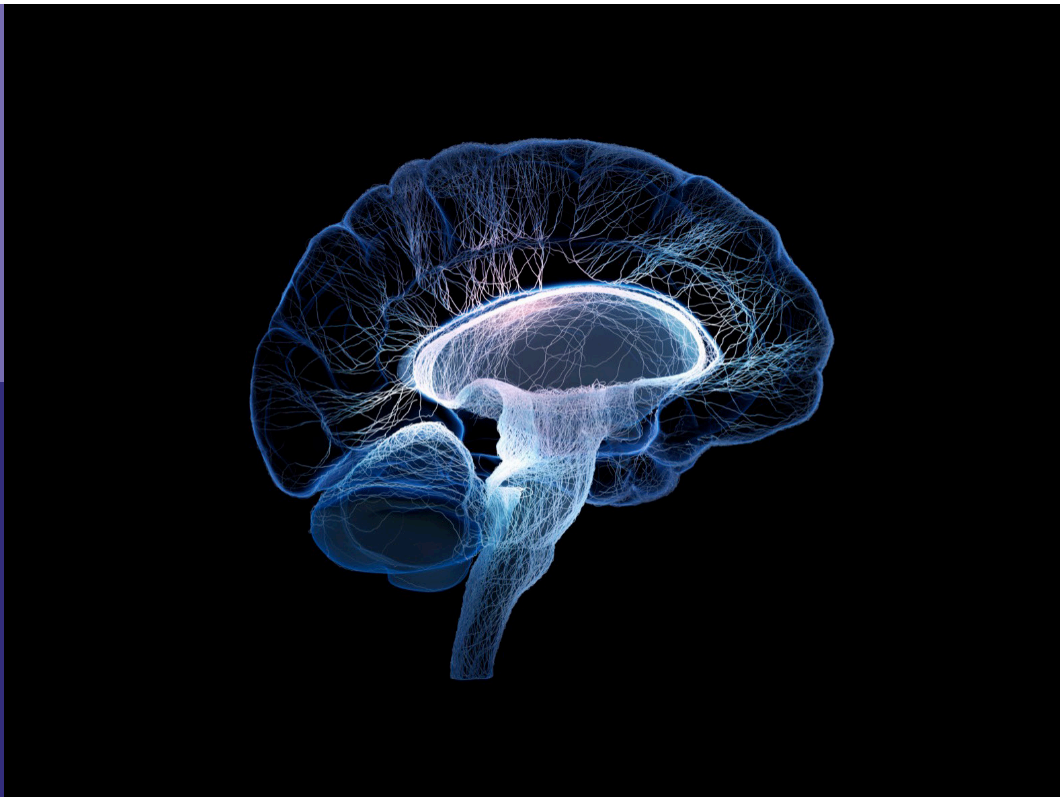
Research Topics

Neural basis of reading acquisition and reading disability

Edited by
Li-Hai Tan, Charles A. Perfetti, Johannes Ziegler and
Bruce D. McCandliss

Published in
Frontiers in Neuroscience
Frontiers in Human Neuroscience



ISSN 1664-8714
ISBN 978-2-8325-2880-8
DOI 10.3389/978-2-8325-2880-8

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Neural basis of reading acquisition and reading disability

Topic editors
Li-Hai Tan — Shenzhen Institute of Neuroscience, China
Charles A. Perfetti — University of Pittsburgh, United States
Johannes Ziegler — Aix-Marseille Université, France
Bruce D. McCandliss — Stanford University, United States

Citation
Tan, L.-H., Perfetti, C. A., Ziegler, J., McCandliss, B. D., eds. (2023). *Neural basis of reading acquisition and reading disability*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-2880-8

Table of contents

frontiers | Frontiers in Neuroscience

TYPE Editorial
PUBLISHED 16 June 2023
DOI 10.3389/fnins.2023.1147156

OPEN ACCESS

EDITED BY
Aaron J. Newman,
Dalhousie University, Canada

REVIEWED BY
Jeremy Purcell,
University of Maryland, United States

*CORRESPONDENCE
Li Hai Tan
✉ tanlh@sions.cn

RECEIVED 18 January 2023
ACCEPTED 01 June 2023
PUBLISHED 16 June 2023

CITATION
Tan LH, Perfetti CA, Ziegler JC and McCandliss B (2023) Editorial: Neural bases of reading acquisition and reading disability. *Front. Neurosci.* 17:1147156.
doi: 10.3389/fnins.2023.1147156

Editorial: Neural bases of reading acquisition and reading disability

Li Hai Tan[1,2]*, Charles A. Perfetti[3], Johannes C. Ziegler[4] and Bruce McCandliss[5]

[1]Center for Language and Brain, Shenzhen Institute of Neuroscience, Shenzhen, China, [2]Guangdong-Hongkong-Macau Institute of Central Nerve System (CNS) Regeneration, Jinan University, Shenzhen, China, [3]Learning Research and Development Center, University of Pittsburgh, Pittsburgh, PA, United States, [4]Laboratoire de Psychologie Cognitive, Aix-Marseille Université and Centre National de la Recherche Scientifique, Marseille, France, [5]Graduate School of Education, Stanford University, Stanford, CA, United States

KEYWORDS

brain, reading development, reading disability, dyslexia, cross-language

Editorial on the Research Topic
Neural bases of reading acquisition and reading disability

Introduction

Reading is an essential skill, necessary not only for success in school, but for maintaining a high quality of life in increasingly literate societies. Changes in technology have altered reading formats and increased the range and complexity of literacy contexts, thus placing even more pressure on foundational reading skills. The development of these skills occurs with corresponding neural development associated with learning the forms and functions of written language and their relation to spoken language. Difficulties in learning to read are associated with neural patterns that differ from those of successful learning. Thus, studies of the neural bases of reading inform the development of literacy and reading disability.

Recent progress builds on the remarkable foundation provided by more than 40 years of research from both behavioral and brain studies (Perfetti and Helder, 2022 is a review of some of this research). This foundation established basic facts about the cognitive processes of reading and learning to read, including the acquisition of orthographic, phonological, and semantic information that comprises the identity of printed words, supported by spoken language and conceptual knowledge.

One robust finding is that phonological knowledge, e.g., awareness of meaningless speech segments affects and predicts reading acquisition and developmental dyslexia. Awareness of the fine-grain level of phonology, the phoneme, is especially important for reading alphabetic writing systems and failures to achieve this awareness may be an indicator of risk for dyslexia.

World-wide, most children learn to read a non-alphabetic language. Thus, for understanding both the universality of reading development and its variation with specific languages and writing systems, cross-language research is important. Chinese reading has received the most research attention and can serve as a comparison with alphabetic reading. For instance, phonological knowledge is associated with Chinese reading development as it is with alphabetic reading development. However, visual-orthographic knowledge, visuo-motor, morphological awareness, vocabulary size, working memory, and some other factors may be as important as phonological knowledge in Chinese children's literacy acquisition.

The visual complexity of the Chinese character, its monosyllabic mapping, and its (limited) ideographic characteristics, require language-specific features to be part of any universal theory of reading disability.

With this background, this Research Topic brings new research and new reviews of research on the neurobiology of reading ability and disability across languages and writing systems. The 14 articles include 4 reviews and 10 articles of original research across a variety of methods (behavioral, fMRI, ERPs, FRPs, and computational modeling) and languages (English, German, Dutch, French, Finnish, and Chinese). These papers contribute significantly to important issues in reading development and dyslexia, as we discuss below.

The development of reading skill

Developing print expertise

In learning to read, children acquire the graphic forms of their writing system, converting them to orthographic units that connect to spoken language. With learning, they acquire precise connections to familiar and whole word orthographic patterns, allowing word identification to shift from computation to memory-based retrieval—fluent word reading. Verhoeven et al. studied the development of this "print tuning" process in learning to read Dutch. They found that fluent word reading did not emerge until children attained a threshold level in decoding accuracy, demonstrating a transition to word reading based on whole word orthographic input.

Complementing this behavioral indicator of print expertise is an ERP component (the N170) measured during word reading. A review by Amora et al. concludes that the distribution of the N170 across left and right hemispheres distinguishes better readers from less able readers and thus serves as an indicator of print expertise.

Developing print expertise involves the interconnections of visual and language brain systems (in occipitotemporal, temporoparietal, and inferior frontal cortex) that support orthographic, phonological, and semantic processing. Reading development must be accompanied by increased connectivity along pathways that implement these processes, e.g., the pathway between the dorsal inferior frontal gyrus and the posterior superior temporal gyrus. In an analysis of fMRI data of 91 native English-speaking children, Wagley and Booth found that, during phonological processing, connectivity along this pathway was related to word reading skill. This result highlights both the functionality of phonological processing along the dorsal pathway during word reading and its relation to actual word reading skill.

If brain organization changes with the development of reading, what happens when the brain has already been organized for reading one language and then acquires a second language? The paper by Cao et al. found that while brain areas active during reading are similar for bilingual readers (L1 Chinese and L2 English), the similarity is greater for adults than children. This is consistent with the convergence hypothesis that increasing proficiency in L2 leads to convergence with the L1 brain network.

Although research attention has focused on cortical areas, subcortical brain areas are also functional in reading through their pathways to cortical areas. In a resting imaging study of 334 Chinese children ages 6–12, Wang et al. found evidence for two different cortico-subcortical pathways, the thalamo-occipital and fronto-striatal circuits. The activation of the thalamus pathway was predicted by reading ability, especially among the younger children. The striatal pathway, perhaps reflecting attention and memory functions, showed a relation that increased with age and became reciprocal. For both younger and older children, reading ability predicted later striatum activation and this association was stronger for older children. For older children only, striatum activation also predicted later reading ability.

A fundamental aspect of typical reading is its dependence on spoken language. This is why phonological processes are so important in the development of reading skill and why spoken language components continue to be present as print expertise develops. In a study of Finnish 12–13-year-old children, Azaiez et al. used a mix of auditory ERP tasks and co-registered eye-tracking/ERPs to provide new insights into the speech-reading relationship. In particular, their study found a correlation between activation in the visual word form area and the superior temporal area (an auditory/speech area) during reading.

Predicting literacy prior to schooling

Predicting a child's reading success prior to schooling is valuable both theoretically and practically. Beyer et al. report a novel study using machine learning applied to behavioral and fMRI measures taken 2 years prior to reading instruction to predict reading outcomes during the 1st year of schooling. They found that future literacy could be predicted by gray matter volume in the left occipito-temporal cortex and local gyrification in the left insular, inferior frontal, and supramarginal gyri. Behaviorally, phonological awareness was also predictive. Thus, the status of the large-scale reading network at a preliterate age can predict how well children learn to read.

A key factor in reading development is a child's language experience. A young child's early language experience can be captured in their conversation with others, especially parents. A longitudinal study by Weiss et al. demonstrates a positive relationship between measures of parental language input during late infancy and the emergence of literacy skills at age 5. They further report that this relationship is probably mediated by the myelination of the left dorsal pathways of the left hemisphere's emerging language network.

Dyslexia: causes and interventions

Problems in phonological processing are a primary causal factor in dyslexia. Some proposals have attempted to explain the phonological deficit itself, and thus provide a deeper explanation of dyslexia. One is the neural noise hypothesis (Hancock et al., 2017). "Neural noise" captures the neural response variability that occurs across repetitions of a specific input to an individual: more variability, more noise. On this hypothesis, too much noise interferes with establishing stable representations of linguistic input. However, Tan et al. well-designed a study that found no evidence that dyslexic readers have noisier neural representations than typical readers.

Another proposal is to explain phonological deficits via visual system problems, although assigning a prominent role for visual system deficits has been controversial. Stein's review paper departs from viewing phonological deficits vs. visual deficits as an either-or dichotomy. It argues that a specific visual "transient" process involving magnocellular neurons in the visual system may be a cause of the phonological deficit, especially with the addition of deficits in processing transient auditory signals.

Attempts to explain phonological deficit by linking them to general neural mechanisms are likely to continue. For now, the state of the knowledge is that phonological processing difficulties are the primary causal factor for basic reading problems.

Research on associations of dyslexia with non-phonological abilities is important in gaining a fuller picture of dyslexia. For example, handwriting has been especially strongly associated with reading development in Chinese (Tan et al., 2005) and handwriting problems are often part of the dyslexia profile. Liu et al. add to this picture with their finding that Chinese children with dyslexia showed reduced connectivity between the sensory-motor network and the visual network during handwriting but not drawing.

Interventions

Well-targeted interventions improve the reading of children and adults with dyslexia. Such improvements are expected to produce neural changes that reflect some degree of "rewiring." However, two papers in the Research Topic conclude that finding specific brain changes, at least at the group level, is elusive. In their review paper, Braid and Richlan, while noting increased activity in RH homologs to the LH reading network, conclude that neuroplasticity effects do not emerge consistently following a successful intervention.

This conclusion is echoed in an original research paper by Krafnick et al. Their intervention raised reading scores, but without producing changes detectable in fMRI. Interestingly, however, reading gains were predicted by pre-intervention brain activity in bilateral supramarginal/angular gyri (and not predicted by pre-intervention behavioral assessments). Both papers emphasise the value of focusing on individual comparisons rather than group data in looking for brain changes.

Finally, although dyslexia is a human condition, Galaburda's review paper argues that animal models make a contribution to the study of its underlying mechanisms. Developmental cortical anomalies, cerebral asymmetries, functional lateralization, sound processing, and visual perception all can be modeled and genetic contributions to cell functions can be studied in animals.

Conclusion

We conclude with the suggestion that, by linking brain studies to behavioral indicators, the multi-method, multi-language approaches represented in these papers, add to knowledge about the development of typical reading and the characterizations of reading problems. They also pose new challenges and leave much to be informed by future research.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

Preparation of part of this Research Topic was supported by Shenzhen Institute of Neuroscience (the STI 2030-Major Projects 2021ZD0200500). JZ is the director of the Institute of Convergence ILCB, which is supported by grants from France 2030 (ANR-16-CONV-0002) and the Excellence Initiative of Aix-Marseille University (A*MIDEX).

Conflict of interest

Publisher's note

References

Hancock, R., Pugh, K. R., and Hoeft, F. (2017). Neural noise hypothesis of developmental dyslexia. *Trends Cogn. Sci.* 21, 434–448. doi: 10.1016/j.tics.2017.03.008

Perfetti, C. A., and Helder, A. (2022). "Progress in reading science: Word identification, comprehension, and universal perspectives," in *The Science of Reading: A Handbook*, 2nd Edn, eds M. Snowling, C. Hulme, and K. Nation (Cambridge: Cambridge University Press), 5–35.

Tan, L. H., Spinks, J. A., Eden, G. F., Perfetti, C. A., and Siok, W. T. (2005). Reading depends on writing, in Chinese. *Proc. Natl. Acad. Sci. U. S. A.* 102, 8781–8785. doi: 10.1073/pnas.0503523102

frontiers | Frontiers in Neuroscience

ORIGINAL RESEARCH
published: 25 April 2022
doi: 10.3389/fnins.2022.816729

Greater Similarity Between L1 and L2's Brain Network in Adults Than in Children

Fan Cao[1]*, Yuyu Fan[1], Xin Yan[2], Wuying Chen[3], Maddie Dodson-Garrett[2], Gregory J. Spray[2], Zhao Wang[4] and Yuan Deng[5]

[1] Department of Psychology, Sun Yat-sen University, Guangzhou, China, [2] Department of Communicative Sciences and Disorders, Michigan State University, East Lansing, MI, United States, [3] School of Education Science, Jiangsu Normal University, Xuzhou, China, [4] School of Chinese Language and Literature, Shaanxi Normal University, Xi'an, China, [5] CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China

It has been documented that processing L2 and L1 engages a very similar brain network in bilingual adults. However, it is not known whether this similarity is evident in bilingual children as well or it develops with learning from children to adults. In the current study, we compared brain activation in Chinese-English bilingual children and adults during L1 and L2 processing. We found greater similarity between L1 and L2 in adults than in children, supporting the convergence hypothesis which argues that when the proficiency of L2 increases, the L2's brain network converges to the L1's brain network. We also found greater differences between adults and children in the brain for L2 processing than L1 processing, even though there were comparable increase in proficiency from children to adults in L1 and L2. It suggests an elongated developmental course for L2. This study provides important insights about developmental changes in the bilingual brain.

Keywords: bilingual, fMRI, development, rhyming, Chinese-English

Edited by:
Charles A. Perfetti,
University of Pittsburgh, United States

Reviewed by:
Lan Shuai,
Haskins Laboratories, United States
Jie Zhuang,
Shanghai University of Sport, China

****Correspondence:***
Fan Cao
caofan3@mail.sysu.edu.cn

Specialty section:
This article was submitted to Neurodevelopment, a section of the journal Frontiers in Neuroscience

Received: *17 November 2021*
Accepted: *09 March 2022*
Published: *25 April 2022*

Citation:
Cao F, Fan Y, Yan X, Chen W, Dodson-Garrett M, Spray GJ, Wang Z and Deng Y (2022) Greater Similarity Between L1 and L2's Brain Network in Adults Than in Children. Front. Neurosci. 16:816729. doi: 10.3389/fnins.2022.816729

INTRODUCTION

There has been a consensus among researchers that brain networks involved in L1 and L2 processing are mostly overlapping in bilingual adults with subtle accommodations to the special features of the language (Perani et al., 1998; Chee et al., 1999a; Klein et al., 1999; Chee et al., 2000; Tettamanti et al., 2002; Musso et al., 2003; Sakai et al., 2004; Perani and Abutalebi, 2005). This neuroimaging finding of great overlap between L1 and L2 is consistent with the repeated observation from behavioral studies that L1 influences the organization of L2's representations (Wu and Thierry, 2010; Costa et al., 2016), suggesting an extensive transfer from L1 to L2 (Koda, 1990; Kroll et al., 2012). The finding of similar brain activation in L1 and L2 also lines up with the unified model proposed based on computational modeling studies, suggesting that the same computational principles can be applied in L1 and L2 language acquisition (MacWhinney, 2012). A recent study further suggests that higher L2 proficiency is related to greater similarity between L1 and L2 in brain activation in Chinese-English late bilingual adults (Cao et al., 2013a). Similarly, another study found that higher proficiency is associated with greater overlap between L1 and L2 in Italian-English bilinguals (Perani et al., 1998). The proficiency effect found in the previous two studies is supportive of the convergence hypothesis proposed by Green (2003), which argues that differences between native speakers and L2 speakers disappear as L2 proficiency increases. This hypothesis has so far been supported by a number of previous studies with a manipulation of L2 proficiency (Chee et al., 1999b; De Bleser et al., 2003; Briellmann et al., 2004; Stein et al., 2009). Taken together, previous studies suggest an overlapped network in L1 and L2, with L2 proficiency being an important variable influencing the degree of similarity (Hernandez et al., 2007).

How L1 and L2 are processed similarly or differently in bilingual children has been studied much less than in bilingual adults. A handful of studies have been published until today, and they suggest an overlapped network between L1 and L2, such as in a semantic judgment task in late Chinese-English bilingual children (Xue et al., 2004), and in a word reading task in simultaneous Hindi-English bilingual children (Cherodath and Singh, 2015). These studies also suggest some activation differences that may be due to different proficiency level. For example, a functional near-infrared spectroscopy (fNIRS) study found greater activation for L1 than L2 in a word repetition task at the bilateral middle/superior temporal gyrus, angular gyrus, supramarginal gyrus, and inferior frontal gyrus (IFG) in 6–10 years old Japanese-English bilingual children who learn English as a foreign language (Sugiura et al., 2011), which may be driven by greater proficiency in L1 than L2. Another study found English-Chinese bilingual children showed similarities between L1 and L2 in an auditory morphological task while the left IFG was more activated for English than Chinese which may be due to the higher proficiency in English than Chinese as well (Ip et al., 2016). In summary, according to the current literature, bilingual children also show an overlap between L1 and L2 brain networks.

However, what is lacking in the literature is the understanding of the dynamic change between L1 and L2 over development. For example, we do not know whether the brain network for L1 and L2 becomes more convergent or more differentiated with learning and development. According to the convergence hypothesis by Green (2003), one would expect greater convergence between L1 and L2 in adults than in children because the differences between the first and second language disappear as L2 proficiency increases. However, according to the interactive specialization model (Johnson, 2011), a specific brain region becomes specialized to a certain type of stimuli or calculations with learning and development. One would expect greater differentiation between L1 and L2 in adults than in children due to increased specialization to each language. Until today, there has been only one study that directly examined the developmental changes of brain activations in both L1 and L2 in bilinguals (Hernandez et al., 2015).

In Hernandez et al.'s (2015) study, it was found that during a single word reading task, Spanish-English bilingual adults showed greater activation in the bilateral middle temporal gyrus in reading English than bilingual children, while only the right middle temporal gyrus was more activated in adults than children in reading Spanish, suggesting greater developmental changes for L2, English than for L1, Spanish. This was explained by the fact that adults had a higher English proficiency than children, whereas their proficiency was comparable to children in Spanish. In other words, this study captured the transition from Spanish dominance in children to English dominance in adults. Therefore, the greater developmental change for English than Spanish in the brain is simply driven by greater proficiency change in English than in Spanish in this study. An ideal situation to study whether L1 and L2 are more convergent or differentiated in the brain over development is to have a bilingual population who have similar amount of growth in L1 and L2 across development, so that the increased/decreased language difference in the brain cannot be driven by unequal proficiency change in the two languages over development.

In the current study, we directly compared brain activation of L1 and L2 in Chinese-English bilingual children and adults who showed comparable increase in the proficiency of Chinese and English over development. We expected to observe greater similarity between L1 and L2 in brain activation in adults than in children if there is increased language convergence as proficiency increases. We expected less similarity between L1 and L2 in adults than in children, if there is increased differentiation between L1 and L2 as proficiency increases.

MATERIALS AND METHODS

Participants

Two groups of native Chinese adults and two groups of native Chinese children were recruited in Beijing, China. Adult participants were undergraduate or graduate students at Beijing Normal University, who were randomly assigned to perform a Chinese word rhyming judgment task [i.e., the adult Chinese (AC) group] (N = 20, mean age = 21 years), or to perform an English pseudoword rhyming judgment task [i.e., the adult English (AE) group] (N = 15, mean age = 22.9 years). The Chinese children were fifth-graders recruited from eight public elementary schools in Beijing. One group of children performed the Chinese word rhyming judgment task [i.e., the child Chinese (CC) group] (N = 14, mean age = 11 years) and the other group performed the English pseudoword rhyming judgment task [i.e., the child English (CE) group] (N = 13, mean age = 11 years). CE and CC were matched on age [$t(25) = 0.528$, $p > 0.05$], performance on a Chinese character naming test [$t(25) = 0.018$, $p > 0.05$] and a Chinese reading fluency test [$t(25) = 0.225$, $p > 0.05$] (**Table 1**).

