within-subject template estimation for unbiased longitudinal image analysis. NeuroImage 61:1402–1418.

Roell M, Bellon E, Polspoel B, Declercq M, De Smedt B (2022) Investigating the association between variability in sulcal pattern and academic achievement. Sci Rep 12:12323.

Ünal ZE, Greene NR, Lin X, Geary DC (2023) What Is the Source of the Correlation Between Reading and Mathematics Achievement? Two Meta-analytic Studies. Educ Psychol Rev 35:4.

Yao JK, Mitchell JL, Davison A, Yeatman JD (2026) Sulcal anatomy of ventral temporal cortex and reading development. :2026.04.06.716640 Available at: *https://www.biorxiv.org/content/10.64898/2026.04.06.716640v1*.

https://doi.org/10.7554/eLife.103007.2.sa0