According to the definition of late bilinguals by previous studies (Wartenburger et al., 2003; Berken et al., 2015), which is people who started to learn a second language at age 5 or later, all participants in our study were late Chinese-English bilinguals. The English proficiency of AE and CE was assessed with a standardized test, Woodcock Johnson-III (Woodcock et al., 2001) that includes five subtests of Antonyms, Synonyms, Word Identification, Word Attack, and Reading Fluency. Both Antonym and Synonym tested participants' vocabulary; Word Identification tested word reading accuracy; Word Attack tested pseudoword reading accuracy; Reading Fluency tested reading speed and comprehension. AE was significantly higher than CE on all five English tests [$t(25) = 6.041$, $p < 0.001$ for antonym; $t(25) = 5.146$, $p < 0.001$ for synonym; $t(26) = 6.850$, $p < 0.001$ for Word ID; $t(26) = 3.936$, $p < 0.01$ for Word Attack; $t(25) = 5.766$, $p < 0.001$ for Reading Fluency] (**Table 1**). Participants in both AE and CE had not lived in an English

TABLE 1 | Demographic information and testing scores for the four groups of participants.

Mean (standard deviation)	AC	AE	CC	CE
N	20	15	14	13
Age	21.45 (2.24) years	22.86 (2.26) years	134.38 (6.05) months	138.71 (5.54) months
AOA (years)		12.25 (0.62)		7.50 (2.10)
Chinese character naming (raw score)	–	–	124.00 (7.49)/150	124.08 (14.20)/150
Chinese reading fluency (raw score)	–	–	54.79 (13.49)/100	56.54 (14.54)/100
English synonym (raw score)	–	11.00 (5.48)/29	–	2.33 (2.19)/29
English Antonym (raw score)	–	15.13 (5.04)/29	–	4.92 (3.32)/29
Word ID (raw score)	–	43.27 (8.56)/60	–	20.00 (9.42)/60
Word Attack (raw score)	–	21.13 (5.51)/31	–	11.38 (7.56)/31
English reading fluency (raw score)	–	46.60 (13.21)/98	–	21.54 (9.02)/98
Rhyming judgment accuracy	0.93 (0.05)	0.82 (0.06)	0.82 (0.13)	0.67 (0.07)
Rhyming judgment reaction time	1224 (328)	1292 (365)	1655 (293)	1372 (375)

AC, adults doing the Chinese task; AE, adults doing the English task; CC, children doing the Chinese task; CE, children doing the English task.

immersion environment, and they do not use English in everyday life outside the classroom. Therefore, the immersion level and usage of English is comparable in AE and CE.

All participants were right-handed, free of any neurological disease or psychiatric disorders, did not have attention deficit hyperactivity disorder, and did not have any learning disabilities. The Institutional Review Board at Beijing Normal University and Michigan State University approved the consent procedures of the proposed study.

Stimuli and Tasks

Participants performed a rhyming judgment task during magnetic resonance imaging (MRI) scanning. For the Chinese word rhyming judgment task, Chinese words consisted of two characters without homophones at the word level. There were 24 trials in each of 4 conditions, 2 rhyming and 2 non-rhyming. The 4 conditions were: (1) similar orthography and phonology (O+P+; e.g., 弥补/mi2bu3/, 纯朴/chun2pu3/), (2) similar orthography and different phonology (O+P-; e.g., 翻译/fan1yi4/, 选择/xuan3ze2/), (3) different orthography and similar phonology (O-P+; e.g., 环保/huan2bao3/, 大炮/da4pao4/), and (4) different orthography and phonology (O-P-; e.g., 损坏/sun3huai4/, 学科/xue2ke1/). Rhyming was defined as the same rime for the second character in each word pair. Orthographically similar words were defined as sharing the same phonetic radical for the second character of the word pair. Participants were told to ignore the tone when making the Chinese rhyming judgment. However, in order to avoid the influence of tonal information on the rhyming judgment, in half of the trials, the second character of the first and second word had the same tone (e.g., 弥补/mi2bu3/, 纯朴/chun2pu3/), and in the other half, they had different tones (e.g., 逮捕/dai4bu3/, 胸脯/xiong1pu2/). The word stimuli and the second character in each word were matched on adult written frequency (Beijing Language and Culture University, 1990), and number of strokes across conditions.

For the English pseudoword rhyming judgment task, all pseudoword stimuli were monosyllabic and there were 2 rhyming conditions: (1) similar orthographic and phonological endings (O+P+; e.g., *weat-yeat*), (2), different orthographic but similar phonological endings (O-P+; e.g., *nean-leen*) and equal number of non-rhyming trials with different orthographic and phonological endings (O-P-; e.g., *wub-hafe*). There were 24 trials in each of the 2 rhyming conditions and 48 trials in the non-rhyming condition. Only the rhyming trials were included in data analysis for both the Chinese word rhyming judgment task and the English pseudoword rhyming judgment task, because a non-rhyming judgment could be made based on the whole syllable without the need to segment the syllable to rime and onset and then compare the rime. Pseudowords were used in the English task, because we wanted to avoid the possibility of different levels of familiarity to English words in children and adults.

Procedures

A pair of stimuli (either English pseudowords for the AE and CE groups, or Chinese words for the AC and CC groups) were sequentially presented in the visual modality and participants were asked to respond whether the two stimuli rhymed or not as quickly and as accurately as possible, using their right index finger for "yes" and their right middle finger for "no." Each stimulus was presented for 800 ms, with a 200 ms interval between stimuli. A red fixation cross appeared on the screen immediately after the offset of the second word/pseudoword in the stimuli pair, indicating the need to respond. The response interval duration was variable (2200, 2600, or 2800 ms), such that each trial lasted for either 4000, 4400, or 4800 ms. Perceptual trials (24) were also included in which it required participants to determine whether two sequentially presented visual symbol patterns were matched or mismatched by pressing the "yes" or "no" buttons. There were also 48 null trials included as a baseline, in which it required participants to press the "yes" button when a fixation cross at the center of the screen turned from black to red. The timing for the perceptual and null trials was the same as the lexical trials.

Magnetic Resonance Imaging Data Acquisition

All images were acquired using a 3.0 Tesla Siemens scanner (Siemens Healthcare, Erlangen, Germany) at Beijing Normal

University where participants lay in the MRI scanner with their head position secured with foam padding. An optical response box was placed in each participant's right hand and a compression alarm ball in the left hand. The head coil was positioned over each participant's head in a way that they could effectively use a mirror to view the projection screen at the rear of the scanner. Gradient echo localizer images were acquired to determine the placement of the functional slices. For the functional images, a susceptibility weighted single-shot echo planar imaging (EPI) method with blood oxygenation level-dependency (BOLD) was used with the following scan parameters: time echo (TE) = 20ms, flip angle = 80°, matrix size = 128 × 128, field of view = 220 × 220 mm, slice thickness = 3 mm (0.48 gap), number of slices = 33, time repetition (TR) = 2,000 ms. These parameters resulted in a 1.7 × 1.7 × 3 mm voxel size. 145 whole-brain volumes were acquired each run with an interleaved bottom to top sequence, with one complete volume collected every 2 s. A high resolution, T1 weighted 3D image was also acquired with the following parameters: TR = 2300 ms, TE = 3.36 ms, flip angle = 9°, matrix size = 256 × 256, field of view = 256 mm, slice thickness = 1 mm, number of slices = 160, resulting voxel size = 1 × 1 × 1 mm. The acquisition of the anatomical scan took approximately 9 min.

Image Analysis

Data analysis was performed using Statistical Parameter Mapping (SPM12).[1] The following steps were used for data preprocessing. (1) Slice timing correction for interleaved acquisition using sinc interpolation. (2) 4th degree b-splice interpolation for realignment to the first volume. (3) Trilinear coregistration with the anatomical image. (4) Segmentation of the anatomical image. (5) Normalization of all functional brains to the Montreal Neurological Institute (MNI) transforms derived from the segmentation of the structural image. (6) 4 × 4 × 8 mm full width half maximum Gaussian kernel smoothing.

Statistical analyses at the first level were calculated using an event-related design with all lexical conditions, the perceptual control condition, and the null baseline condition. A high pass filter with a cutoff period of 128 s was applied. Trials were modeled using a canonical hemodynamic response function (HRF). Data from each subject were entered into a general linear model using an event-related analysis procedure. Group results were obtained using random-effects analysis by combining subject-specific summary statistics across the group as implemented in SPM12. The contrast of rhyming trials (including O+P+ and O-P+) versus the perceptual trials was entered into a flexible factorial design: an age (adults, children) by language (Chinese, English) ANCOVA with task accuracy as a covariate. Age effect, language effect and the interaction between them were calculated and reported. All reported results were at uncorrected $p < 0.001$ at the voxel level, and FDR corrected $p < 0.05$ at the cluster level.

In order to examine the similarity between brain activation of L1 and L2 within each age group, we calculated the similarity between CE and CC, and between AE and AC using the Dice Similarity Coefficient (DSC) formula (Dice, 1945; Craddock et al., 2012): $2|A \cap B|/|A| + |B|$, where $A \cap B$ is the number of voxels activated in the overlap between A and B, and $|A| + |B|$ is the sum of brain voxels that were activated in A and B. The DSC ranges between 0 (no similarity) to 1 (perfect similarity). Therefore, the DSC is expected to inform the degree of similarity between A and B.

In order to examine what drove the interaction of age and language, we ran separate contrasts to examine developmental changes within each language (i.e., English: adults > children, and children > adults; Chinese: adults > children, and children > adults). We also broke down the interaction by looking at language Differences within each age group (i.e., adults: Chinese > English, and English > Chinese; children: Chinese > English, and English > Chinese). These contrasts were calculated with accuracy on the task regressed out. Then by identifying the contrasts that show common effect as the interaction, we can explain what drove the interaction.

For regions that showed a significant interaction in the last step, we also examined brain-behavioral correlation to examine whether these regions are also associated with English proficiency separately in AE and CE. The averaged accuracy on Synonym, Antonym, Word Attack, Word ID, and Reading Fluency subtests represented English proficiency.

RESULTS

Behavioral Results

We conducted an age (children, adults) by language (Chinese, English) ANOVA on accuracy and reaction time of the rhyming judgment task. We found a significant main effect of age with adults being more accurate [$F(1,58) = 47.289$, $p < 0.001$] and faster [$F(1,58) = 8.512$, $p < 0.01$] than children. We found a significant main effect of language with Chinese being more accurate [$F(1,58) = 67.370$, $p < 0.001$] than English. The main effect of language was not significant for reaction time [$F(1,58) = 1.518$, $p > 0.05$]. The interaction was not significant for either accuracy [$F(1,58) = 3.253$, $p > 0.05$] or reaction time [$F(1,58) = 4.009$, $p > 0.05$] (**Table 1**).

Brain Activation Results

The Main Effect of Language

We found greater activation for Chinese than English at the left STG, bilateral lingual gyri and cuneus, bilateral putamen, and the left IFG. We found greater activation for English than Chinese in the right superior parietal lobule (SPL) (**Table 2** and **Figure 1**).

When we examined language differences separately for adults and children, we found greater activation in Chinese than in English in the bilateral cuneus and left STG in adults, and greater activation in bilateral lingual/MOG, bilateral STG/MTG and left caudate in Chinese than in English for children. Neither children nor adults showed greater activation in English than in Chinese at the current threshold (**Table 3** and **Figure 2**). However, when we lowered the threshold to $p < 0.005$ uncorrected, Cluster = FDR corrected $P < 0.05$, children showed greater activation in English than in Chinese in the right SPL. The DSC between Chinese and

---

[1] http://www.fil.ion.ucl.ac.uk/spm

TABLE 2 | Brain regions that showed a significant main effect of language, a significant main effect of age, and a significant interaction effect between language and age.

Anatomical Region	H	BA	Voxels	x	y	z	Z
Main effect of language: *Chinese* > *English*							
Superior temporal gyrus	L	22	126	−51	−40	8	5.93
Lingual gyrus, cuneus	L, R	17, 18, 19	1217	−6	−70	−4	5.77
Putamen	L		168	−18	11	−4	5.31
Putamen	R		117	9	8	−1	4.31
Inferior frontal gyrus	L	45	35	−48	29	8	3.99
Main effect of language: *English* > *Chinese*							
Superior parietal lobule	R	7	44	24	−67	44	4.26
Main effect of age: *Adults* > *children*							
–							
Main effect of age: *Children* > *Adults*							
Fusiform gyrus	R	37	106	30	−52	−13	4.64
Interaction effect							
Inferior frontal gyrus, Middle frontal gyrus,	R	9	31	54	17	38	4.62
Post-central gyrus	R	2	25	45	−31	47	4.42
Medial frontal gyrus	L/R		52	−12	38	−7	4.07

H, hemisphere; L, left; R, right; BA, Brodmann area; Voxels, the number of voxels in each cluster; x,y,z, coordinates in the MNI atlas; Z: z-value of the peak voxel in the contrast.

English is 0.26 for adults and 0 for children at the threshold of $p < 0.001$ uncorrected voxel level and 0 at the cluster level.

The Main Effect of Age

We found greater activation in children than in adults in the right fusiform gyrus (**Table 2** and **Figure 3**). We found that no regions showed greater activation in adults than in children.

When we examined age effect separately in Chinese and English, we found no age differences in Chinese. However, for English, we found greater activation in adults than in children in bilateral medial frontal gyri, bilateral STG, and bilateral posterior cingulate gyri. We found greater activation in children than in adults in the right post-central gyrus, right IFG/MFG and the right fusiform gyrus for English (**Table 4** and **Figure 4**).

TABLE 3 | Brain regions that showed significant differences between Chinese and English in either adults or children.

Anatomical region	H	BA	Voxels	x	y	z	Z
Adults*: Chinese* > *English*							
Cuneus	L/R		527	15	−91	8	5.02
Superior temporal gyrus	L	22	36	−51	−40	8	5.01
Adults*: English* > *Chinese*							
–							
Children*: Chinese* > *English*							
Superior temporal gyrus	R	22	38	39	−28	−7	4.94
Middle temporal gyrus	L	22	91	−51	−37	2	4.93
Lingual gyrus	L	18	338	−6	−67	−1	4.84
Caudate	L		126	−15	11	−1	4.77
Middle occipital gyrus	R	19	27	24	−91	2	3.90
Children*: English* > *Chinese*							
–							

Interaction Effect

At the whole brain level, we found three regions that showed a significant interaction effect between age and language (**Table 2** and **Figure 4**). They are the right IFG/MFG, right post-central gyrus, and bilateral medial frontal gyrus. Break-down of the interaction showed that the interaction at the right IFG/MFG and the right post-central gyrus was due to greater activation in children than in adults in English but not Chinese, while the interaction at the bilateral medial frontal gyrus was due to greater activation in adults than in children in English but not in Chinese (**Table 4** and **Figure 4**).

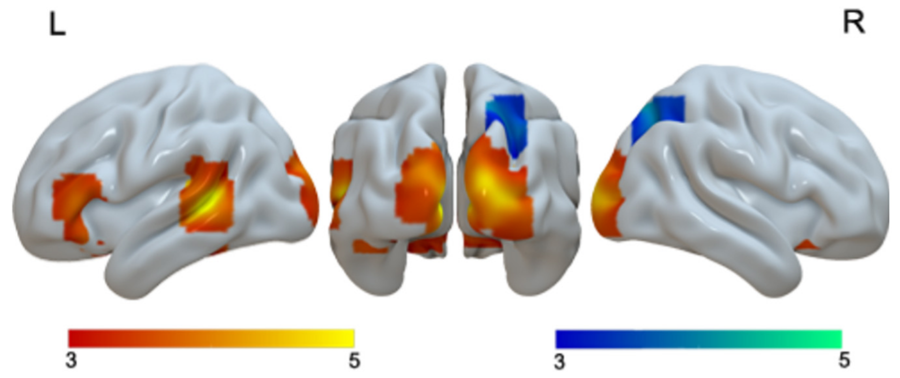


FIGURE 1 | Brain activation in the main effect of language. Brain regions that were more activated in Chinese than in English are in red; brain regions that were more activated in English than in Chinese are in blue.

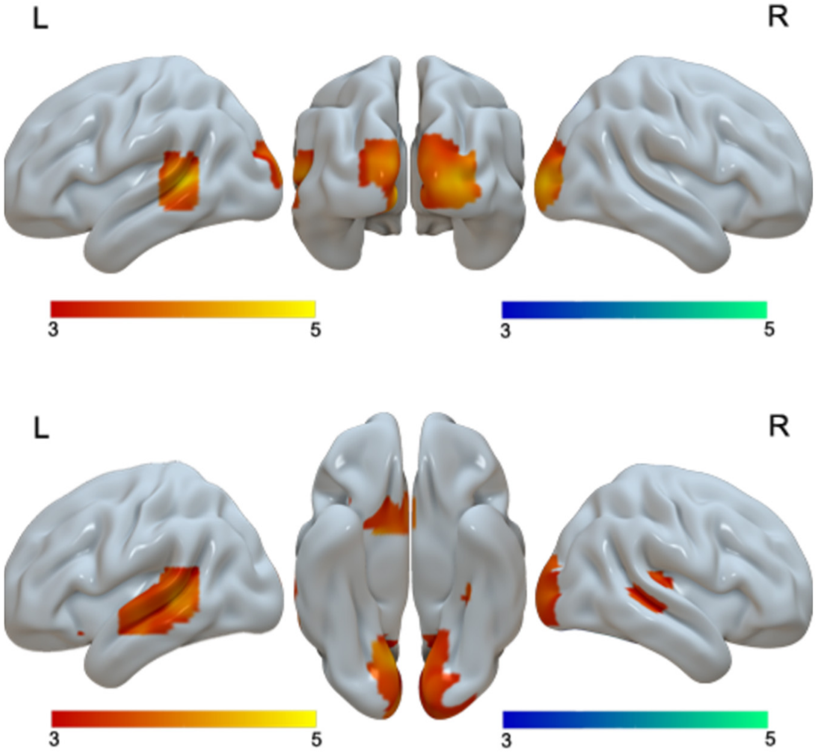


FIGURE 2 | First row, brain regions that showed greater activation in Chinese than in English for adults. Second row, brain regions that showed greater activation in Chinese than in English for children. No brain regions showed greater activation in English than in Chinese for either adults or children.

Brain-Behavioral Correlation

Since the interactions were due to greater developmental differences in English than in Chinese, we correlated brain activation at the three regions that showed a significant interaction effect with English proficiency separately for children and adults. For the CE group, we calculated correlation between English proficiency and brain activation at the ROI of the right inferior frontal gyrus, post-central gyrus and the bilateral medial frontal gyrus. We found a significant negative correlation at the medial frontal gyrus ($r = -0.635$, $p = 0.008$) (**Figure 5**). We also found a marginally significant positive correlation at the right inferior frontal gyrus ($r = 0.471$, $p = 0.066$). The correlations in the AE group were not significant ($r = -0.186$, $p = 0.543$ for the right inferior frontal gyrus; $r = 0.338$, $p = 0.259$ for the right post-central gyrus; $r = -0.419$, $p = 0.154$ for the bilateral medial frontal gyrus). A direct comparison between the correlation coefficients in AE and CE at the medial frontal gyrus revealed a non-significant difference ($z = -0.721$, $p = 0.47$). It was not significantly different between AE and CE either at the right inferior frontal gyrus ($z = 1.663$, $p = 0.09$).

DISCUSSION

In this study, we examined the developmental changes in brain regions involved in L1 and L2 processing by directly comparing bilingual children and bilingual adults. Behaviorally, we found no interaction between age and language, which simply suggests that the proficiency level increased similarly in L1 and L2 from children to adults. However, brain imaging data revealed greater developmental changes in English than in Chinese. In terms of comparison between L1 and L2 within each age group, we found greater similarity between L1 and L2 in adults than in children, which supports the convergence hypothesis. When L2 proficiency increases, the L2 network and L1 network becomes more convergent.

Language Differences in the Brain

For the main effect of language, we found greater activation in the left STG, IFG and bilateral lingual gyrus for Chinese than for English, and greater activation in the right superior parietal lobule for English than for Chinese. The differences at the left STG and IFG should be driven by different levels of semantic and phonological activation in Chinese and English. The stimuli for the English task were English pseudowords which elicit less semantic activation than Chinese words. The left STG is involved in phonological representation and language comprehension (Binder et al., 1997), while the left IFG is associated with phonological processing, such as phonological retrieval and phonological manipulation (Pugh et al., 1996; Fiez and Petersen, 1998). Greater activation in bilateral lingual gyri in Chinese than in English is due to more complex visuo-orthographic configuration in Chinese than in English, which is well documented in previous research (Bolger et al., 2005;

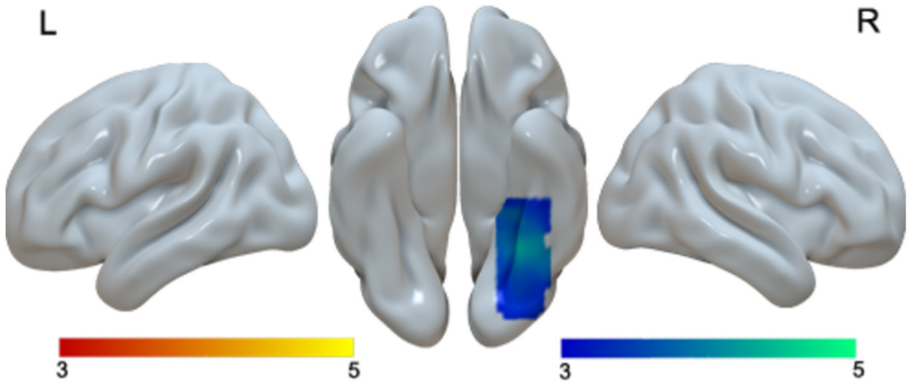


FIGURE 3 | Brain activation in the main effect of age. Brain regions that were more activated in children than in adults (blue). No brain regions were more activated in adults than in children.

Cao et al., 2010; Cao et al., 2013b). The greater activation in the right SPL in English than in Chinese might be due to compensation mechanisms of greater right hemispheric involvement in less proficient performers. The right SPL is associated with visuo-spatial processing as found in previous studies (Suchan et al., 2002). This region has been found to be more involved in Chinese than English (Cao et al., 2013b; Cao et al., 2015). Therefore, our finding suggests that these bilingual children and adults rely on visual spatial analysis important for their L1 to an even greater degree when they process L2, English, implicating a carry-over effect in L2 processing.

When we examined language difference in adults and children separately, we found greater language differences in bilingual children than bilingual adults. These additional language differences in children included greater activation in Chinese than in English at the right MTG and left caudate, as well as greater activation in English than in Chinese in the right SPL if we lowered the threshold. The DSC was 0.26 in adults and 0 in children. This is consistent with the convergence hypothesis that increased proficiency in L2 is associated with greater similarity to L1 brain activation (Perani et al., 1998; Golestani et al., 2006; Stein et al., 2009; Cao et al., 2013a). A previous study found that early bilinguals showed greater similarity between L1 and L2 at the left IFG than late bilinguals (Kim et al., 1997), suggesting an AOA effect. In our study, children had an earlier AOA than adults, however, we found greater similarity of brain activation in L1 and L2 in bilingual adults than bilingual children, suggesting a proficiency effect rather than an AOA effect. Our study adds to the literature that the overlap between L1 and L2 brain networks actually develops with learning and development. It provides important evidence from a developmental perspective for the convergence hypothesis, which argues that the difference between L1 and L2 in the brain disappears as L2 proficiency increases.

TABLE 4 | Brain regions that showed significant differences between adults and children in either English or Chinese.

Anatomical region	H	BA	Voxels	x	y	z	Z
English*: Adults* > *Children*							
Medial frontal gyrus	R/L	11	83	3	44	−13	4.38
Superior temporal gyrus	L	22	48	−57	−13	−1	4.32
Superior temporal gyrus	R	41,22	76	60	−25	5	4.12
Posterior cingulate gyrus	L/R	23	58	−6	−58	11	4.04
English*: Children* > *Adults*							
Post-central gyrus, Inferior parietal lobule	R	40, 2	201	42	−34	44	5.02
Fusiform gyrus	R	37	80	33	−49	−16	4.74
Inferior frontal gyrus, Pre-central gyrus	R	6, 45	51	48	8	11	4.24
Chinese: *Adults* > *children*							
–							
Chinese: *Children* > *Adults*							
–							

Greater Developmental Changes in L2 Than in L1

The main effect of age was primarily driven by age differences in English, since we found no age differences in Chinese. For English, children had greater activation than adults in the right post-central gyrus, right inferior frontal gyrus and right fusiform gyrus, while adults had greater activation than children in bilateral medial frontal gyrus, bilateral STG and bilateral posterior cingulate gyrus. It suggests that child L2 learners who had a lower proficiency level tend to involve more right hemisphere regions. The involvement of the right hemisphere is common in beginning readers or low proficiency readers, such as those with dyslexia (Waldie et al., 2013). The involvement of the right IFG/MFG in younger normal children declines with age during reading (Shaywitz et al., 2007), suggesting that its importance decreases as reading develops in the left hemisphere. It has also been suggested that the right temporo-parietal-frontal regions are especially important

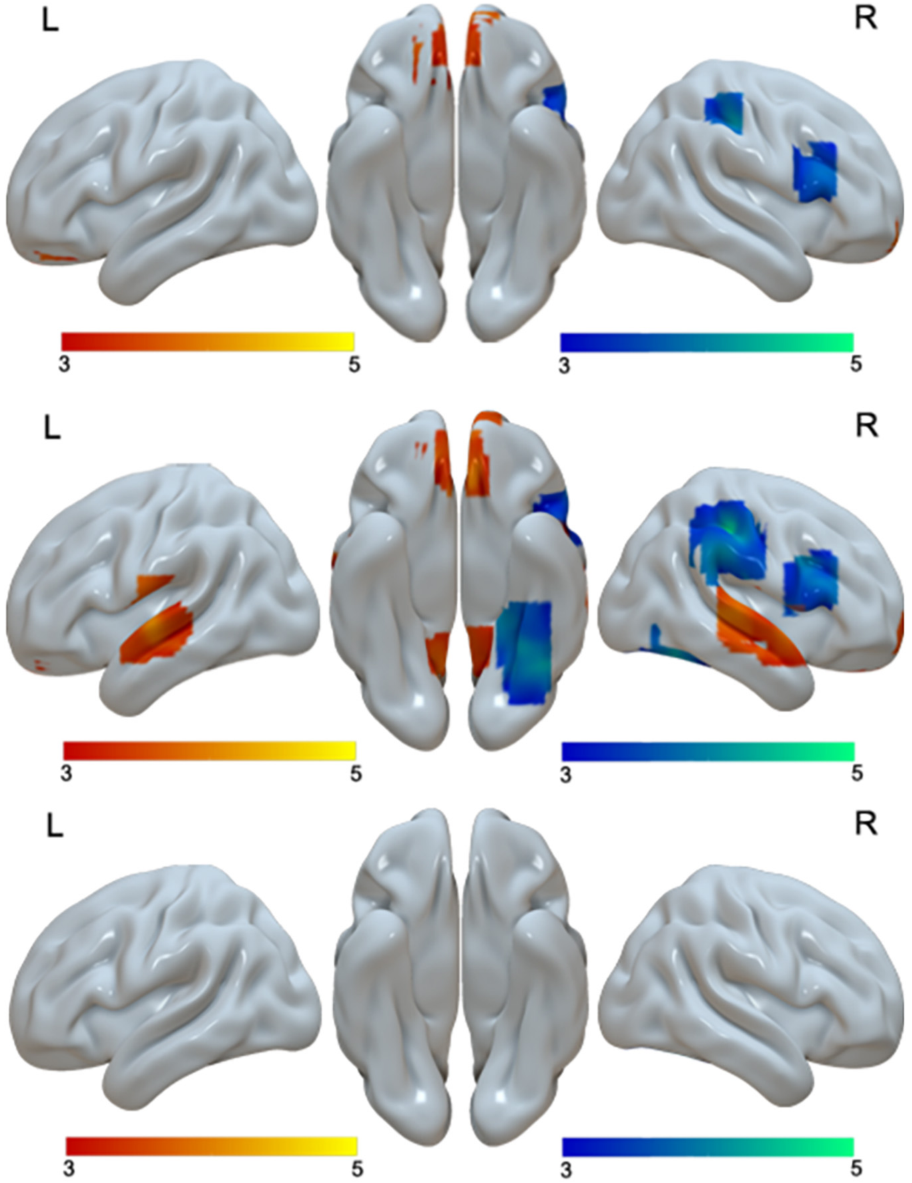


FIGURE 4 | Brain activation in the interaction of language by age (first row), including brain activation in the positive interaction (red), and brain regions in the negative interaction (blue). Brain activation in the comparison between adults and children in English (second row). Brain regions that were more activated in adults than in children in English are in red; brain regions that were more activated in children than in adults are in blue. No brain regions showed age differences in Chinese (third row). The interactions were driven by greater age differences in English than in Chinese.

for the early stage of language learning/acquisition in both L1 and L2 (Sugiura et al., 2011). Our finding is consistent with this pattern of shifting with learning. We found that the right IFG and post-central gyrus were more involved in bilingual children than bilingual adults for English, presumably due to compensation for deficient phonological processing in the left language network in children. Furthermore, at the right inferior frontal gyrus, there is a marginally significant positive correlation with English proficiency in children, suggesting that higher skilled children tend to use this right IFG compensation strategy to a greater degree. The right fusiform gyrus was also more involved in bilingual children than bilingual adults for English. This region is more involved in Chinese reading than in alphabetic reading (Bolger et al., 2005; Tan et al., 2005), presumably due to the more complex visuo-orthographic features of Chinese. It is also more involved in higher skilled Chinese readers than lower skilled readers (Cao et al., 2009; Cao et al., 2010), suggesting its essential role in Chinese reading. Therefore, greater activation in this region in children than in adults during English pseudoword processing suggests a greater reliance on the L1's visuo-orthographic strategy when processing unfamiliar L2's orthography.

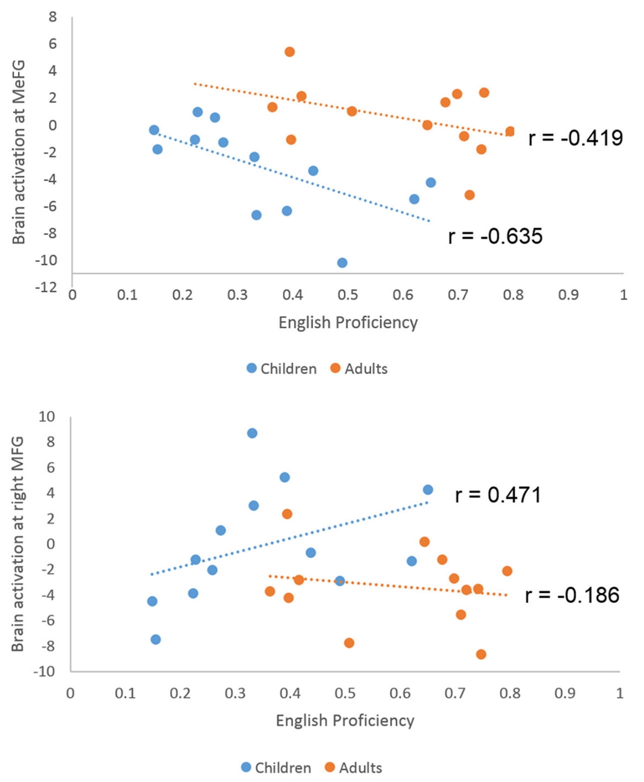


FIGURE 5 | Scatter plots for the correlation between English proficiency and brain activation level at the medial frontal gyrus and right middle frontal gyrus in children and adults. Significant negative correlation was found in children but not in adults at the medial frontal gyrus. The correlation difference between children and adults is not significant. A marginally significant positive correlation was found in the right middle frontal gyrus for children.

On the other hand, adults showed greater activation than children in bilateral STG, bilateral medial frontal gyrus and bilateral posterior cingulate gyri during the English task. Greater activation in bilateral STG implicates greater activation level of English phonology which should be driven by greater familiarity to the English phonology in adults than in children. Bilateral medial frontal gyri and bilateral posterior cingulate gyri are critical parts of the default mode network (Raichle et al., 2001), which are deactivated during active cognitive tasks. Greater activation in these regions in adults than in children might be due to reduced task difficulty for adults than for children. We found a negative correlation with English proficiency in children in the bilateral medial frontal gyrus, suggesting that higher skilled children tend to deactivate this region to a greater degree.

Limitations

One limitation of the current study is the cross-sectional between-subject design. It would be ideal to study developmental changes in the brain with language learning in a longitudinal design, however, due to obvious constraints, most studies conducted in the field of cognitive developmental neuroscience are cross-sectional. Even though it cannot provide strong causal inferences of brain development, the current study showed differences in the brain between adults and children during language processing, implicating age and proficiency influence in the brain. More importantly, the main finding of the study is the interaction between age and language; specifically, the age effect in L1 and L2 is different. Therefore, even if we employed a within-subject design, the key finding is still a between- subject comparison. Another limitation of the current study is the unmatched AOA in adults and children. If the AOA was matched, the proficiency effect would be even greater in the current study, because earlier AOA was associated with greater similarity between L1 and L2 (Kim et al., 1997), and adults had a later AOA but greater similarity than children in the current study. However, it is impossible to match AOA in Chinese bilingual children and adults due to national policy change. Future research should match AOA while examining age and proficiency effect.

CONCLUSION

We examined the developmental change of the dynamics between L1 and L2 in the brain in Chinese-English late bilingual children and adults. We found greater similarity between L1 and L2 in adults than in children, supporting the convergence hypothesis, which argues that the difference between L1 and L2 disappears as L2 proficiency increases. Moreover, we found greater differences between children and adults in the L2's brain network than in the L1's brain network. Our study, for the first time, demonstrated that there might be different brain mechanisms underlying L1 and L2 development.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors upon request, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Beijing Normal University and Michigan State University. Written informed assent and consent were obtained from participants and their parents, respectively.

AUTHOR CONTRIBUTIONS

FC: conceptualization, writing, and supervision. XY: data collection. YF, WC, MD-G, GS, and ZW: data analysis. YD: supervision of data collection. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the "National Social Science Fund of China (21BYY204)" awarded to FC.

REFERENCES

Beijing Language and Culture University (1990). *Frequency for Modern Chinese Words* (Unpublished corpus). Beijing: Beijing Language and Culture University.

Berken, J. A., Gracco, V. L., Chen, J. K., Watkins, K. E., Baum, S., Callahan, M., et al. (2015). Neural activation in speech production and reading aloud in native and non-native languages. *Neuroimage* 112, 208–217. doi: 10.1016/j.neuroimage.2015.03.016

Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M., and Prieto, T. (1997). Human brain language areas identified in functional magnetic resonance imaging. *J. Neurosci.* 17, 353–362. doi: 10.1523/jneurosci.17-01-00353.1997

Bolger, D. J., Perfetti, C. A., and Schneider, W. (2005). Cross-cultural effect on the brain revisited: universal structures plus writing system variation. *Hum. Brain Mapp.* 25, 92–104. doi: 10.1002/hbm.20124

Briellmann, R. S., Saling, M. M., Connell, A. B., Waites, A. B., Abbott, D. F., and Jackson, G. D. (2004). A high-field functional MRI study of quadri-lingual subjects. *Brain Lang.* 89, 531–542. doi: 10.1016/j.bandl.2004.01.008

Cao, F., Brennan, C., and Booth, J. R. (2015). The brain adapts to orthography with experience: evidence from English and Chinese. *Dev. Sci.* 18, 785–798. doi: 10.1111/desc.12245

Cao, F., Lee, R., Shu, H., Yang, Y., Xu, G., Li, K., et al. (2010). Cultural constraints on brain development: evidence from a developmental study of visual word processing in Mandarin Chinese. *Cereb. Cortex* 20, 1223–1233. doi: 10.1093/cercor/bhp186

Cao, F., Peng, D., Liu, L., Jin, Z., Fan, N., Deng, Y., et al. (2009). Developmental differences of neurocognitive networks for phonological and semantic processing in Chinese word reading. *Hum. Brain Mapp.* 30, 797–809. doi: 10.1002/hbm.20546

Cao, F., Tao, R., Liu, L., Perfetti, C. A., and Booth, J. R. (2013a). High proficiency in a second language is characterized by greater involvement of the first language network: evidence from Chinese learners of English. *J. Cogn. Neurosci.* 25, 1649–1663. doi: 10.1162/jocn_a_00414

Cao, F., Vu, M., Chan, D. H., Lawrence, J. M., Harris, L. N., Guan, Q., et al. (2013b). Writing affects the brain network of reading in Chinese: a functional magnetic resonance imaging study. *Hum. Brain Mapp.* 34, 1670–1684. doi: 10.1002/hbm.22017

Chee, M. W., Caplan, D., Soon, C. S., Sriram, N., Tan, E. W., Thiel, T., et al. (1999a). Processing of visually presented sentences in Mandarin and English studied with fMRI. *Neuron* 23, 127–137. doi: 10.1016/s0896-6273(00)80759-x

Chee, M. W., Tan, E. W., and Thiel, T. (1999b). Mandarin and English single word processing studied with functional magnetic resonance imaging. *J. Neurosci.* 19, 3050–3056. doi: 10.1523/JNEUROSCI.19-08-03050.1999

Chee, M. W., Weekes, B., Lee, K. M., Soon, C. S., Schreiber, A., Hoon, J. J., et al. (2000). Overlap and dissociation of semantic processing of Chinese characters, English words, and pictures: evidence from fMRI. *Neuroimage* 12, 392–403. doi: 10.1006/nimg.2000.0631

Cherodath, S., and Singh, N. C. (2015). The influence of orthographic depth on reading networks in simultaneous biliterate children. *Brain Lang.* 143, 42–51. doi: 10.1016/j.bandl.2015.02.001

Costa, A., Pannunzi, M., Deco, G., and Pickering, M. J. (2016). Do bilinguals automatically activate their native language when they are not using it? *Cogn. Sci.* 41, 1629–1644. doi: 10.1111/cogs.12434

Craddock, R. C., James, G. A., Holtzheimer, P. E. III, Hu, X. P., and Mayberg, H. S. (2012). A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Hum. Brain Mapp.* 33, 1914–1928. doi: 10.1002/hbm.21333

De Bleser, R., Dupont, P., Postler, J., Bormans, G., Speelman, D., Mortelmans, L., et al. (2003). The organisation of the bilingual lexicon: a PET study. *J. Neurolinguistics* 16, 439–456. doi: 10.1007/s11910-002-0037-y

Dice, L. R. (1945). Measures of the amount of ecologic association between species. *Ecology* 26, 297–302.

Fiez, J. A., and Petersen, S. E. (1998). Neuroimaging studies of word reading. *Proc. Natl. Acad. Sci. U.S.A* 95, 914–921. doi: 10.1073/pnas.95.3.914

Golestani, N., Alario, F. X., Meriaux, S., Le Bihan, D., Dehaene, S., and Pallier, C. (2006). Syntax production in bilinguals. *Neuropsychologia* 44, 1029–1040. doi: 10.1016/j.neuropsychologia.2005.11.009

Green, D. W. (2003). "The neural basis of the lexicon and the grammar in L2 acquisition," in *The Interface Between Syntax and the Lexicon in Second Language Acquisition*, eds A. H. R. van Hout, F. Kuiken, and R. Towell (Amsterdam: John Benjamins), 197–208. doi: 10.1080/02643294.2014.882814

Hernandez, A. E., Hofmann, J., and Kotz, S. A. (2007). Age of acquisition modulates neural activity for both regular and irregular syntactic functions. *Neuroimage* 36, 912–923. doi: 10.1016/j.neuroimage.2007.02.055

Hernandez, A. E., Woods, E. A., and Bradley, K. A. (2015). Neural correlates of single word reading in bilingual children and adults. *Brain Lang.* 143C, 11–19. doi: 10.1016/j.bandl.2015.01.010

Ip, K. I., Hsu, L. S., Arredondo, M. M., Tardif, T., and Kovelman, I. (2016). Brain bases of morphological processing in Chinese-English bilingual children. *Dev. Sci.* 20, 12449.

Johnson, M. H. (2011). Interactive specialization: a domain-general framework for human functional brain development? *Dev. Cogn. Neurosci.* 1, 7–21. doi: 10.1016/j.dcn.2010.07.003

Kim, K. H. S., Relkin, N. R., Lee, K. M., and Hirsch, J. (1997). Distinct cortical areas associated with native and second languages. *Nature* 388, 171–174. doi: 10.1038/40623

Klein, D., Milner, B., Zatorre, R. J., Zhao, V., and Nikelski, J. (1999). Cerebral organization in bilinguals: a PET study of Chinese-English verb generation. *NeuroReport* 10, 2841–2846. doi: 10.1097/00001756-199909090-00026

Koda, K. (1990). Factors affecting second language text comprehension. *Natl. Read. Conference Yearb.* 39, 419–427.

Kroll, J. F., Bogulski, C. A., and McClain, R. (2012). "Psycholinguistic perspectives on second language learning and bilingualism: the course and consequence of cross-language competition," in *Linguistic Approaches to Bilingualism* (Amsterdam: John Benjamnins), 1–24. doi: 10.1075/lab.2.1.01kro

MacWhinney, B. (2012). "The logic of the unified model," in *The Routledge Handbook of Second Language Acquisition*, eds S. M. Gass and A. Mackey (Abingdon: Routledge), 17.

Musso, M., Moro, A., Glauche, V., Rijntjes, M., Reichenbach, J., Buchel, C., et al. (2003). Broca's area and the language instinct. *Nat. Neurosci.* 6, 774–781. doi: 10.1038/nn1077

Perani, D., and Abutalebi, J. (2005). The neural basis of first and second language processing. *Curr. Opin. Neurobiol.* 15, 202–206. doi: 10.1016/j.conb.2005.03.007

Perani, D., Paulesu, E., Galles, N. S., Dupoux, E., Dehaene, S., Bettinardi, V., et al. (1998). The bilingual brain: proficiency and age of acquisition of the second language. *Brain* 121, 1841–1852. doi: 10.1093/brain/121.10.1841

Pugh, K. R., Shaywitz, B. A., Shaywitz, S. E., Constable, R. T., Skudlarski, P., Fulbright, R. K., et al. (1996). Cerebral organization of component processes in reading. *Brain* 119, 1221–1238. doi: 10.1093/brain/119.4.1221

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., and Shulman, G. L. (2001). A default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A* 98, 676–682.

Sakai, K. L., Miura, K., Narafu, N., and Muraishi, Y. (2004). Correlated functional changes of the prefrontal cortex in twins induced by classroom education of second language. *Cereb. Cortex* 14, 1233–1239. doi: 10.1093/cercor/bhh084

Shaywitz, B. A., Skudlarski, P., Holahan, J. M., Marchione, K. E., Constable, R. T., Fulbright, R. K., et al. (2007). Age-related changes in reading systems of dyslexic children. *Ann. Neurol.* 61, 363–370. doi: 10.1002/ana.21093

Stein, M., Federspiel, A., Koenig, T., Wirth, M., Lehmann, C., Wiest, R., et al. (2009). Reduced frontal activation with increasing 2nd language proficiency. *Neuropsychologia* 47, 2712–2720. doi: 10.1016/j.neuropsychologia.2009.05.023

Suchan, B., Yaguez, L., Wunderlich, G., Canavan, A. G., Herzog, H., Tellmann, L., et al. (2002). Neural correlates of visuospatial imagery. *Behav. Brain Res.* 131, 163–168. doi: 10.1016/s0166-4328(01)00373-4

Sugiura, L., Ojima, S., Matsuba-Kurita, H. I, Tsuzuki, D., Katura, T., and Hagiwara, H. (2011). Sound to language: different cortical processing for first and second languages in elementary school children as revealed by a large-scale study using fNIRS. *Cereb. Cortex* 21, 2374–2393. doi: 10.1093/cercor/bhr023

Tan, L. H., Laird, A. R., Karl, L., and Fox, P. T. (2005). Neuroanatomical correlates of phonological processing of Chinese characters and alphabetic words: a meta-analysis. *Hum. Brain Mapp.* 25, 83–91. doi: 10.1002/hbm.20134

Tettamanti, M., Alkadhi, H., Moro, A., Perani, D., Kollias, S., and Weniger, D. (2002). Neural correlates for the acquisition of natural language syntax. *Neuroimage* 17, 700–709. doi: 10.1006/nimg.2002.1201

Waldie, K. E., Haigh, C. E., Badzakova-Trajkov, G., Buckley, J., and Kirk, I. J. (2013). Reading the wrong way with the right hemisphere. *Brain Sci.* 3, 1060–1075. doi: 10.3390/brainsci3031060

Wartenburger, I., Heekeren, H. R., Abutalebi, J., Cappa, S. F., Villringer, A., and Perani, D. (2003). Early setting of grammatical processing in the bilingual brain. *Neuron* 37, 159–170. doi: 10.1016/s0896-6273(02)01150-9

Wu, Y. J., and Thierry, G. (2010). Chinese-English bilinguals reading English hear Chinese. *J Neurosci* 30, 7646–7651. doi: 10.1523/JNEUROSCI.1602-10.2010

Woodcock, R. W., McGrew, K. S., and Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: The Riverside Publishing Company.

Xue, G., Dong, Q., Jin, Z., Zhang, L., and Wang, Y. (2004). An fMRI study with semantic access in low proficiency second language learners. *Neuroreport* 15, 791–796. doi: 10.1097/00001756-200404090-00010

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Neuroscience

ORIGINAL RESEARCH
published: 13 June 2022
doi: 10.3389/fnins.2022.898661

The Role of Brain Activity in Characterizing Successful Reading Intervention in Children With Dyslexia

Anthony J. Krafnick[†]*, Eileen M. Napoliello, D. Lynn Flowers and Guinevere F. Eden**

Center for the Study of Learning, Department of Pediatrics, Georgetown University Medical Center, Washington, DC, United States

Studies of reading intervention in dyslexia have shown changes in performance and in brain function. However, there is little consistency in the location of brain regions associated with successful reading gains in children, most likely due to variability/limitations in methodologies (study design, participant criteria, and neuroimaging procedures). Ultimately for the results to be meaningful, the intervention has to be successful, be assessed against a control, use rigorous statistics, and take biological variables (sex) into consideration. Using a randomized, crossover design, 31 children with dyslexia were assigned to a phonological- and orthographic-based tutoring period as well as a within-subjects control period to examine: (1) intervention-induced changes in behavior (reading performance) and in brain activity (during reading); and (2) behavioral and brain activity pre-intervention data that predicted intervention-induced gains in reading performance. We found gains in reading ability following the intervention, but not following the control period, with no effect of participants' sex. However, there were no changes in brain activity following the intervention (regardless of sex), suggesting that individual brain changes are too variable to be captured at the group level. Reading gains were not predicted by pre-intervention behavioral data, but were predicted by pre-intervention brain activity in bilateral supramarginal/angular gyri. Notably, some of this prediction was only found in females. Our results highlight the limitations of brain imaging in detecting the neural correlates of reading intervention in this age group, while providing further evidence for its utility in assessing eventual success of intervention, especially if sex is taken into consideration.

Keywords: dyslexia, reading disability, intervention, fMRI, children

Edited by:
Li-Hai Tan,
Shenzhen Institute of Neuroscience, China

Reviewed by:
Fan Cao,
Sun Yat-sen University, China
Heikki Juhani Lyytinen,
Niilo Mäki Institute, Finland

Correspondence:
Guinevere F. Eden
edeng@georgetown.edu

[†]***Present address:***
Anthony J. Krafnick,
Department of Psychology, Dominican University, River Forest, IL, United States

Specialty section:
This article was submitted to Neurodevelopment, a section of the journal Frontiers in Neuroscience

Received: *17 March 2022*
Accepted: *23 May 2022*
Published: *13 June 2022*

Citation:
Krafnick AJ, Napoliello EM, Flowers DL and Eden GF (2022) The Role of Brain Activity in Characterizing Successful Reading Intervention in Children With Dyslexia. Front. Neurosci. 16:898661. doi: 10.3389/fnins.2022.898661

INTRODUCTION

Developmental dyslexia is a common learning disability, affecting approximately between 5 and 13% of the general United States population (Katusic et al., 2001). It is defined by difficulties in word recognition and word decoding, that are incongruent with other cognitive skills, classroom experience, and motivation to learn how to read (Lyon et al., 2003). The word decoding problems (sounding out of novel words) are believed to be due to difficulties with understanding how sounds in speech are isolated, manipulated, and recalled (phonological coding or phonological awareness; Scarborough and Brady, 2002) and other difficulties with representing the speech stream (Peterson and Pennington, 2012), therefore impeding their mapping onto their corresponding graphemes during reading. Further, word form recognition is also impaired in dyslexia, not only as a consequence of poor phonological coding skills, but also because of difficulties in establishing a "sight word vocabulary" through memorization of the visual (orthographic) word forms (Badian, 1995, 2001). Brain imaging studies have revealed hypoactivation in brain regions associated with phonological mapping in temporo-parietal cortex (TPC) and in regions associated with visual word form recognition in the occipito-temporal cortex (OTC) (Pugh et al., 2001; Sandak et al., 2004; Eden et al., 2015). The role of these areas in reading acquisition and in dyslexia continue to be an active area of discussion (Richlan, 2012).

When it comes to addressing the reading difficulties of children with dyslexia, explicit instructions in phonological coding (Alexander and Slinger-Constant, 2004), letter-speech sound training (Brem et al., 2010; Kyle et al., 2013) and orthographic facilitation (Baron et al., 2018) are often key elements in the interventions administered. Ideally, such tutoring occurs in small groups (or one-on-one) with an emphasis on early intervention (Wanzek and Vaughn, 2007). Understanding the kinds of reading interventions that work and how they work, is of continued interest. Investigations into the neural correlates of reading intervention in dyslexia have shown widespread increases in brain activity in children (Aylward et al., 2003; Temple et al., 2003; Shaywitz et al., 2004; Richards et al., 2006; Meyler et al., 2008; Odegard et al., 2008; Gebauer et al., 2012) and adults (Eden et al., 2004). A narrative review by Barquero et al. (2014) describes these to include left and right hemisphere inferior, middle, and superior frontal gyri, superior and middle temporal gyri, occipital cortices, inferior parietal lobule, post central gyrus, and insulae (Barquero et al., 2014). Notably it has been suggested that gains in reading in dyslexia are associated with increases in left-hemisphere regions typically involved in reading, while other regions, such as right frontal cortex (Temple et al., 2003; Richards et al., 2006; Meyler et al., 2008), perhaps serve in a compensatory role (Shaywitz et al., 2004; Hoeft et al., 2007; Barquero et al., 2014). However, a meta-analysis of intervention neuroimaging research demonstrates a lack of convergence across studies for many of these implicated brain regions (Barquero et al., 2014). The strongest results from this meta-analysis of eight studies (173 participants) were left thalamus (three studies contributing), right insula/inferior frontal gyrus (four studies contributing), and left inferior frontal gyrus (three studies contributing). Right posterior cingulate and left middle occipital gyrus were also identified (though with only two contributing studies). Notably, other left-hemisphere regions typically involved in reading, that is left OTC and PTC, were not found to change. A more recent meta-analysis of changes in brain activation following reading intervention of eight studies (151 participants) and differing from the Barquero meta-analysis by two out of eight studies, found no results (Perdue et al., 2022).

There are also methodological limitations that question the validity of prior findings on changes in brain activity following reading intervention in dyslexia. As a whole, the imaging thresholds used are far less stringent than those used today. Of the studies in children and adolescents included in the Barquero and Perdue meta-analyses, only four used any correction of cluster size in their whole brain analysis (Aylward et al., 2003; Gebauer et al., 2012; Heim et al., 2015; Partanen et al., 2019), and voxel level thresholds vary considerably. Another concern is the variability in reported behavioral gains associated with the interventions. Not all studies report on changes in single word reading ability (though most report comprehension level data), and only two have examined whether these gains persist in the long-term (Shaywitz et al., 2004; Meyler et al., 2008). Importantly, while many studies of dyslexia include a control group, none include a within-subject control intervention to assess specificity of these changes, a design that is considered best practice in clinical research.

In addition to examining the brain bases for reading disability and successful reading intervention, neuroimaging has also been used in a small number of studies to examine whether brain function can be used to "predict" later reading outcome in typically reading children as well as children with dyslexia; and these studies have been done either with or without conducting a formal intervention. For example, in typically developing children both left visual word system (fusiform gyrus) event-related potentials and functional magnetic resonance imaging (fMRI) signal attained during a lexical decision task in kindergarten (prior to learning to read) were found to be predictive of how well children (who participated in a speech-sound association training) could read in second grade (Bach et al., 2013). Specific to dyslexia, a study in children and adolescents with dyslexia found that right inferior frontal gyrus activation during a written word rhyming task predicted single word reading measures 2.5 years later (no intervention was provided) (Hoeft et al., 2011). Two studies have examined reading intervention in children with dyslexia and tested whether gains in reading following the intervention were predicted by pre-intervention brain activity. One study found gains in untimed pseudoword reading were predicted by pre-intervention activity during a phonological processing task in left inferior frontal gyrus, and gains in timed word reading were predicted by activity in left and right inferior frontal gyri (Farris et al., 2016). Another study found gains in basic reading were predicted by pre-intervention functional connectivity between middle temporal gyrus and left inferior parietal lobule during a lexical decision task (Aboud et al., 2018). Like studies investigating changes in brain activity with intervention, the use of cluster-level correction for the whole-brain analysis is mixed, with only the last two of the above mentioned studies using cluster level correction (Farris et al., 2016; Aboud et al., 2018).

There have been recent calls to pay more attention to sex as a biological variable in all research (Cahill, 2006, 2012) and especially in research of language processing and dyslexia (Ramus et al., 2018; Krafnick and Evans, 2019). Sex has played a role in language research, where converging evidence suggest sex-specific differences in language acquisition and development (Martin and Hoover, 1987; Bornstein et al., 2000; Dionne et al., 2003), as well as sex-specific organization of the brain for language (Shaywitz et al., 1995; Jaeger et al., 1998; Kansaku and Kitazawa, 2001; Burman et al., 2008). For example, males have been shown to have more left-lateralized activation of perisylvian brain regions during language tasks, whereas females activate bilateral perisylvian brain regions (Shaywitz et al., 1995; Jaeger et al., 1998).

Specific to dyslexia, prevalence differs amongst boys and girls with odds ratios ranging from 1.39 to 3.19 in favor of higher prevalence in boys (Rutter et al., 2004; Quinn and Wagner, 2015), even when controlling for ascertainment bias (Liederman et al., 2005; Quinn and Wagner, 2015). Neuroimaging studies of dyslexia have on average recruited more male subjects, as reflected in 65% male participants contributing to the meta-analysis in children, and 95% to the meta-analysis in adults reported by Richlan et al. (2011), and 59% of subjects in a meta-analysis of neuroimaging studies of reading intervention (Barquero et al., 2014). Most importantly, evidence of sex-specific differences in dyslexia from studies of gray matter volume (Evans et al., 2014) and cortical thickness (Altarelli et al., 2013; Clark et al., 2014) suggest that the brain bases of dyslexia may not be the same in males and females, with females, but not males, showing differences in anatomy in brain regions associated with early sensory processing (Altarelli et al., 2013; Clark et al., 2014; Evans et al., 2014). However, sex has not been accounted for in investigations into the brain-based correlates of successful reading intervention in dyslexia. Sex-specific differences in dyslexia prior to an intervention could lead to sex-specific differences in the neural correlates of successful intervention. Critical to this study, failure to identify any antagonist interactions for sex could result in failure to register significant (sex-specific) changes following intervention. That is, if sex-specific changes are in opposite directions, the changes during intervention could appear small or non-existent. Lastly, if the behavioral response to reading intervention is the same for males and females, it does not mean that the neural substrates underlying that change in performance is the same for both sexes (Cahill, 2006).

In the present study we report behavioral data for reading and reading-related skills as well as fMRI data during a word processing task in 31 children with dyslexia. These data were acquired in all children prior to and following (i) an intensive intervention focused on promoting reading through phonological and orthographic skills, and (ii) an intensive intervention focused on promoting math (active control) or, instead, a null period (developmental control), using a randomized, crossover design. Our study of dyslexia allowed us to ask: (1) What are the brain activation changes that follow a successful reading intervention, and are these changes specific to the reading intervention? (2) Can brain activity during reading indicate whether children will subsequently reap benefits from the reading intervention? And (3) Are these findings affected by sex? Together these findings should advance our understanding of the location and specificity of the neural correlates that underlie successful reading intervention in males and females with dyslexia, as well as whether brain activity signals a readiness to benefit from such a reading intervention.

MATERIALS AND METHODS

Participants

Thirty-one dyslexic children (14 female; age average 9.6 and range 7.4–12.6 years) were recruited from a private school specializing in students with learning disabilities. School records were used to identify children who had a score of less than or equal to 92 on the Woodcock–Johnson Test of Achievement III Letter-Word Identification (W-J WID) and/or Word Attack (W-J WA) (Woodcock et al., 2001), and a documented diagnosis of dyslexia. In order to be included in the study, children had to score at least 80 on Verbal, Performance, and Full IQ on the Wechsler Abbreviated Scale of Intelligence (WASI) (Wechsler, 1999). All children were in good health and free of other developmental disabilities, neurological and psychiatric disorders or any disease affecting brain function, except for ADHD (children taking medication for ADHD had to refrain taking it prior to the scans). Other exclusion criteria included contraindications to MRI scanning such as metallic implants or severe claustrophobia. fMRI data for some of these children using the same reading task have been published previously in a comparison with typically reading children (Olulade et al., 2015).

Behavioral Tests

All subjects received a battery of psychoeducational tests to evaluate intelligence quotient (IQ), reading, and skills that are related to reading. Except for IQ, the entire testing battery was administered at all three visits (prior to and following interventions). The WASI (Wechsler, 1999) was used to measure IQ. The Woodcock–Johnson Test of Achievement III was used to assess reading ability: Word Identification (W-J WID) subtest for single real word reading, Word Attack (W-J WA) subtest for single pseudoword reading, and Passage Comprehension (W-J PC) for understanding of written text (Woodcock et al., 2001). In addition, we measured skills that play a role in acquiring reading and are typically impaired in dyslexia: the Lindamood Auditory Conceptualization Test (LAC) for phonemic awareness (Lindamood and Lindamood, 1971), the Rapid Automatized Naming test (RAN L&N and C&O) for naming fluency of letters/numbers and colors/objects (Denckla and Rudel, 1976a,b), the Digit Span test for working memory (Wechsler, 1999), and the Symbol Imagery (SI) test for visual imagery (memory for letters and orthographic patterns) (Bell, 1997). These measures were used to gauge improvement in reading and reading-related skills, which were expected to increase following the reading intervention but not following the math intervention. To also assess changes in mathematical performance, we used the Calculation subtest for computational ability, the Math Fluency subtest for timed arithmetic, and the Applied Problems subtest of the Woodcock–Johnson Test of Achievement III (Woodcock et al., 2001) for mathematical word problems. All scores reported are standard scores (Population Mean = 100, SD = 15). Researchers acquiring these data were blind to each child's group assignment.

Study Design

The children were randomly assigned to one of three groups. Each of the three groups received the same reading intervention (3 h a day, for a total of 90 h). For Group 1 (n = 10) this 6-week reading intervention was followed by a 6-week math intervention, and for Group 2 (n = 9) it was preceded by a math intervention (math intervention was also 3 h a day, for a total of 90 h). As such, these 19 children received the intervention of interest (reading) and an active control intervention (math), with the order counterbalanced (a 10th child originally assigned to Group 2 left the study after it had begun). Group 3 (*n* = 12) received the same reading intervention followed by a 6-week null period (no intervention) to provide a developmental control period (Krafnick et al., 2011). As such, we would be able to weigh any benefits resulting from the reading intervention against the possibility of a Hawthorn effect and/or a placebo effect (by comparison to the active control math intervention). The latter effects could result from participating in a study that involves intensive work on the part of the participants as well as strong encouragement by others for their efforts. Further, both the reading intervention and the active control math intervention could be assessed relative to no intervention (null period) to be able to assess changes relative to the normal developmental changes that would occur during this time span. Three behavioral testing/scanning sessions were scheduled eight weeks apart (one prior to any intervention/control period and another after each intervention/control period; see **Figure 1**). One-way ANOVAs showed that randomization to group was successful in keeping the groups similar in age, IQ, reading, and reading-related skills prior to intervention (**Table 1**). As such, age and IQ were not included in the analyses looking at gains in performance measures following the interventions. Further, a Chi-square test revealed no significant difference in sex amongst the groups (**Table 1**). Most subjects (26 of the original 31) returned for behavioral testing 1 year later, allowing us to gauge long-term outcome of the intervention.

Reading and Math Interventions

All 31 children received the reading intervention Seeing Stars® (Bell, 1997) purchased by us from Lindamood-Bell Learning Processes® and delivered by their employees at the children's school in small groups. This intervention uses a "multisensory approach" to promote integration of internal visual and phonological representations of letters and letter strings. The imagery portion increases in difficulty starting with single letter imagery and increasing through two and three syllable words. A tactile/motor portion involves finger tracing of visualized letters, and a language production portion involves aloud verbalization of letter and syllable sounds while they are finger traced in the air. The use of imagery/visualization in this reading intervention is based on several studies involving the use of imagery in reading, including self-report of imagery during reading (Long et al., 1989), imagery in semantic retrieval (Kosslyn, 1976) and the use of imagery to improve processing and comprehension (Linden and Wittrock, 1981; Sadoski, 1983).

The math intervention was On Cloud Nine® (Bell and Tuley, 1997) and was delivered by the same Lindamood-Bell Learning Processes® staff at the school. It utilizes a multisensory approach focusing on imagery, tracing and verbalization, similar to the reading intervention, but with a focus on numbers and number lines instead of letters and syllables, thus serving as a suitable active control for the reading intervention.

Behavioral Analyses

To test for performance changes brought about by the interventions, we conducted 2 × 2 repeated measures ANOVAs with Time Point (pre- versus post intervention) as a within-subjects factor and sex (male versus female) as a between-subjects factor (Time Point × Sex). Each analysis was specific to a reading or reading-related measure and a particular intervention, e.g., children's pre- and post-reading intervention data for a given measure were included as "Time Points" to examine changes during the reading intervention. We refer to Time Points here as opposed to Visits in the description of the intervention design above (**Figure 1**) because of the counterbalanced design. For example, some participants' pre-reading intervention visit was Visit 1, whereas others it was Visit 2; to investigate changes following the reading intervention, we use data from each subject's pre- and post-reading intervention Time Points. We employed a Holm–Bonferroni correction for multiple comparisons (Holm, 1979). These pre- versus post intervention comparisons are similar to those presented in previous reading intervention studies (that did not include a control period) and are presented here for the purpose of comparison with behavioral gains following intervention in those studies.

However, to test whether any such gains in reading during the reading intervention are significantly greater than any gains during the control period, we conducted a 2 × 2 repeated measures ANOVA using intervention (reading intervention period versus math intervention/null period) and performance measure (change in reading ability on W-J WID versus change in math ability on Calculation standard score) as within-subjects factors and tested for an interaction.

To test if performance measures predicted reading gains, the three reading and the five reading-related measures at Visit 1 (prior to any intervention), age, IQ, and sex were entered into a single multiple regression with change in reading ability on W-J WID standard score as the dependent variable.

Behavioral analyses and visualization were carried out in SPSS (IBM SPSS Statistics 22), jamovi (version 2.5.5), and Microsoft Excel.

Functional Magnetic Resonance Imaging Data Acquisition and Preprocessing

During acquisition of fMRI data, subjects performed an implicit reading task (Price et al., 1996). The children saw single real words (Word) or false font strings (False Font) and responded with a button press in their right hand if the Word or False Fonts contained a "tall" letter or character (e.g., "alarm" contains the tall letter "l") and a button press in their left hand if it did not (e.g., sauce has no tall letters). This task has been used previously in our studies of reading and reading disability (Turkeltaub et al., 2003, 2004; Olulade et al., 2013, 2015; Evans et al., 2016). Blocks of Word and blocks of False Font stimuli alternated (twice each) and were separated by blocks of Fixation. Blocks of Words and False Font contained 10 trials each, lasting 42 s, and Fixation blocks lasted 18 s each (with additional Fixation scans at the beginning and end of the run, resulting in a total scan time of four minutes, twenty-seven seconds). Each child underwent two

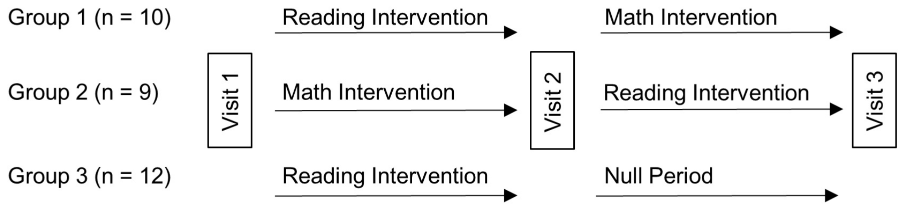


FIGURE 1 | Intervention study design. Thirty-one children with dyslexia were randomly assigned to one of three intervention arms. All groups received the reading intervention, and either a math intervention (active control, Groups 1 and 2) or null period (developmental control, Group 3). Each visit consisted of acquisition of behavioral and imaging data, with 6-week intervention periods between Visits 1 and 2, and Visits 2 and 3.

TABLE 1 | Behavioral profile prior to intervention (Visit 1) for entire group and by intervention group.

	Mean (SD)					
	All subjects (*n* = 31)	**Group 1 (*n* = 10)**	**Group 2 (*n* = 9)**	**Group 3 (*n* = 12)**	***F*-statistic/Chi-square**	***p*-Value**
Age (years)	9.6 (1.5)	10.0 (1.6)	9.9 (1.5)	9.0 (1.3)	1.414	0.260
Sex (M/F)	14/17	7/3	4/5	3/9	4.46	0.107
IQ: Wechsler Abbreviated Scale of Intelligence (WASI)*					*F* (2,27)	
Verbal IQ	110.2 (9.0)	111.9 (11.6)	107.5 (11.4)	110.7 (3.4)	0.538	0.590
Performance IQ	101.9 (10.2)	100.7 (10.8)	104.8 (8.9)	101.1 (10.9)	0.403	0.672
Full IQ	106.9 (8.4)	107.2 (11.6)	106.8 (7.6)	106.8 (6.2)	0.007	0.993
Measures of reading: Woodcock–Johnson					*F* (2,28)	
Word Identification (single real words)	77.4 (8.0)	80.3 (7.9)	75.8 (7.3)	76.2 (8.6)	0.982	0.387
Word Attack (single pseudowords)	91.8 (6.4)	93.0 (5.1)	89.2 (6.0)	92.7 (7.5)	1.018	0.374
Passage Comprehension (reading comprehension)	78.4 (13.9)	84.4 (9.9)	77.0 (14.7)	74.4 (15.4)	1.528	0.235
Measures of skills that support reading					*F* (2,28)	
Lindamood Auditory Conceptualization Test (phonemic awareness)	98.4 (8.4)	102.4 (9.9)	96.1 (8.4)	96.8 (6.4)	1.783	0.187
Rapid Naming (naming fluency for letters and numbers)	78.6 (12.4)	84.4 (11.6)	77.8 (9.7)	74.3 (13.7)	1.974	0.158
Rapid Naming (naming fluency for colors and objects)	84.5 (12.4)	90.7 (13.5)	79.2 (10.1)	83.3 (11.7)	2.313	0.118
Digit Span (working memory)	93.7 (11.8)	99.0 (12.9)	92.8 (12.8)	90.0 (9.3)	1.693	0.202
Symbol Imagery (visual imagery/orthographic processing)	80.3 (9.7)	84.4 (9.8)	74.7 (11.8)	80.8 (8.5)	2.502	0.100

**WASI scores were missing for one subject.*

scan acquisitions (two runs yielding 28 whole-head echo planar imaging (EPI) volumes for each condition, Word, False Font and Fixation) at three different times over the study (one at each visit as described above, see **Figure 1**).

Functional magnetic resonance imaging fMRI data was acquired using an EPI sequence using a 3 Tesla Siemens Trio whole-body MRI system [TE = 30 ms, TR = 3 s, 64 × 64 matrix, 192 mm FOV, 50 axial slices, 2.8 mm slice thickness (0.2 mm interslice gap) yielding 3 mm cubic voxels, flip angle 90°]. A high resolution, 3D T1-weighted MPRAGE image obtained at the outset of the study (Visit 1, prior to any intervention) on the same Siemens Trio whole-body MRI system was used to aid in anatomical localization of the fMRI data.

Pre-processing for functional analysis began by segmenting the subjects' MPRAGE images and normalizing to a standard template brain (Montreal Neurological Institute, MNI). For all functional runs, the first five scans were removed, and the remaining scans were corrected for head motion by realigning to the mean image, co-registered to the subjects MPRAGE, normalized using the same parameters for the MPRAGE image and finally smoothed using a 6 mm × 6 mm × 5.8 mm Gaussian kernel. For each subject's first level analysis, both runs were included, and contrasts were generated for the Word versus False Font condition, Word versus Fixation condition and False Font versus Fixation condition. Motion parameters and global mean signal were included as regressors of no interest to account for subject movement and global signal variation during each run.

Functional Magnetic Resonance Imaging Group Level Analyses

All analyses were carried out in SPM (Statistical Parametric Mapping, Wellcome Trust Centre for Neuroimaging, London, United Kingdom). All group analyses (differences in activation pre- versus post the intervention, and activation to predict intervention-induced changes in reading performance) were performed on Words > False Font contrasts at an uncorrected height threshold of $p < 0.001$, and an extent threshold of $p < 0.05$ family wise error (FWE) corrected. For the analysis on activation to predict intervention-induced reading gains, the MarsBaR toolbox (Brett et al., 2002) was used to extract the signal from clusters identified in the analyses (described below), to display the mean percent signal change. Again, sex was included as a between-subjects factor in the intervention Time Point comparisons (same as in the behavioral analyses).

RESULTS

Behavioral Measures Change Following the Reading Intervention

To evaluate the impact of the reading intervention, the 2 × 2 repeated measures ANOVAs conducted on the standard scores of the three measures of reading and the five measures of reading-related skills, as well as the three math skills, immediately prior to and following the reading intervention (Time Point as within-subjects factor and Sex as between-subjects factor) found six of the eight reading/reading-related measures showed a significant main effect of Time Point and increased scores following the reading intervention (**Table 2**): real word reading (W-J WID), pseudoword reading (W-J WA), reading comprehension (W-J PC), phonemic awareness (LAC), naming fluency of letters/numbers (RAN L&N), and visual imagery (SI). One of the math measures (W-J Math Fluency) showed a significant main effect of Time Point, decreasing after the reading intervention. After Holm–Bonferroni correction (Holm, 1979), all three reading measures (real word reading, pseudoword reading, and passage comprehension), two of the five reading-related measures (phonemic awareness and visual imagery), and none of the math measures remained significant. As such the children made gains on a range of measures of reading, as well as the skills targeted by the intervention and known to promote reading acquisition (see **Figure 2**). There were no significant interactions for Time-Point × Sex.

When the same analyses were conducted following the control periods (math and no intervention) none of the reading or reading-related measures changed (**Table 2**), demonstrating the specificity of the above-described effects of the reading intervention. To ensure this result was not because one type of control period (e.g., math intervention) had effects which were canceled out or diluted by opposite effects of the other control period (e.g., no intervention developmental control), a one-way ANOVA for Group on changes following the reading intervention was conducted for these two specific arms of the study and showed no significant differences (**Table 3**). While there were no gains made in reading in the groups receiving the math intervention, this intervention resulted in significant gains on measures of mathematics. Specifically, there were significant main effects of Time Point for the measures of mathematic computational ability (W-J Calculation), timed arithmetic (W-J Math Fluency), and mathematical word problems (W-J Applied Problems), with all measures increasing over this time period. The first two results remain significant after Holm–Bonferroni correction (**Table 2**), demonstrating that the math intervention was successful and specific in bringing about gains in the domain of math (see **Figure 2**). There were no significant interactions for Time-Point × Sex during the math intervention.

While these analyses allow for direct comparison with previous studies that did not include control periods in their experimental design, we ultimately wanted to test whether gains in reading were statistically greater during the reading intervention period compared to the control periods. For this we conducted a 2 × 2 repeated measures ANOVA for intervention (reading intervention versus math intervention/null period) × measure (change in reading ability on W-J WID versus change in math ability on Calculation standard score) and specifically examined the interaction. The interaction was significant [$F(1,30) = 18.52$, $p < 0.001$], and the *post hoc* test of reading change during the reading intervention versus reading change during the math intervention/null period was significant [$t(30) = 3.81$, $p = 0.004$). During the reading intervention the average reading (W-J WID) standard score change was 6.84, while during the math intervention/null period it was −0.29. Similarly, for the average math (Calculation) standard score change during the math intervention/null period there was an average increase of 7.84, while during the reading intervention scores decreased by −0.26.

Lastly, turning to the follow-up visit, there were no significant differences on the standardized reading measures between the time the children completed the reading intervention and 1 year later (12.4 months on average; assessed in 26 of the original 31 participants; Digit Span was not assessed) indicating that the children maintained the same level of performance they had reached at the end of the intervention. Specifically, as the raw scores on average increased, the standardized measures revealed no significant changes ($p > 0.05$) for any of the eight measures listed above.

Behavioral Measures as Predictors of Intervention-Induced Gains in Reading

We next investigated whether our behavioral measures of reading and reading-related skills prior to the intervention were predictive of the reading gains made in single real word reading (W-J WID) by the completion of the reading intervention. That is, in the whole group ($n = 31$), each of the eight reading or reading-related measures at Visit 1, along with age, Full IQ, and sex were entered into a multiple regression with change in single real word reading (W-J WID difference before and after the reading intervention) standard score as the dependent variable. The model was not significant [$F(11,18) = 0.523$, $p = 0.863$], and none of the variables contributed to predicting single word reading score change (all p-values > 0.05). There was evidence of collinearity among several of the reading and reading-related measures (VIF > 2.5). Removing these specific variables (pseudoword reading, passage comprehension, SI, and age) showed no improvement [$F(7,22) = 0.788$, $p = 0.605$) and again none of the variables contributed to predicting single real word reading score change (all p-values > 0.05).

Brain Activation Changes Following the Reading Intervention

A 2 × 2 ANOVA (Time Point × Sex) to examine changes in brain activity following the reading intervention, and whether sex played a role, yielded

TABLE 2 | Changes in behavior following intervention.

	Mean (SD)		Main effect of time point	Time point × sex interaction
Pre versus post period of reading intervention	**Pre**	**Post**	***p*-Value**	***p*-Value**
Measures of reading: Woodcock–Johnson				
Word Identification (single real words)	77.5 (7.9)	84.4 (9.2)	*3.0×10^{-6}	0.154
Word Attack (single pseudowords)	91.0 (7.0)	96.9 (7.3)	*1.0×10^{-6}	0.610
Passage Comprehension (reading comprehension)	79.0 (12.3)	85.6 (7.8)	*2.3×10^{-4}	0.710
Measures of skills that support reading				
Lindamood Auditory Conceptualization Test (phonemic awareness)	98.0 (8.4)	102.7 (10.7)	*0.005	0.123
Rapid Naming (naming fluency for letters and numbers)	79.0 (12.6)	82.7 (13.5)	0.029	0.399
Rapid Naming (naming fluency for colors and objects)	85.3 (12.6)	85.7 (15.6)	0.909	0.314
Digit Span (working memory)	92.3 (10.6)	93.7 (10.5)	0.533	0.879
Symbol Imagery (visual imagery/orthographic processing)	81.4 (8.4)	94.1 (12.3)	*1.0×10^{-8}	0.077
Math skills				
Calculation (computational ability)	96.6 (13.5)	95.3 (10.6)	0.753	0.258
Math Fluency (timed arithmetic)	86.4 (12.8)	80.7 (15.0)	0.019	0.929
Applied Problems (mathematical word problems)	97.7 (9.5)	96.8 (7.9)	0.719	0.470
Pre versus post period of math intervention/no intervention	**Mean (SD)**			
	Pre-MI/NI	**Post-MI/NI**	***p*-Value**	***p*-Value**
Measures of reading: Woodcock–Johnson				
Word Identification (single real words)	82.9 (9.6)	82.6 (11.6)	0.924	0.146
Word Attack (single pseudowords)	96.0 (7.2)	96.1 (8.9)	0.923	0.874
Passage Comprehension (reading comprehension)	84.0 (10.7)	84.2 (10.6)	0.772	0.066
Measures of skills that support reading				
Lindamood Auditory Conceptualization Test (phonemic awareness)	102.4 (11.0)	100.1 (9.4)	0.226	0.088
Rapid Naming (naming fluency for letters and numbers)	83.0 (13.3)	82.1 (12.3)	0.392	0.515
Rapid Naming (naming fluency for colors and objects)	86.3 (15.1)	85.5 (13.9)	0.647	0.192
Digit Span (working memory)	93.7 (10.6)	92.7 (12.0)	0.725	0.408
Symbol Imagery (visual imagery/orthographic processing)	90.7 (14.6)	89.4 (12.8)	0.532	0.387
Math skills^				
Calculation (computational ability)	90.7 (12.0)	106.1 (15.3)	*3.0×10^{-6}	0.445
Math Fluency (timed arithmetic)	80.8 (13.9)	90.0 (16.0)	*1.8×10^{-4}	0.478
Applied Problems (mathematical word problems)	96.4 (7.2)	99.0 (8.6)	0.027	0.485

**p-Values in survive Holm–Bonferroni correction for multiple comparisons. ^Results limited to the two groups that received the math intervention, to evaluate effectiveness of the math intervention.*

no significant results for main effect of Time Point, or for Time Point × Sex interaction (height threshold of $p < 0.001$, uncorrected, extent threshold $p < 0.05$ FWE corrected).

As there were no significant findings in this first analysis examining changes in activation following the reading intervention, there no longer was a need to assess these pre-post differences in comparison to the control periods (math intervention/null).

Brain Activation as Predictors of Intervention-Induced Gains in Reading

A simple regression analysis of brain activity during reading task at Visit 1 with change in single real word reading (W-J WID difference before and after the reading intervention), revealed two clusters, one in left and the other in right supramarginal/angular gyri (BA 39/40) (height threshold of $p < 0.001$ uncorrected, extent threshold of $p < 0.05$ FWE, corrected) as depicted in **Figure 3** (see **Table 4**). To examine if this predictive relationship was specific to the reading intervention, the signal in these two clusters (activity during reading task at Visit 1) was submitted for correlations with change in single real word reading standard scores during the math intervention (as above, W-J WID standard score differences prior to and immediately after intervention), but neither cluster was significant, showing that the predictive powers in this region were specific to outcomes following the reading intervention.

This relationship between brain activity during reading task at Time Point 1 and change in reading score was examined for females and males separately. In females only, the relationships were very strong in both the left hemisphere ($r = 0.848$, $p = 1.3 \times 10^{-4}$) and right hemisphere ($r = 0.803$, $p = 5.4 \times 10^{-4}$). In males only, the relationships was not significant in the left hemisphere ($r = 0.409$, $p = 0.103$) and barely significant in the right hemisphere ($r = 0.483$, $p = 0.049$) as visualized by scatterplots in **Figure 3**.

DISCUSSION

In the present study we studied children with dyslexia: (1) to test for intervention-induced changes in reading performance and in

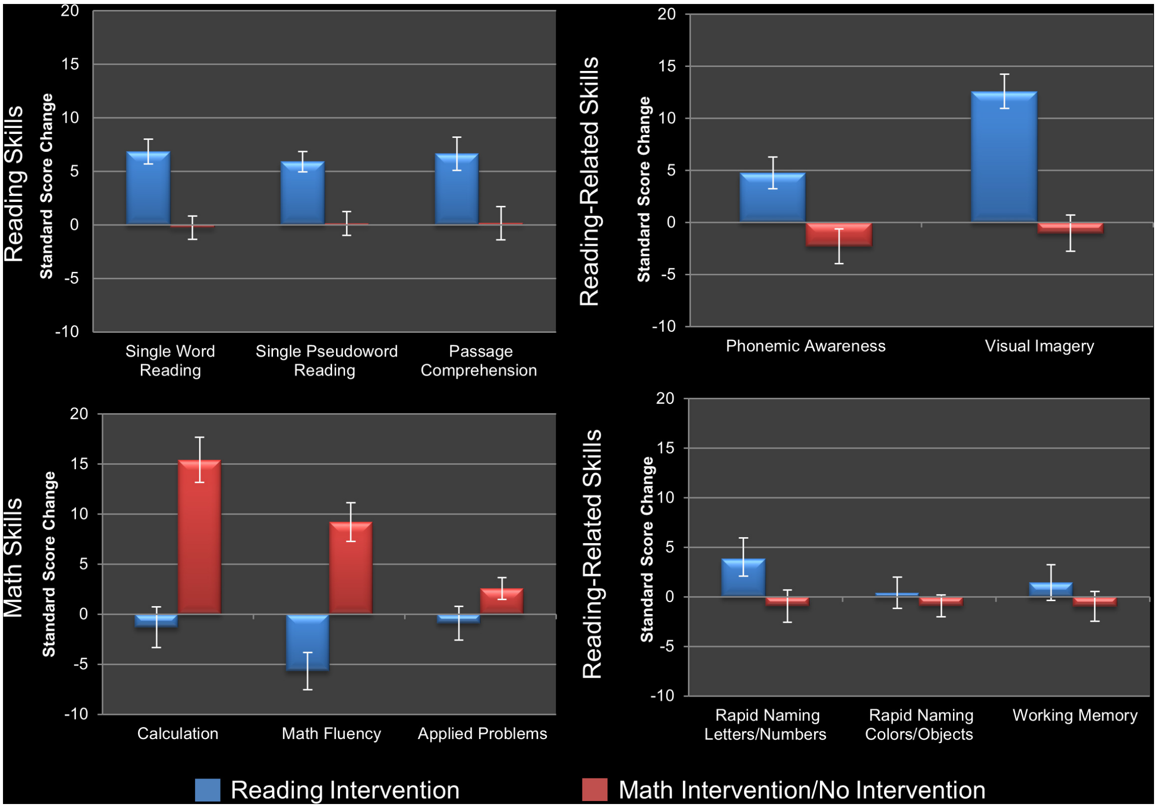


FIGURE 2 | Changes in behavioral measures following intervention. Visualization of standard score changes in measures of reading skills (top left), phonological and orthographic reading-related skills (top right), other reading-related skills (bottom right), and math skills (bottom left) for the analyses reported in **Table 2**. Reading skills: single real word reading (Word Identification), single pseudoword reading (Word Attack), and reading comprehension (Passage Comprehension). Phonological and orthographic reading-related skills: phonemic awareness (Lindamood Auditory Conceptualization Test), and visual imagery/orthographic processing (Symbol Imagery). Other reading-related skills: naming fluency of letters/numbers and colors/objects (Rapid Automatized Naming test, and working memory (Digit Span). Math skills: computational ability (Calculation), timed arithmetic (Math Fluency), and mathematical word problems (Applied Problems). Error bars show standard error for the average change in standard score. For statistical tests see text.

brain activity during reading; and (2) to determine if behavioral measures or brain activity prior to the reading intervention predicted intervention-induced gains in reading. We used a cross-over design allowing us to directly compare reading intervention outcomes with a control period (within-subject control). Overall, the children made strong gains in reading performance (single real word reading, single pseudoword decoding and reading comprehension) as well as the two skills trained during the intervention (phonological and orthographic processing). These gains were specific to the reading intervention as the control math intervention resulted in gains on math but not reading measures, with an ANOVA showing a clear dissociation of the effects of the reading intervention period versus the math/null control period on reading performance versus math performance. However, there were no significant changes in brain activity following the reading intervention. On the other hand, while behavioral measures prior to the onset of the intervention did not predict reading gains made during the reading intervention, brain activation during reading prior to the reading intervention did predict reading gains made during the reading intervention (in left and right supramarginal/angular gyri). Interestingly, while sex was not a significant factor in any of the analyses up until this point, this predictive relationship between pre-intervention brain activity and reading gains following the reading intervention was significant in female subjects, whereas males showed no significant relationship in the left hemisphere and barely in the right hemisphere. These results show that brain activity does not shed light on the neural bases of a successful and enduring reading intervention, but unlike measures of behavior, it identifies regions that signal a level of brain activity that indicates eventual treatment success; and this predictive signal is manifest strongly in females, but not males.

Behavioral Measures Change Following Reading Intervention

Our study showed performance gains in reading and reading-related skills following the reading intervention. Notably gains occurred on all three measures of reading, namely single real word reading, pseudoword decoding, as well as reading comprehension, the latter ultimately being the raison d'être for reading. No such reading gains occurred in the same children

TABLE 3 | Pre versus post period of math intervention/null period.

	Mean (SD)			*F* (2,28)	
	Group 1	Group 2	Group 3	*F*-statistic	*p*-Value
Measures of reading: Woodcock–Johnson					
Word Identification (single real words)	1.9 (7.1)	−2.6 (6.3)	0.4 (3.5)	1.639	0.212
Word Attack (single pseudowords)	0.4 (5.5)	2.1 (7.3)	−2.8 (4.7)	1.659	0.209
Passage Comprehension (reading comprehension)	0.1 (8.9)	−1.3 (6.7)	2.1 (11.0)	0.374	0.691
Measures of skills that support reading					
Lindamood Auditory Conceptualization Test (phonemic awareness)	−0.9 (11.4)	−4.1 (10.0)	−1.4 (5.9)	0.354	0.705
Rapid Naming (naming fluency for letters/numbers)	−1.7 (4.8)	−2.3 (6.6)	1.6 (6.9)	1.075	0.355
Rapid Naming (naming fluency for colors/objects)	−0.4 (8.4)	−4.2 (6.3)	2.8 (9.6)	1.958	0.160
Digit Span (working memory)	2.0 (13.0)	−0.4 (13.0)	−5.0 (10.3)	0.787	0.465
Symbol Imagery (visual imagery/orthographic processing)	−3.7 (11.0)	−3.1 (9.5)	3.9 (6.9)	1.920	0.166

following their control period where some were engaged in a math intervention (active control) with others receiving no intervention at all (null period for developmental control). While one might have expected small carryover effects from the reading intervention into the control periods (due to the within-subject, cross-over design), we did not see gains in reading during the math intervention control period or the null period developmental control. Importantly, an ANOVA confirmed that gains in single word reading performance following the reading intervention was significantly different from any changes in single word reading following the math intervention. Both interventions required the participants' attention; and they both involved the tutor motivating the child to learn. As such, we can be assured that the reading gains can unequivocally be attributed to the information learned during the reading intervention and were not due to domain-general effects such as attention, or the result of a Hawthorn, or placebo effect. Lastly, these behavioral gains were independent of sex, demonstrating no sex-specific effects on intervention-induced gains in reading.

Overall, our results are similar to those in a prior study of adults with dyslexia (Eden et al., 2004), which used a similar tutoring approach and resulted in measurable gains in single real and pseudoword word reading (but not reading comprehension) as well as in the skills that were trained by the intervention (phonemic awareness and visual imagery). However, children with reading disability in a recent study (Christodoulou et al., 2017; Romeo et al., 2017) did not make gains on these (Word Identification subtest and Word Attack) or other measures of reading, even though the same intervention was used as the one employed here. This underscores the challenges of dyslexia and the fact that not all intervention studies result in a favorable outcome.

Turning to the gains in reading-related measures, it is no surprise that we found gains in those skills trained by the reading intervention, namely, phonemic awareness and visual imagery, demonstrating task-specific training. However, these gains were accompanied by gains in reading, indicating that these improvements in skills that support reading generalized and transferred to reading. Other skills known to support reading acquisition were studied: naming fluency and digit span, which together with phonemic awareness have been described as a set of interrelated phonological processing skills that are impaired in dyslexia due to a core phonological deficit (Wagner and Torgesen, 1987; Stanovich, 1988). As such our results shed light on the fact that these three skills are interrelated yet separate, with gains in one not necessarily accompanied by an equal gain in another. It is also possible that some of these other skills are not as pliable. For example, while some have advocated that working memory can be improved through training (Spencer-Smith and Klingberg, 2015), the strength and generalizability of these gains has been debated (Shipstead et al., 2012; Spencer-Smith and Klingberg, 2015; Nutley and Söderqvist, 2017). However, most likely changes in naming fluency and digit span did not occur because they were not targeted by the intervention. Nevertheless, it is interesting that gains can be made in reading without advancing these two skills.

None of these gains in reading, phonemic awareness or visual imagery were influenced by sex (male versus female). So, while there are behavioral (Wolf and Gow, 1986; Voyer et al., 1995; Weiss et al., 2003; Bornstein et al., 2004) and brain imaging studies (Clements et al., 2006; Sato, 2020; however, see Wallentin, 2009) showing sex-specific effects for language and spatial processing skills, in children with reading disability one sex does not have an advantage over the other when it comes to training the understanding of the sound structure of spoken language and how it maps to print, or the ability to visualize words in one's mind.

Benefits of the intervention appeared to have longevity, as the standard scores were maintained a year later (as a function of their raw scores increasing), indicating that the students' progress since the intervention ended was of a magnitude that is consistent with all children in their age group based on this normed testing instrument.

No Brain Activation Changes Following Reading Intervention

There has been much interest in brain changes following skill acquisition and training. It is known that brain function during object processing in adults who are literate is dramatically different from that of illiterate adults, speaking to the adaptations that occur in the brain as a consequence of learning to read (Dehaene et al., 2010). Successful treatment of dyslexia, it would

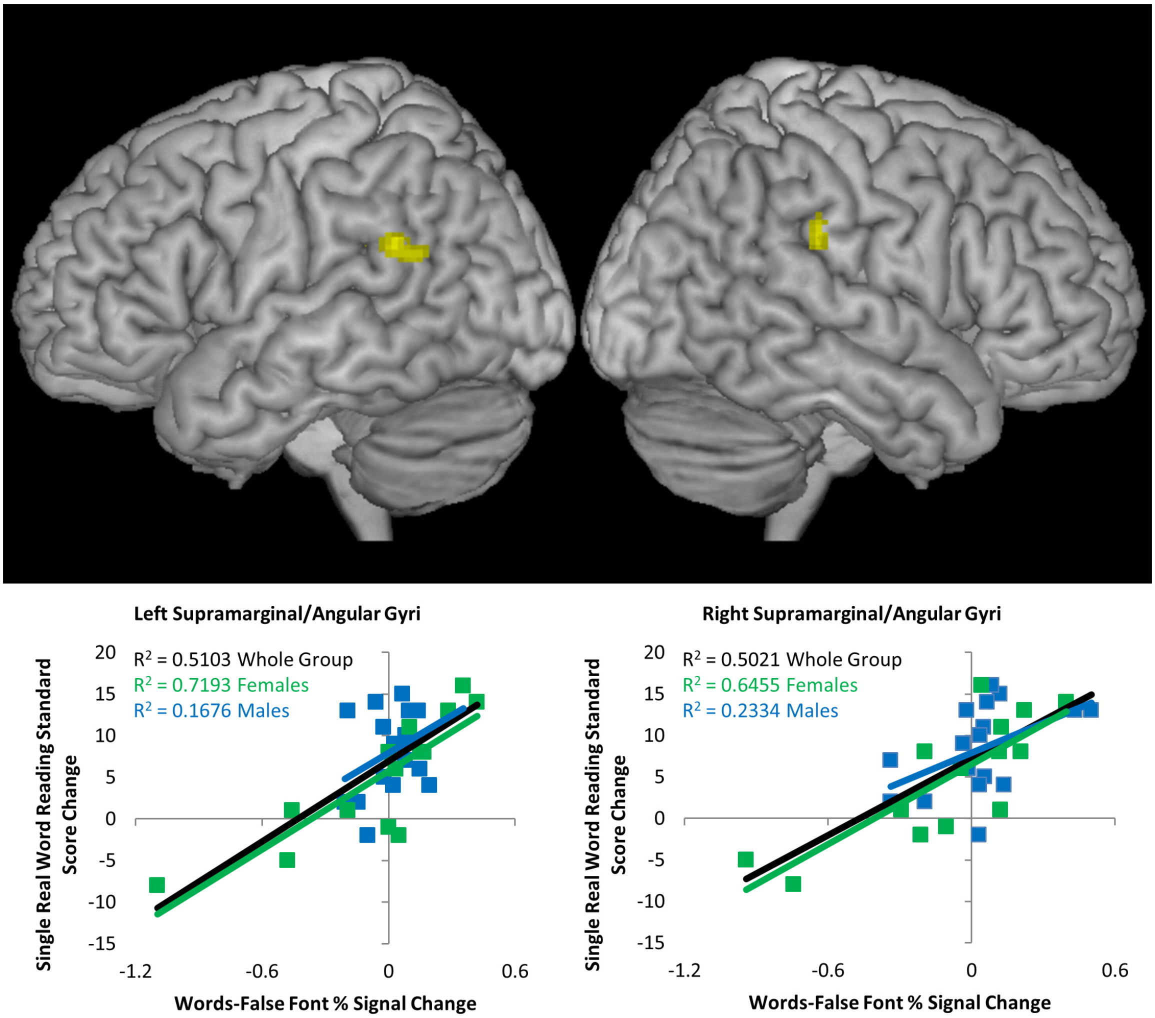


FIGURE 3 | Brain activation predictors of reading gain. Simple regression of Words > False Font activation prior to the intervention versus change in single real word reading (Word Identification) standard score following the reading intervention ($p < 0.001$ uncorrected height threshold, FWE corrected extent threshold $p < 0.05$). **Top:** lateral whole brain views of the whole group relationship between left and right supramarginal/angular gyrus activation and change in score. **Bottom:** signal extracted from each cluster with scatterplots for the whole group (black), females only (green), and males only (blue). R^2 values show strong relationships for both brain regions for females, while in males on the right hemisphere cluster was (barely) significant.

TABLE 4 | Coordinates and statistics for regression analysis.

Talairach peak coordinate			Cluster size (voxels)	*T*-statistic	*Z*-score	Anatomical location
X	***Y***	***Z***				
Activity during reading at Visit 1 versus change in single real word reading following reading intervention (whole group)						
Left hemisphere						
−46	−45	24	205	5.76	4.66	Supramarginal/angular gyri, BA 39/40
Right hemisphere						
46	−28	25	126	5.80	4.69	Supramarginal/angular gyri, BA 39/40

seem, should be accompanied by changes in brain function. Not only did we observe changes in brain function in adults with dyslexia undergoing a similar intervention (Eden et al., 2004), but prior studies in children have shown increases in activation during letter or word stimuli following reading interventions (Aylward et al., 2003; Temple et al., 2003; Shaywitz et al., 2004). However, as already noted in the Introduction, findings from individual intervention studies are variable. While one meta-analysis of brain imaging studies of reading interventions identified five regions of overlap, specifically in left thalamus, right insula/inferior frontal gyrus, left inferior frontal gyrus, right posterior cingulate, and left middle occipital gyrus, some foci in the meta-analysis were found in only two of the eight studies included. The authors themselves expressed caution in their interpretation of these results because of the variability in the methodologies used in these studies (Barquero et al., 2014). A more recent meta-analysis found no such convergence of results anywhere in the brain (Perdue et al., 2022) and pointed out significant limitations in the existing studies. It is worth noting that ours is the first study to use a within-subject control period to assess activation changes in dyslexia. We also applied a more stringent threshold than previous studies, recognizing that all older studies were accustomed to less stringent practices. Taking all of these factors into consideration our absence of a finding should perhaps not be all that surprising.

However, the question arises whether using a more targeted approach than the whole-brain analysis would have yielded a finding. As such we also conducted a *post hoc* region-of-interest (ROI) analysis. The details of this analysis were not described in section "Materials and Methods" because they followed later, but the approach is similar to that used in other studies (e.g., Brem et al., 2010) and in our prior work (Olulade et al., 2015), and is described in the **Supplementary Material** with results presented in **Supplementary Figure 1**. The results yielded no significant changes in specific regions of the fusiform gyrus (home of the visual word form area), even though this very region has shown to be altered by learning to read (Dehaene et al., 2010).

Our interpretation of these results is that there are several possible mechanisms at work, which cannot be differentiated in the current study and could also account for prior variability in the published results. One possibility is that even though gains were made in reading performance, the brain has remained unchanged, such that improved behavior occurs despite persisting functional aberrations. Based on prior, varied findings, we think it is more likely, however, that it has changed, but we are not able to measure these changes because they are too variable to be captured in a group analysis. Such variability would reflect the fact that reading intervention promotes functional changes, but that they occur in different brain regions for different individuals. A likely reason for this would be that if brain regions typically involved in reading do not change following the intervention and instead other brain areas compensate, this compensation may fall to different regions in different individuals. A subset of individuals may be mobilizing traditional reading networks, however, they represent enough of a minority that they are not captured in the group results. In addition to these mechanisms, there will always be some children who did not have brain changes because they did not make significant gains in reading. In this context, however, it is worth noting that a *post hoc* analysis correlating change in reading with change with activity did not yield any findings either. Considering that profiles of dyslexia can be unique, and prior studies on differences in brain anatomy and function in dyslexia have not entirely converged, it is not unreasonable to expect that changes following remediation could show similar variability. In fact, children who have struggled with reading will have received reading instruction or intervention of varying quality, and these prior experiences (which are very difficult to control for) will also be reflected in this variability. It is also possible that any of these sources of variability is reduced in adults, where lower plasticity in adulthood constrains intervention-induced changes to a limited set of brain regions, a possibility that merits further investigations by studying children and adults with dyslexia in the same study. It is also possible that heterogeneity amongst children can be reduced by identifying children at risk for dyslexia (based on a family history of dyslexia) an approach that is helpful in the quest to determine the etiology of dyslexia (Lohvansuu et al., 2021).

In Females, Brain Activity but Not Behavioral Measures, Predict Gains in Reading Following Intervention

Behavioral measures were unsuccessful at predicting later reading outcome, yet brain activity was. While prior behavioral studies, which demonstrated phonological skills to predict word level reading skills (Hatcher and Hulme, 1999; Torgesen et al., 1999; Catts et al., 2001) set expectations that we would find skills like single real and single pseudoword reading and phonemic awareness to be indicative of later reading gains, we did not. This aligns with a prior report by Hoeft et al. (2011), who found none of 17 reading measures to predict changes in single real word reading over 2.5 years, yet brain activation in right inferior frontal gyrus during a written word rhyming task predicted change in single word reading over the same time period. The current study found that activity in inferior parietal cortex in the left and right hemisphere predicted post-reading intervention gains in single word reading. However, closer examination showed that this effect was driven by the females in the group. Females had strong predictive relationships between activity during reading in the left and right supramarginal/angular gyri and later reading gains, while there was no result for males in the left hemisphere and the effect in the right supramarginal/angular gyrus barely meet significance.

Left temporoparietal cortex represents the indirect route in the dual-route model for reading and is thought to subserve phoneme-grapheme mapping (Pugh et al., 2000, 2001; Coltheart et al., 2001; Jobard et al., 2003). It is thought to be especially important during the early process of learning to read (when new words need to be "sounded out") and remains engaged into adulthood (Pugh et al., 2001; Turkeltaub et al., 2003; Sandak et al., 2004; Frost et al., 2009). In dyslexia however, this region is underactivated: The left inferior parietal region identified here maps precisely onto the location of less activity in those with dyslexia relative to typical readers identified by meta-analysis (Maisog et al., 2008). Greater engagement of left inferior parietal cortex while processing orthographic and phonological representations of words represent a sign of a brain that is ready to make greater gains in reading once targeted, structured and intense instructions are provided. Of the studies reviewed in the Introduction, one showed left inferior parietal lobule connectivity with middle temporal gyrus during a lexical decision task was predictive of basic reading skill gains following a reading intervention (Aboud et al., 2018). Our results are also consistent with two magnetoencephalography (MEG) studies showing activity in temporoparietal regions (and others) at baseline predicted reading fluency gains following intervention at a 1 year follow-up (Rezaie et al., 2011a,b). Specifically, signal in left and right middle and superior temporal gyri, left supramarginal and angular gyri, left ventral occipitotemporal regions, and right mesial temporal cortex were related to gains in reading fluency.

Why girls but not boys show a relationship where more engagement of left inferior parietal cortex during reading leads to reaping greater benefits from the intervention is not clear. Turning to the literature on the role of sex hormones on brain development, it has been shown that there is a negative correlation between fetal testosterone (*in utero*) and early childhood gray matter volume (8–11 years old) in right TPC, suggesting that the development of this region in males may be modulated by this sex hormone (Lombardo et al., 2012). This in turn may have an impact on brain function, possibly even in the contralateral hemisphere. Post-mortem studies in adults known to have had dyslexia during their lifetime revealed neuronal ectopias (attributed to developmental errors in neuronal migration), primarily in perisylvian regions (Galaburda et al., 1985) and primarily in males (Humphreys et al., 1990). It has been shown that estrogen treatment in women results in increased activation for verbal stimuli, and decreased activation for non-verbal stimuli in the left and right inferior parietal lobule during working memory tasks (Shaywitz et al., 1999). Based on these factors it has been suggested that males and females with dyslexia may have different etiological profiles due to their different hormonal environments (Krafnick and Evans, 2019).

However, it is important to note that the males and females did not differ in the gains they made following the reading intervention. While they may have reached these identical goals in different ways, it is not clear whether there are changes in brain activity following the intervention that are sex-dependent but if there are, we did not capture them. Yet in females, but not males, we were able to identify a left inferior parietal brain region that signals a level or readiness of brain function, promising that the introduction of an intervention will lead to a successful outcome. This suggests some separation of brain function in regions that bring about gains in reading, and regions that signal what may be a certain level of brain function that is required in order to harness the benefits of the intervention, but just in females. Future studies will need to disentangle the relationship between these and directly examine if this predictive relationship is under hormonal influence.

CONCLUSION

The results of this study suggest that while it is possible to see significant, specific and enduring gains in reading performance in children/adolescents with dyslexia following intensive treatment, individual variability may explain the fact that we did not observe any change in brain activity following the intervention. On the other hand, brain activity in left TPC predicted reading gains resulting from the intervention, while behavioral measures did not. Interestingly, the predictive powers of brain activity for reading outcome were attributed to the females but not males in our group, suggesting sexual dimorphism in the relationship between brain function during reading and the ability to reap benefits from intensive, structured reading intervention. As a whole, this work suggests there is considerable work to be done to understand brain changes related to reading intervention in order to determine what mechanisms are at work to promote these gains.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Georgetown University (IRB). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

GE and DLF conceived and designed the study. GE and EN were involved in overall study logistics and data collection. AK performed the statistical analyses and together with GE drafted the manuscript. All authors contributed to the manuscript and approved the submission.

FUNDING

This work was supported by the Eunice Kennedy Shriver Institute for Child Health and Human Development (P50 HD040095 and HD081078).

ACKNOWLEDGMENTS

We would like to thank Ashley Wall, Emma Cole, Corinna Moore, Jenni Rosenberg, Iain DeWitt, Robert Twomey, and Alison Merikangas for aiding with data acquisition. We also thank our participants and their families for volunteering their time, the Jemicy School for their support and for allowing us to conduct the intervention at their school and the staff from Lindamood-Bell Learning Processes for administering the intervention as determined by the research protocol.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2022.898661/full#supplementary-material

REFERENCES

Aboud, K. S., Barquero, L. A., and Cutting, L. E. (2018). Prefrontal mediation of the reading network predicts intervention response in dyslexia. *Cortex* 101, 96–106. doi: 10.1016/j.cortex.2018.01.009

Alexander, A. W., and Slinger-Constant, A. M. (2004). Current status of treatments for dyslexia: critical review. *J. Child Neurol.* 19, 744–758. doi: 10.1177/08830738040190100401

Altarelli, I., Monzalvo, K., Iannuzzi, S., Fluss, J., Billard, C., Ramus, F., et al. (2013). A functionally guided approach to the morphometry of occipitotemporal regions in developmental dyslexia: evidence for differential effects in boys and girls. *J. Neurosci.* 33, 11296–11301. doi: 10.1523/JNEUROSCI.5854-12.2013

Aylward, E. H., Richards, T. L., Berninger, V. W., Nagy, W. E., Field, K. M., Grimme, A. C., et al. (2003). Instructional treatment associated with changes in brain activation in children with dyslexia. *Neurology* 61, 212–219. doi: 10.1212/01.wnl.0000068363.05974.64

Bach, S., Richardson, U., Brandeis, D., Martin, E., and Brem, S. (2013). Print-specific multimodal brain activation in kindergarten improves prediction of reading skills in second grade. *Neuroimage* 82, 605–615. doi: 10.1016/j.neuroimage.2013.05.062

Badian, N. A. (1995). Predicting reading ability over the long term: the changing roles of letter naming, phonological awareness and orthographic processing. *Ann. Dyslexia* 45, 79–96. doi: 10.1007/BF02648213

Badian, N. A. (2001). Phonological and orthographic processing: their roles in reading prediction. *Ann. Dyslexia* 51, 179–202.

Baron, L. S., Hogan, T. P., Alt, M., Gray, S., Cabbage, K. L., Green, S., et al. (2018). Children with dyslexia benefit from orthographic facilitation during spoken word learning. *J. Speech Lang. Hear. Res.* 61, 2002–2014. doi: 10.1044/2018_JSLHR-L-17-0336

Barquero, L. A., Davis, N., and Cutting, L. E. (2014). Neuroimaging of reading intervention: a systematic review and activation likelihood estimate meta-analysis. *PLoS One* 9:e83668. doi: 10.1371/journal.pone.0083668

Bell, N. (1997). *Seeing Stars: Symbol Imagery for Phonemic Awareness, Sight Words, and Spelling*. San Luis Obispo, CA: Gander Publishing.

Bell, N., and Tuley, K. (1997). *On Cloud Nine: Visualizing and Verbalizing for Math*. San Luis Obispo, CA: Gander Publishing.

Bornstein, M. H., Hahn, C.-S., and Haynes, O. M. (2004). Specific and general language performance across early childhood: stability and gender considerations. *First Lang.* 24, 267–304

Bornstein, M. H., Haynes, O. M., Painter, K. M., and Jenevro, J. L. (2000). Child language with mother and with stranger at home and in the laboratory: a methodological study. *J. Child Lang.* 27, 407–420. doi: 10.1017/s0305000900004165

Brem, S., Bach, S., Kucian, K., Guttorm, T. K., Martin, E., Lyytinen, H., et al. (2010). Brain sensitivity to print emerges when children learn letter-speech sound correspondences. *PNAS* 107, 7939–7944. doi: 10.1073/pnas.0904402107

Brett, M., Valabregue, R., and Poline, J. (2002). "Region of interest analysis using an SPM toolbox," in *Proceedingds of the 8th International Conference on Functional Mapping of the Human Brain*, Sendai.

Burman, D. D., Bitan, T., and Booth, J. R. (2008). Sex differences in neural processing of language among children. *Neuropsychologia* 46, 1349–1362. doi: 10.1038/nmeth.2250.Digestion

Cahill, L. (2006). Why sex matters for neuroscience. *Nat. Rev. Neurosci.* 7, 477–484. doi: 10.1038/nrn1909

Cahill, L. (2012). A half-truth is a whole lie: on the necessity of investigating sex influences on the brain. *Endocrinology* 153, 2541–2543. doi: 10.1210/en.2011-2167

Catts, H. W., Fey, M. E., Zhang, X., and Tomblin, J. B. (2001). Estimating the risk of future reading difficulties in kindergarten children. *Lang. Speech Hear. Serv. Sch.* 32, 38–50. doi: 10.1044/0161-1461(2001/004)

Christodoulou, J. A., Cyr, A., Murtagh, J., Chang, P., Lin, J., Guarino, A. J., et al. (2017). Impact of intensive summer reading intervention for children with reading disabilities and difficulties in early elementary school. *J. Learn. Disabil.* 50, 115–127. doi: 10.1177/0022219415617163

Clark, K. A., Helland, T., Specht, K., Narr, K. L., Manis, F. R., Toga, A. W., et al. (2014). Neuroanatomical precursors of dyslexia identified from pre-reading through to age 11. *Brain* 137, 3136–3141. doi: 10.1093/brain/awu229

Clements, A. M., Rimrodt, S. L., Abel, J. R., Blankner, J. G., Mostofsky, S. H., Pekar, J. J., et al. (2006). Sex differences in cerebral laterality of language and visuospatial processing. *Brain Lang.* 98, 150–158. doi: 10.1016/j.bandl.2006.04.007

Coltheart, M., Rastle, K., Perry, C., Langdon, R., and Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychol. Rev.* 108, 204–256.

Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Nunes Filho, G., Jobert, A., et al. (2010). How learning to read changes the cortical networks for vision and language. *Science* 330, 1359–1364. doi: 10.1126/science.1194140

Denckla, M. B., and Rudel, R. G. (1976b). Rapid "automatized" naming (RAN): dyslexia differentiated from other learning disabilities. *Neuropsychologia* 14, 471–479.

Denckla, M. B., and Rudel, R. G. (1976a). Naming of object-drawlings by dyslexic and other learning disabled children. *Brain Lang.* 3, 1–15.

Dionne, G., Dale, P. S., Boivin, M., and Plomin, R. (2003). Genetic evidence for bidirectional effects of early lexical and grammatical development. *Child Dev.* 74, 394–412. doi: 10.1111/1467-8624.7402005

Eden, G. F., Jones, K. M., Cappell, K., Gareau, L., Wood, F. B., Zeffiro, T. A., et al. (2004). Neural changes following remediation in adult developmental dyslexia. *Neuron* 44, 411–422. doi: 10.1016/j.neuron.2004.10.019

Eden, G. F., Olulade, O. A., Evans, T. M., Krafnick, A. J., and Alkire, D. R. (2015). "Developmental Dyslexia," in *Neurobiology of Language*, eds G. Hickok and S. L. Small (Cambridge, MA: Academic Press), 815–826.

Evans, T. M., Flowers, D. L., Luetje, M. M., Napoliello, E. M., and Eden, G. F. (2016). Functional anatomy of arithmetic and word reading and its relationship to age. *Neuroimage* 143, 304–315. doi: 10.1016/j.neuroimage.2016.08.048

Evans, T. M., Flowers, D. L., Napoliello, E. M., and Eden, G. F. (2014). Sex-specific gray matter volume differences in females with developmental dyslexia. *Brain Struct. Funct.* 219, 1041–1054. doi: 10.1007/s00429-013-0552-4

Farris, E. A., Ring, J., Black, J., Lyon, G. R., and Odegard, T. N. (2016). Predicting growth in word level reading skills in children with developmental dyslexia using an object rhyming functional neuroimaging task. *Dev. Neuropsychol.* 41, 145–161. doi: 10.1080/87565641.2016.1158264

Frost, S. J., Landi, N., Mencl, W. E., Sandak, R., Fulbright, R. K., Tejada, E. T., et al. (2009). Phonological awareness predicts activation patterns for print and speech. *Ann. Dyslexia* 59, 78–97. doi: 10.1007/s11881-009-0024-y

Galaburda, A. M., Sherman, G. F., Rosen, G. D., Aboitiz, F., and Geschwind, N. (1985). Developmental dyslexia: four consecutive patients with cortical anomalies. *Ann. Neurol.* 18, 222–233. doi: 10.1002/ana.410180210

Gebauer, D., Fink, A., Kargl, R., Reishofer, G., Koschutnig, K., Purgstaller, C., et al. (2012). Differences in brain function and changes with intervention in children with poor spelling and reading abilities. *PLoS One* 7:e38201. doi: 10.1371/journal.pone.0038201

Hatcher, P. J., and Hulme, C. (1999). Phonemes, rhymes, and intelligence as predictors of children's responsiveness to remedial reading instruction: evidence from a longitudinal intervention study. *J. Exp. Child Psychol.* 72, 130–153. doi: 10.1006/jecp.1998.2480

Heim, S., Pape-Neumann, J., van Ermingen-Marbach, M., Brinkhaus, M., and Grande, M. (2015). Shared vs. specific brain activation changes in dyslexia after training of phonology, attention, or reading. *Brain Struct. Funct.* 220, 2191–2207. doi: 10.1007/s00429-014-0784-y

Hoeft, F., McCandliss, B. D., Black, J. M., Gantman, A., Zakerani, N., Hulme, C., et al. (2011). Neural systems predicting long-term outcome in dyslexia. *PNAS* 108, 361–366. doi: 10.1073/pnas.1008950108

Hoeft, F., Ueno, T., Reiss, A. L., Meyler, A., Whitfield-Gabrieli, S., Glover, G. H., et al. (2007). Prediction of children's reading skills using behavioral, functional, and structural neuroimaging measures. *Behav. Neurosci.* 121, 602–613. doi: 10.1037/0735-7044.121.3.602

Holm, S. (1979). Board of the foundation of the scandinavian journal of statistics a simple sequentially rejective multiple test procedure a simple sequentially rejective multiple test procedure. *Scand. J. Statist.* 6, 65–70.

Humphreys, P., Kaufmann, W. E., and Galaburda, A. M. (1990). Developmental dyslexia in women: neuropathological findings in three patients. *Ann. Neurol.* 28, 727–738. doi: 10.1002/ana.410280602

Jaeger, J. J., Lockwood, A. H., van Valin, R. D., Kemmerer, D. L., Murphy, B. W., and Wack, D. S. (1998). Sex differences in brain regions activated by grammatical and reading tasks. *NeuroReport* 9, 2803–2807. doi: 10.1097/00001756-199808240-00022

Jobard, G., Crivello, F., and Tzourio-Mazoyer, N. (2003). Evaluation of the dual route theory of reading: a metanalysis of 35 neuroimaging studies. *Neuroimage* 20, 693–712. doi: 10.1016/S1053-8119(03)00343-4

Kansaku, K., and Kitazawa, S. (2001). Imaging studies on sex differences in the lateralization of language. *Neurosci. Res.* 41, 333–337. doi: 10.1016/s0168-0102(01)00292-9

Katusic, S. K., Colligan, R. C., Barbaresi, W. J., Schaid, D. J., and Jacobsen, S. J. (2001). Incidence of reading disability in a population-based birth cohort, 1976-1982, Rochester, Minn. *Mayo Clin. Proc.* 76, 1081–1092. doi: 10.4065/76.11.1081

Kosslyn, S. M. (1976). Using imagery to retrieve semantic information : a developmental study. *Child Dev.* 47, 434–444.

Krafnick, A. J., Flowers, D. L., Napoliello, E. M., and Eden, G. F. (2011). Gray matter volume changes following reading intervention in dyslexic children. *Neuroimage* 57, 733–741.

Krafnick, A. J., and Evans, T. M. (2019). Neurobiological sex differences in developmental dyslexia. *Front. Psychol.* 9:2669. doi: 10.3389/fpsyg.2018.02669

Kyle, F., Kujala, J., Richardson, U., Lyytinen, H., and Goswami, U. (2013). Assessing the effectiveness of two theoretically motivated computerassisted reading interventions in the United Kingdom: GG Rime and GG Phoneme. *Read. Res. Q.* 48, 61–76. doi: 10.1002/rrq.038

Liederman, J., Kantrowitz, L., and Flannery, K. (2005). Male vulnerability to reading disability is not likely to be a myth: a call for new data. *J. Learn. Disabil.* 38, 109–129. doi: 10.1177/00222194050380020201

Lindamood, C., and Lindamood, P. (1971). *Lindamood Auditory Conceptualization (LAC) Test*. Austin, TX: Pro-Ed.

Linden, M., and Wittrock, M. C. (1981). The teaching of reading comprehension according to the model of generative learning. *Read. Res. Q.* 17, 44–57.

Lohvansuu, K., Torppa, M., Ahonen, T., Eklund, K., Hämäläinen, J. A., Leppänen, P. H. T., et al. (2021). Unveiling the mysteries of dyslexia—lessons learned from the prospective jyväskylä longitudinal study of dyslexia. *Brain Sci.* 11:427. doi: 10.3390/brainsci11040427

Lombardo, M. V., Ashwin, E., Auyeung, B., Chakrabarti, B., Taylor, K., Hackett, G., et al. (2012). Fetal testosterone influences sexually dimorphic gray matter in the human brain. *J. Neurosci.* 32, 674–680. doi: 10.1523/JNEUROSCI.4389-11.2012

Long, S. A., Winograd, P. N., and Bridge, C. A. (1989). The effects of reader and text characteristics on imagery reported during and after reading. *Read. Res. Q.* 24, 353–372.

Lyon, G. R., Shaywitz, S. E., and Shaywitz, B. A. (2003). Defining dyslexia, comorbidity, teachers' knowledge of language and reading. A definition of dyslexia. *Ann. Dyslexia* 53, 1–14.

Maisog, J. M., Einbinder, E. R., Flowers, D. L., Turkeltaub, P. E., and Eden, G. F. (2008). A meta-analysis of functional neuroimaging studies of dyslexia. *Ann. N. Y. Acad. Sci.* 1145, 237–259. doi: 10.1196/annals.1416.024

Martin, D. J., and Hoover, H. D. (1987). Sex differences in educational achievement: a longitudinal study. *J. Early Adolesc.* 7, 65–83. doi: 10.1177/0272431687071007

Meyler, A., Keller, T. A., Cherkassky, V. L., Gabrieli, J. D. E., and Just, M. A. (2008). Modifying the brain activation of poor readers during sentence comprehension with extended remedial instruction: a longitudinal study of neuroplasticity. *Neuropsychologia* 46, 2580–2592. doi: 10.1016/j.neuropsychologia.2008.03.012

Nutley, S. B., and Söderqvist, S. (2017). How is working memory training likely to influence academic performance? Current evidence and methodological considerations. *Front. Psychol.* 8:69. doi: 10.3389/fpsyg.2017.00069

Odegard, T. N., Ring, J., Smith, S., Biggan, J., and Black, J. (2008). Differentiating the neural response to intervention in children with developmental dyslexia. *Ann. Dyslexia* 58, 1–14. doi: 10.1007/s11881-008-0014-5

Olulade, O. A., Flowers, D. L., Napoliello, E. M., and Eden, G. F. (2013). Developmental differences for word processing in the ventral stream. *Brain Lang.* 125, 134–145. doi: 10.1016/j.bandl.2012.04.003

Olulade, O. A., Flowers, D. L., Napoliello, E. M., and Eden, G. F. (2015). Dyslexic children lack word selectivity gradients in occipito-temporal and inferior frontal cortex. *Neuroimage Clin.* 7, 742–754. doi: 10.1016/j.nicl.2015.02.013

Partanen, M., Siegel, L., and Giaschi, D. (2019). Effect of reading intervention and task difficulty on orthographic and phonological reading systems in the brain. *Neuropsychologia* 130, 13–25. doi: 10.1016/j.neuropsychologia.2018.07.018

Perdue, M. V., Mahaffy, K., Vlahcevic, K., Wolfman, E., Erbeli, F., Richlan, F., et al. (2022). Reading intervention and neuroplasticity: a systematic review and meta-analysis of brain changes associated with reading intervention. *Neurosci. Biobehav. Rev.* 132, 465–494. doi: 10.1016/j.neubiorev.2021.11.011

Peterson, R. L., and Pennington, B. F. (2012). Developmental dyslexia. *Lancet* 379, 1997–2007. doi: 10.1016/S0140-6736(12)60198-6

Price, C. J., Wise, R. J., and Frackowiak, R. S. (1996). Demonstrating the implicit processing of visually presented words and pseudowords. *Cereb. Cortex* 6, 62–70. doi: 10.1093/cercor/6.1.62

Pugh, K. R., Mencl, W. E., Jenner, A. R., Katz, L., Frost, S. J., Lee, J. R., et al. (2000). Functional neuroimaging studies of reading and reading disability(developmental dyslexia). *Ment. Retard. Dev. Disabil. Res. Rev.* 213, 207–213. doi: 10.1002/1098-2779(2000)6:3<207::AID-MRDD8>3.0.CO;2-P

Pugh, K. R., Mencl, W. E., Jenner, A. R., Katz, L., Frost, S. J., Lee, J. R., et al. (2001). Neurobiological studies of reading and reading disability. *J. Commun. Disord.* 34, 479–492. doi: 10.1016/s0021-9924(01)00060-0

Quinn, J. M., and Wagner, R. K. (2015). Gender differences in reading impairment and in the identification of impaired readers: results from a large-scale study of at-risk readers. *J. Learn. Disabil.* 48, 433–445. doi: 10.1177/0022219413508323

Ramus, F., Altarelli, I., Jednoróg, K., Zhao, J., Scotto, and di Covella, L. (2018). Neuroanatomy of developmental dyslexia: pitfalls and promise. *Neurosci. Biobehav. Rev.* 84, 434–452. doi: 10.1016/j.neubiorev.2017.08.001

Rezaie, R., Simos, P. G., Fletcher, J. M., Cirino, P. T., Vaughn, S., and Papanicolaou, A. C. (2011b). Temporo-parietal brain activity as a longitudinal predictor of response to educational interventions among middle school struggling readers. *J. Int. Neuropsychol. Soc.* 17, 875–885. doi: 10.1017/S1355617711000890

Rezaie, R., Simos, P. G., Fletcher, J. M., Cirino, P. T., Vaughn, S., and Papanicolaou, A. C. (2011a). Engagement of temporal lobe regions predicts response to educational interventions in adolescent struggling readers. *Dev. Neuropsychol.* 36, 869–888. doi: 10.1080/87565641.2011.606404

Richards, T. L., Aylward, E. H., Berninger, V. W., Field, K. M., Grimme, A. C., Richards, A. L., et al. (2006). Individual fMRI activation in orthographic mapping and morpheme mapping after orthographic or morphological spelling treatment in child dyslexics. *J. Neurolinguist.* 19, 56–86. doi: 10.1016/j.jneuroling.2005.07.003

Richlan, F. (2012). Developmental dyslexia: dysfunction of a left hemisphere reading network. *Front. Hum. Neurosci.* 6:120. doi: 10.3389/fnhum.2012.00120

Richlan, F., Kronbichler, M., and Wimmer, H. (2011). Meta-analyzing brain dysfunctions in dyslexic children and adults. *Neuroimage* 56, 1735–1742. doi: 10.1016/j.neuroimage.2011.02.040

Romeo, R. R., Christodoulou, J. A., Halverson, K. K., Murtagh, J., Cyr, A. B., Schimmel, C., et al. (2017). Socioeconomic status and reading disability: neuroanatomy and plasticity in response to intervention. *Cereb. Cortex* 28, 2297–2312. doi: 10.1093/cercor/bhx131

Rutter, M., Caspi, A., Fergusson, D., Horwood, L. J., Goodman, R., Maughan, B., et al. (2004). Sex differences in developmental reading disability: new findings from 4 epidemiological studies. *J. Am. Med. Assoc.* 291, 2007–2012. doi: 10.1001/jama.291.16.2007

Sadoski, M. (1983). An exploratory study of the relationships between reported imagery and the comprehension and recall of a story. *Read. Res. Q.* 19, 110–123.

Sandak, R., Mencl, W. E., Frost, S. J., and Pugh, K. R. (2004). The neurobiological basis of skilled and impaired reading: recent findings and new directions. *Sci. Stud. Read.* 8, 273–292.

Sato, M. (2020). The neurobiology of sex differences during language processing in healthy adults: a systematic review and a meta-analysis. *Neuropsychologia* 140:107404 doi: 10.1016/j.neuropsychologia.2020.107404

Scarborough, H. S., and Brady, S. A. (2002). Toward a common terminology for talking about speech and reading: a glossary of the "phon" words and some related terms. *J. Lit. Res.* 34, 299–336. doi: 10.1207/s15548430jlr3403

Shaywitz, B. A., Shaywitz, S. E., Blachman, B. A., Pugh, K. R., Fulbright, R. K., Skudlarski, P., et al. (2004). Development of left occipitotemporal systems for skilled reading in children after a phonologically- based intervention. *Biol. Psychiatry* 55, 926–933. doi: 10.1016/j.biopsych.2003.12.019

Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Constable, R. T., Skudlarski, P., Fulbright, R. K., et al. (1995). Sex differences in the functional organization of the brain for language. *Nature* 373, 607–609.

Shaywitz, S. E., Shaywitz, B. A., Pugh, K. R., Fulbright, R. K., Skudlarski, P., Mencl, W. E., et al. (1999). Effect of estrogen on brain activation patterns in postmenopausal women during working memory tasks. *JAMA* 281, 1197–1202. doi: 10.1001/jama.281.13.1197

Shipstead, Z., Hicks, K. L., and Engle, R. W. (2012). Cogmed working memory training: Does the evidence support the claims? *J. Appl. Res. Mem. Cogn.* 1, 185–193. doi: 10.1016/j.jarmac.2012.06.003

Spencer-Smith, M., and Klingberg, T. (2015). Benefits of a working memory training program for inattention in daily life: a systematic review and meta-analysis. *PLoS One* 10:e0119522. doi: 10.1371/journal.pone.0119522

Stanovich, K. E. (1988). Explaining the differences between the dyslexic and the garden-variety poor reader: the phonological-core variable-difference model. *J. Learn. Disabil.* 21, 590–604. doi: 10.1177/002221948802101003

Temple, E., Deutsch, G. K., Poldrack, R. A., Miller, S. L., Tallal, P., Merzenich, M. M., et al. (2003). Neural deficits in children with dyslexia ameliorated by behavioral remediation: evidence from functional MRI. *PNAS* 100, 2860–2865. doi: 10.1073/pnas.0030098100

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Rose, E., Lindamood, P., Conway, C., et al. (1999). Preventing reading failure in young children with phonological processing disabilities: group and individual responses to instruction. *J. Educ. Psychol.* 91, 589–593.

Turkeltaub, P. E., Flowers, D. L., Verbalis, A., Miranda, M., Gareau, L., and Eden, G. F. (2004). The neural basis of hyperlexic reading: an fMRI case study. *Neuron* 41, 11–25. doi: 10.1016/s0896-6273(03)00803-1

Turkeltaub, P. E., Gareau, L., Flowers, D. L., Zeffiro, T. A., and Eden, G. F. (2003). Development of neural mechanisms for reading. *Nat. Neurosci.* 6, 767–773. doi: 10.1038/nn1065

Voyer, D., Voyer, S., Bryden, M. P., thank Ann Bigelow, W., Brooks, G. P., Bulman-Fleming, B., et al. (1995). Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. *Psychol. Bull.* 117, 250–270. doi: 10.1037/0033-2909.117.2.250

Wagner, R. K., and Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychol. Bull.* 101, 192–212. doi: 10.1037//0033-2909.101.2.192

Wallentin, M. (2009). Putative sex differences in verbal abilities and language cortex: a critical review. *Brain Lang.* 108, 175–183.

Wanzek, J., and Vaughn, S. (2007). Research-based implications from extensive early reading interventions. *Sch. Psychol. Rev.* 36, 541–561.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio: Psychological Corporation.

Weiss, E. M., Kemmler, G., Deisenhammer, E. A., Fleischhacker, W. W., and Delazer, M. (2003). Sex differences in cognitive functions. *Pers. Individ. Differ.* 35, 863–875.

Wolf, M., and Gow, D. (1986). A longitudinal investigation of gender differences in language and reading development. *First Lang.* 6, 81–110. doi: 10.1177/014272378600601701

Woodcock, R. W., McGrew, K., and Mather, N. (2001). *The Woodcock-Johnson Tests of Achievement*, 3rd Edn. Itasca, IL: Riverside.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Neuroscience

MINI REVIEW
published: 16 June 2022
doi: 10.3389/fnins.2022.921931

The Functional Neuroanatomy of Reading Intervention

Jeremias Braid† *and Fabio Richlan**†

Department of Psychology, Centre for Cognitive Neuroscience, Paris Lodron University of Salzburg, Salzburg, Austria

The present article reviews the literature on the brain mechanisms underlying reading improvements following behavioral intervention for reading disability. This includes evidence of neuroplasticity concerning functional brain activation, brain structure, and brain connectivity related to reading intervention. Consequently, the functional neuroanatomy of reading intervention is compared to the existing literature on neurocognitive models and brain abnormalities associated with reading disability. A particular focus is on the left hemisphere reading network including left occipito-temporal, temporo-parietal, and inferior frontal language regions. In addition, potential normalization/compensation mechanisms involving right hemisphere cortical regions, as well as bilateral sub-cortical and cerebellar regions are taken into account. The comparison of the brain systems associated with reading intervention and the brain systems associated with reading disability enhances our understanding of the neurobiological basis of typical and atypical reading development. All in all, however, there is a lack of sufficient evidence regarding rehabilitative brain mechanisms in reading disability, which we discuss in this review.

Keywords: reading intervention, reading disability, developmental dyslexia, neuroimaging, review

Edited by:
Johannes Ziegler,
Aix-Marseille Université, France

Reviewed by:
Eraldo Paulesu,
University of Milano-Bicocca, Italy
Johanna Liebig,
Freie Universität Berlin, Germany

***Correspondence:**
Fabio Richlan
Fabio.Richlan@plus.ac.at

†*These authors have contributed equally to this work*

Specialty section:
This article was submitted to Neurodevelopment,
a section of the journal
Frontiers in Neuroscience

Received: *16 April 2022*
Accepted: *31 May 2022*
Published: *16 June 2022*

Citation:
Braid J and Richlan F (2022) The Functional Neuroanatomy of Reading Intervention.
Front. Neurosci. 16:921931.
doi: 10.3389/fnins.2022.921931

INTRODUCTION

Reading acquisition or learning to read is a complex endeavor requiring the integration of orthographic, phonological, and semantic information about written words together with knowledge of spoken language and conceptual knowledge. In a considerable number of cases, however, children struggle with the acquisition of foundational reading skills—a condition known as reading disability (RD) or developmental dyslexia. Specifically, RD is characterized by severe and persistent problems in reading acquisition.

In children with RD, performance in standardized reading tests is significantly below the age-expected norm. In addition, people affected by RD often present a mixture of different manifestations of problems in diverse aspects of literacy including reading fluency, accuracy, comprehension, and/or spelling (e.g., Lyon et al., 2003). Importantly, the difficulties cannot be explained by problems regarding intelligence, motivation, vision, or educational environment. Finally, these difficulties markedly impair academic achievement or activities in everyday life requiring reading skills (American Psychiatric Association, 2013; World Health Organization, 2016).

The present mini-review aims to concisely summarize the literature on neuroplasticity following reading intervention and to relate it to the functional neuroanatomical models of reading and RD. For that purpose, we review the systematic findings regarding brain mechanisms underlying reading improvements following behavioral intervention for RD (covering multiple rehabilitation techniques). This includes evidence of neuroplasticity concerning functional brain activation, brain structure, and brain connectivity. Finally, we discuss limitations, open issues, and future perspectives in order to pave the way for further progress in this field.

THE FUNCTIONAL NEUROANATOMY OF READING AND READING DISABILITY

Functional Brain Activation

During the last years, there has been considerable progress in understanding the neurocognitive and neurobiological mechanisms underlying reading and RD. Using brain imaging techniques such as functional magnetic resonance imaging (fMRI), electroencephalography (EEG), and magnetoencephalography (MEG), studies have largely converged on the brain circuits involved in typical and atypical reading. Specifically, the functional neuroanatomical model of typical reading involves a predominantly left-lateralized network including occipito-temporal (OT), temporo-parietal (TP), and frontal language regions (e.g., Dolan et al., 1997; Paulesu et al., 2000; Cattinelli et al., 2013; Martin et al., 2015; Schuster et al., 2016; Chyl et al., 2021).

With respect to RD, qualitative reviews and quantitative meta-analyses have identified altered brain activation in atypical readers during reading or reading-related tasks in this left-hemisphere network. In particular, the most consistent finding across studies was underactivation in people affected by RD compared with their age-matched peers in the left ventral OT cortex (fusiform gyrus, FFG and posterior inferior temporal gyrus, ITG), the left posterior middle and superior temporal gyrus (MTG and STG), and the left inferior frontal gyrus (IFG) (e.g., Paulesu et al., 2001, 2014; Maisog et al., 2008; Richlan et al., 2009, 2011; Martin et al., 2016).

Underactivation of the left hemisphere reading network—in particular the language-universal dysfunction of the left ventral OT cortex—most probably reflects the phonological speed deficit characteristic of RD. This is in line with evidence showing that in typical readers the ventral OT cortex subserves both lexical whole-word recognition and sublexical serial decoding (e.g., Richlan et al., 2010; Schurz et al., 2010; Wimmer et al., 2010). Conversely, overactivation in atypical compared with typical readers was identified in the left precentral cortex and the bilateral frontal striatum (including caudate and putamen), perhaps reflecting overreliance on sub-vocal articulatory-based reading processes (Richlan, 2012, 2014, 2020; Hancock et al., 2017).

There is an increasing number of hints on the existence of additional functional activation abnormalities in cortical, sub-cortical, and cerebellar regions in RD (e.g., Danelli et al., 2012; Mascheretti et al., 2017; Alvarez and Fiez, 2018; De Vos et al., 2020), but this has not yet been evidenced by objective quantification through systematic meta-analysis. The reasons for this absence most probably lie more in methodological limitations of the meta-analyses themselves, than in the primary studies. Obviously, any alterations in functional brain activation strongly depend on the in-scanner tasks and baseline conditions, as well as several other experimental considerations related to stimulus types, presentation modalities, instructions, sample sizes, analytical techniques, statistical thresholds and last but not least diagnosis/inclusion criteria for the RD groups (see section "Limitations, Open Issues, and Future Perspectives").

Gray and White Matter Structure and Connectivity

Quantitative meta-analyses on gray matter (GM) structural abnormalities in RD as investigated by means of voxel-based morphometry showed a similar picture, with limited convergence across studies (for an in-depth discussion see Ramus et al., 2018). The most robust and consistent finding was GM volume reduction in atypical compared with typical readers in the right STG and the left superior temporal sulcus (STS), but only about half of the primary studies contributed to these meta-analytic clusters (Linkersdörfer et al., 2012; Richlan et al., 2013; Eckert et al., 2016).

Across different languages, the left STS is assumed to play an important role in the integration of auditory and visual information (e.g., Van Atteveldt et al., 2004; Blomert, 2011; Holloway et al., 2013; Richlan, 2019). Therefore, in typical reading acquisition, it plays a pivotal role during self-reliant learning processes based on serial grapheme-phoneme conversion. The STG/STS GM volume reduction found in RD might be related to a deficit in this sublexical self-teaching reading strategy, specifically in the development of a brain system for efficient interactive processing of auditory and visual linguistic inputs (Blau et al., 2010).

With respect to white matter (WM) structure and connectivity, the major pathways supporting skilled reading are found in left TP areas and in posterior callosal tracts including the superior longitudinal fasciculus (including the arcuate fasciculus, AF), occipital and temporal callosal fibers, and corona radiata fibers passing through the posterior limb of the internal capsule (Ben-Shachar et al., 2007). In RD, these pathways have been identified with lower fractional anisotropy values (indicating reduced structural integrity) in diffusion tensor imaging (DTI) studies. A prime candidate fiber tract most consistently associated with RD is the left AF, which connects left TP and left frontal language regions (Silani et al., 2005; Vandermosten et al., 2012; Dehaene et al., 2015). Additional findings point to deficits in visual thalamo-cortical connections (Müller-Axt et al., 2017).

The left AF was reported to be among the first brain circuits to anatomically change during reading acquisition. Specifically, learning to read has been shown to be accompanied by an increase in fractional anisotropy (FA) and a decrease in perpendicular diffusivity (PD) (reflecting a microstructural improvement) of this fiber tract (Thiebaut de Schotten et al., 2012; Yeatman et al., 2012). Based on these findings, the left AF is assumed to play a crucial role, especially during the early stages of literacy development by supporting letter-speech sound integration and grapheme-phoneme coding, which, in turn, is required for self-reliant phonological word decoding in beginning readers (Richlan, 2019).

Functional Connectivity and Integration

Besides structural connectivity by means of DTI, studies on functional and effective connectivity provide interesting insights into how brain regions interact with each other in order to support skilled reading (e.g., Mechelli et al., 2005; Schlaggar and McCandliss, 2007; Vogel et al., 2013; Carreiras et al., 2014; Schurz et al., 2014). Put simply, in typical readers, left OT, TP, and IFG regions are functionally connected, whereas in RD this functional coupling is disrupted, either as a cause or consequence (or both) of reading difficulties. Reduced functional connectivity in RD within the typical left-hemisphere reading network was found both during reading and reading-related tasks (e.g., Paulesu et al., 1996; Van der Mark et al., 2011; Boets et al., 2013; Olulade et al., 2015; Cao et al., 2017; Morken et al., 2017) as well as in the absence of a task, that is, during rest (e.g., Koyama et al., 2013; Schurz et al., 2015).

The idea that RD results from disrupted connections between brain regions supporting vision and brain regions supporting language has been around for decades (Geschwind, 1965a,b; Paulesu et al., 1996). As evidenced by modern-day neuroimaging, this disruption of brain systems might reflect the characteristic visual-verbal speed deficit in the behavioral manifestation of RD, which, in turn, is attributed to inefficient access from letters to speech sounds. This deficit in RD was hypothesized to underlie the universal reading speed impairment across languages (Wimmer, 1993; Ramus and Szenkovits, 2008; Blomert, 2011; Richlan, 2019). As pointed out in the following section, behavioral interventions for people affected by RD often focus on letter-speech sound integration and on linking sub-lexical and lexical orthographic and phonological information (e.g., Fraga González et al., 2015).

READING INTERVENTION AND NEUROPLASTICITY

Behavioral Effects of Reading Intervention

RD poses a significant burden for those affected (American Psychiatric Association, 2013; World Health Organization, 2016). Fortunately, many studies have shown that reading intervention can be beneficial for people with RD (e.g., Wanzek et al., 2018). By and large, explicit phonics instruction can be regarded as the gold standard in reading intervention programs due to its beneficial effects on a large amount of RD people (Galuschka et al., 2014). This includes interventions aimed at teaching (a) letter-speech sound correspondences, (b) decoding strategies that involve blending or segmenting individual letters or phonemes, and (c) dividing spoken or written words into syllables or onsets and rimes.

Systematic meta-analyses revealed moderate effect sizes regarding improvement in reading ability after reading intervention (Wanzek et al., 2013, 2016, 2018). The examined intervention programs, however, differed significantly in a number of aspects such as skills targeted, duration, intensity, modality, and group size. In addition, marked individual differences between participants within particular studies impede generalization. Therefore, specific conclusions on the efficacy of intervention programs must be drawn with caution.

Brain Effects of Reading Intervention

Functional Brain Activation

Recently, Perdue et al. (2022) conducted a quantitative meta-analysis using seed-based d mapping (Albajes-Eizagirre et al., 2019) on changes in brain activation pre/post reading intervention in people with—or at risk for—RD. In sum, eight fMRI studies that met predefined inclusion criteria (total aggregated sample size = 151 participants, mean age per study = 5.6–44 years) were included in the meta-analysis, which followed the strict PRISMA statement for transparent reporting of systematic reviews and meta-analyses evaluating the effects of interventions (Page et al., 2021). Intervention duration lasted from three to twelve weeks and various (in part commercially available) training programs were used, aimed at different reading component skills (e.g., phoneme awareness, morpheme-based spelling, grapheme-phoneme conversion, or reading fluency).

No statistically significant brain effects of reading intervention could be observed in this meta-analysis. According to the authors, one possible explanation could be the small set of included studies due to the exclusion of studies for methodological reasons. Additionally, even the studies that met the inclusion criteria suffered from small sample sizes. The primary limiting factor, however, is the use of region/volume of interest (ROI/VOI) analysis instead of whole-brain analysis, which renders objective coordinate-based meta-analysis difficult if not impossible and therefore has been a methodological exclusion criterion. Discussing their findings, Perdue et al. (2022) suggest, that future reading intervention studies should employ exploratory, spatially unrestricted whole-brain analysis in larger samples to adequately assess the effects of reading intervention on brain activation.

Furthermore, Barquero et al. (2014) reported an activation likelihood estimation (ALE) meta-analysis with a slightly different set of eight fMRI studies (seven studies with children and adolescents and one study with adults, total aggregated sample size = 173 participants) assessing functional activation patterns after reading intervention. Across the included studies, intervention periods ranged from 3 weeks up to two school years. As in the Perdue et al. (2022) meta-analysis, various different intervention programs were administered in the single studies.

Increased activation in RD participants following reading intervention was observed in the following brain regions of the typical reading network: left thalamus, right insula/IFG, left IFG, right posterior cingulate gyrus, and left middle occipital gyrus. In conclusion, and similar to the previously discussed meta-analysis, the authors note that the results must be interpreted with caution due to several methodological limitations at this relatively early stage of research, such as the high degree of heterogeneity in data acquisition and analysis methodology across studies and the generally limited number of published studies.

Despite the slightly disappointing and inconclusive meta-analytic (null-) results, a systematic qualitative review of reading intervention studies—also including MEG studies, which could not be part of the coordinate-based meta-analysis due to methodological reasons—essentially showed the following findings: fMRI and MEG studies identified pre-to-post changes in (a) the typical reading network as detailed in section "The Functional Neuroanatomy of Reading and Reading Disability", thus indicating normalization of functional activation in RD and (b) additional cortical, sub-cortical, and cerebellar regions usually not included in this network, probably associated with compensatory reading mechanisms (Perdue et al., 2022).

In particular, multiple studies reported elevated levels of activation following reading intervention in the left hemisphere reading areas (Shaywitz et al., 2004; Richards et al., 2006b; Horowitz-Kraus et al., 2014; Heim et al., 2015). Importantly, this indicates that—through specific training—functions of the typical reading network can recover in people with RD. Additionally, initial group differences in activation levels between RD and typically developing controls were normalized in some studies, that is, differences before intervention were no longer detectable after intervention (Aylward et al., 2003; Richards et al., 2006a; Meyler et al., 2008).

In some of these studies, this normalization of functional activation also involved increases in the right hemisphere and sub-cortical regions (e.g., Meyler et al., 2008; Gebauer et al., 2012; Nugiel et al., 2019; Partanen et al., 2019). Equal levels of right-hemispheric activation in RD following reading intervention—compared with typical readers—could indicate a shift toward the typical engagement of these regions. Previous literature instead largely suggested that such changes may reflect compensatory processes, in the sense that people with RD engage regions outside of the typical reading network in order to make up for their deficits.

Across studies, the most consistent normalization effects could be observed in the right IFG (Temple et al., 2003; Meyler et al., 2008; Odegard et al., 2008; Horowitz-Kraus et al., 2014; Partanen et al., 2019). The right IFG is already activated during reading and reading-related processes before intervention and people with higher initial activation showed greater engagement after intervention (Hoeft et al., 2011). Functionally, the right IFG is thought to support articulatory recoding, working memory, and attention during reading (Shaywitz et al., 2002; Hancock et al., 2017).

Neuroplasticity associated with reading intervention in the right hemisphere was also identified in homologous regions of the left hemisphere reading network, that is, STG, OT cortex, and inferior parietal lobule (IPL) (Perdue et al., 2022). The exact functional role of greater activation following reading intervention in these right hemisphere sub-components, however, still remains unclear (for an in-depth discussion on this topic see Perdue et al., 2022). To conclude, contrary to previous findings, newer studies suggest that enhanced right-hemispheric activation in RD following reading intervention might reflect normalization rather than compensation.

Gray and White Matter Structure and Connectivity

With respect to GM volume, structural changes related to reading intervention in children with RD were identified in hubs of the typical reading network, sub-cortical and right hemisphere regions. This included increases in GM volume relative to the pre-intervention assessments in the left anterior OT cortex extending into the hippocampus, bilateral precuneus, right hippocampus, and right cerebellum (Krafnick et al., 2011). After an 8-week period without intervention, these effects were stable and an additional cluster of GM volume increase was identified in the right caudate.

Romeo et al. (2017) investigated neuroplasticity by means of cortical thickness in 65 children with RD (aged 6–9 years) of which n = 40 participated in a summer reading intervention program, that lasted for 6 weeks. The remaining n = 25 children constituted the waiting-list control group. A commercial multisensory program (centered on orthographic and phonological processing) was used. Results showed that the intervention group maintained their reading scores, whereas the waiting control group decreased in performance. On an individual basis, children who improved their reading scores—in the intervention group—had lower socioeconomic backgrounds than children that declined in reading performance. Comparing responders with non-responders, greater change in cortical thickness could be observed in responders in the following regions: bilateral middle-inferior temporal cortex, IPL, precentral cortex, and paracentral/posterior cingulate cortex, right STG and insula, and left MTG.

Regarding WM structure and connectivity, neuroplasticity associated with reading intervention could be observed in a number of studies (Perdue et al., 2022). Specifically, several DTI studies reported changes in structural connectivity and WM integrity linked to enhancement of reading performance after intervention (Davis et al., 2010; Richards et al., 2017; Huber et al., 2018). Increased FA and decreased mean and radial diffusivity might indicate that WM pathways increased in efficiency by improving communication among distant cortical and sub-cortical structures involved in reading.

Impressively, structural changes already occurred after only 2–3 weeks of intervention (same program as in Romeo et al., 2017) when children aged 7–12 years were scanned multiple times over the period of 8 weeks (Huber et al., 2018). Therefore, the brain delivers a fast adaptation response following the high demands of intensive training, i.e., detectable neuroanatomical rewiring processes as a consequence of reading intervention. The links between reading skill improvement and WM microstructure deviated from typical developmental trajectories during intervention. Consequently, this does not support the assumption of neuroanatomical normalization as reported in some functional activation studies. Study designs similar to the one employed by Huber et al. (2018), however, are costly and therefore rarely used, even though they provide important insights into the temporal progress of ongoing brain changes.

Davis et al. (2010) reported that changes in structural connectivity in response to a small group reading intervention (duration = 17 weeks) in eleven first graders (mean age = 7.5 years) were consistent with behavioral changes. Moreover, associations between functional connectivity and WM structure (Richards et al., 2018), together with incremental changes in WM microstructure during reading intervention as described before provide valuable insights into possible mechanisms of neuroplasticity in brain networks that enhance reading. Essentially, neural optimization in terms of rewiring of network connections might be related to the establishment of stronger brain circuits on the one hand and the reduction of inefficient connections on the other hand. Therefore, the strict distinction between neuroanatomical normalization vs. compensation mechanisms may not apply in these studies.

Functional Connectivity and Integration

Evidence regarding altered functional connectivity following reading intervention suggests that integrating dispersed functional networks facilitates reading improvements in RD (Perdue et al., 2022). Intervention-related neuroplasticity effects were found both during task-based and resting-state fMRI in diverse brain systems including fronto-parietal and cingulo-opercular networks (Horowitz-Kraus et al., 2015; Richards et al., 2016, 2017), and among low-level visual, dorsal attentional, and executive function networks distributed in various brain regions (Horowitz-Kraus et al., 2019).

Specifically, Horowitz-Kraus et al. (2019) examined changes in functional connectivity during task-based fMRI using a lexical decision task. They compared three groups (n = 18 each): RD, comorbid attention-deficit and hyperactivity disorder and RD (ADHD + RD), and typically developing (TD) in a computer-based intervention program targeting reading skills and executive functions, which lasted for 4 weeks. Independent component analysis was used to extract networks for connectivity analysis. Across the three groups, results showed positive correlations between reading speed gains and both increased functional network connectivity between the executive function component (bilateral superior frontal gyri) and the low-level visual component (bilateral FFG) and increased functional connectivity between the dorsal attention component (bilateral precuneus/posterior cingulate) and the low-level visual component.

In contrast, Richards et al. (2018) also found decreases in local functional connectivity following a computerized program focused on reading and writing (duration = 18 lessons). The sample consisted of N = 42 students (mean age = 11 years, 10 months). For example, during a multi-sentence reading comprehension task, local functional connectivity in the right middle frontal gyrus decreased in two RD groups, whereas it increased in a dysgraphia and a TD group. The above-reported findings were interpreted as reflecting modulation of attention-linked networks during reading. Since both increases and decreases in functional connectivity were observed, one could argue that this pattern reflects a process of re-adjustment toward an optimal level of integration and separation within and between different functional brain networks.

LIMITATIONS, OPEN ISSUES, AND FUTURE PERSPECTIVES

Studies on the brain mechanisms underlying reading improvements following behavioral intervention for RD have provided tremendously valuable insights into the neurobiology of typical and atypical reading development. Taken together, however, there is only limited consistency across studies regarding possible neuroplasticity effects, as illustrated by the absence of (or only weak) meta-analytic evidence (Barquero et al., 2014; Perdue et al., 2022). Reasons for this heterogeneity of results are discussed below.

Meta-analyses are generally limited in scope due to strict inclusion/exclusion criteria. This is particularly evident in meta-analyses of brain effects. Specifically, in the recent meta-analysis by Perdue et al. (2022), 31 out of 39 thematically relevant primary studies had to be excluded because of ineligible imaging modalities, regionally restricted analysis strategies, imaging time points, and other methodological considerations. In addition, even the included studies used a variety of different fMRI activation tasks and methodological parameters for image preprocessing and statistical analysis, and generally suffered from small sample sizes, thus increasing the probability of both false positive and false negative results (Button et al., 2013).

The next issue concerns the participants in these studies themselves. In the reviewed studies, participants differed in terms of several aspects known to have an influence on reading development, such as age, home literacy environment, socioeconomic status, and initial skills. With respect to age, Suggate (2010) reported an interaction between grade at intervention and focus of intervention. In earlier grades, greater effects were elicited by phonics training, whereas in later grades, greater effects were elicited by comprehension training. Orthographic depth of the written language may also play a considerable role in this regard (e.g., Paulesu et al., 2001; Richlan, 2014, 2020; Martin et al., 2016).

In addition, there is the potential problem of (mis-) diagnosis and comorbidity. In particular, RD is often comorbid with atypical or delayed oral language development (Catts et al., 2009; Peterson et al., 2009), writing disabilities, ADHD, and math disabilities (e.g., Landerl and Moll, 2010; Willcutt et al., 2010). This, together with the generally large inter-individual differences with respect to responsiveness to reading intervention, may lead to higher variability of (potential) neuroplasticity effects, which, in turn, may lead to weaker meta-analytic results.

There is no clarity about whether specific regions or patterns of activation are required in order to provoke improvements in reading ability in RD. Numerous studies have shown effects within the typical reading network as well as outside. The differentiation between "normalization" vs. "compensation" effects is more complex than detecting activation in certain brain areas because multiple brain regions linked to the typical reading network are associated with other, more general cognitive networks as well (e.g., attention and executive function networks).

Future studies should try and identify networks of activation in addition to fundamental structural changes linked to improvement in reading ability. Furthermore, several factors regarding individual differences and interventions should guide research on the neural mechanisms of reading intervention. One way of providing more thorough evidence would be via longitudinal studies with a longer time frame (i.e., going beyond sole pre-/post-intervention assessments). Although extremely expensive and challenging to conduct, such studies would be desperately needed (Chyl et al., 2021).

Another desirable and extremely worthwhile approach would be to conduct multi-center studies with sufficient sample sizes, where the same methodologies are applied in a concerted and standardized effort. For example, Paulesu et al. (2001) investigated cultural differences across people with RD in the course of a cross-European PET project, whereas Jednoróg et al. (2015) conducted a large-scale multi-center, multi-language VBM study. Recently, another study showed that brain-wide association between inter-individual differences in brain structure or function and complex cognitive or mental health phenotypes, such as reading disability and its remediation, requires thousands of individuals (Marek et al., 2022).

Last but not least, the general issue of publication bias (i.e., withholding null findings and publishing statistically significant results) might create a false impression of substantial and reliable brain changes linked to reading intervention. This could explain some of the observed contradictory findings between studies reviewed here. To summarize, systematic and robust neuroplasticity effects in response to reading improvements across many studies could not yet be found. Therefore, further (pre-registered) research on the interplay between behavioral reading intervention and the brain mechanisms underlying typical and atypical reading is needed.

CONCLUSION

In recent years, outstanding progress has been made in understanding the functional neuroanatomy of typical reading, RD, and reading intervention for RD. Our review of studies suggests that enhanced activation in right-hemispheric homologous regions of the typical left hemisphere reading network following behavioral intervention might reflect functional neuroanatomical normalization rather than compensation of brain mechanisms for reading. With respect to rewiring of white matter network connections in response to intervention, neural optimization might be related to both, the establishment of stronger brain circuits and the reduction of inefficient connections in RD.

Nevertheless, the field suffers from a lack of consistent neuroplasticity effects associated with improvement in reading ability across studies. Future studies should examine inter-individual differences and developmental trajectories more closely over a longer time frame. Additionally, the common dichotomy between "normalization" vs. "compensation" seems to be insufficient to explain the complex underlying neurobiology and a more integrated view of the brain mechanisms related to reading intervention should be employed.

AUTHOR CONTRIBUTIONS

JB and FR conceived and wrote the manuscript. Both authors contributed to the article and approved the submitted version.

REFERENCES

Albajes-Eizagirre, A., Solanes, A., Fullana, M. A., Ioannidis, J. P. A., Fusar-Poli, P., Torrent, C., et al. (2019). Meta-analysis of Voxel-Based Neuroimaging Studies using Seed-based d Mapping with Permutation of Subject Images (SDM-PSI). *J. Vis. Exp.* 153:59841. doi: 10.3791/59841

Alvarez, T. A., and Fiez, J. A. (2018). Current perspectives on the cerebellum and reading development. *Neurosci. Biobehav. Rev.* 92, 55–66. doi: 10.1016/j.neubiorev.2018.05.006

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders*, 5th Edn. Virginia: American Psychiatric Association, doi: 10.1176/appi.books.9780890425596

Aylward, E., Richards, T., Berninger, V., Nagy, W., Field, K., Grimme, A., et al. (2003). Instructional treatment associated with changes in brain activation in children with dyslexia. *Neurology* 61, 212–219. doi: 10.1212/01.wnl.0000068363.05974.64

Barquero, L. A., Davis, N., and Cutting, L. E. (2014). Neuroimaging of Reading Intervention: a Systematic Review and Activation Likelihood Estimate Meta-Analysis. *PLoS One* 9:e83668. doi: 10.1371/journal.pone.0083668

Ben-Shachar, M., Dougherty, R. F., and Wandell, B. A. (2007). White matter pathways in reading. *Curr. Opin. Neurob.* 17, 258–270. doi: 10.1016/j.conb.2007.03.006

Blau, V., Reithler, J., van Atteveldt, N., Seitz, J., Gerretsen, P., Goebel, R., et al. (2010). Deviant processing of letters and speech sounds as proximate cause of reading failure: a functional magnetic resonance imaging study of dyslexic children. *Brain* 133, 868–879. doi: 10.1093/brain/awp308

Blomert, L. (2011). The neural signature of orthographic–phonological binding in successful and failing reading development. *NeuroImage* 57, 695–703. doi: 10.1016/j.neuroimage.2010.11.003

Boets, B., Beeck, H. P., Vandermosten, M., Scott, S. K., Gillebert, C. R., Mantini, D., et al. (2013). Intact But Less Accessible Phonetic Representations in Adults with Dyslexia. *Science* 342, 1251–1254. doi: 10.1126/science.1244333

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., et al. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nat. Rev. Neurosci.* 14, 365–376. doi: 10.1038/nrn3475

Cao, F., Yan, X., Wang, Z., Liu, Y., Wang, J., Spray, G. J., et al. (2017). Neural signatures of phonological deficits in Chinese developmental dyslexia. *NeuroImage* 146, 301–311. doi: 10.1016/j.neuroimage.2016.11.051

Carreiras, M., Armstrong, B. C., Perea, M., and Frost, R. (2014). The what, when, where, and how of visual word recognition. *Trends Cogn. Sci.* 18, 90–98. doi: 10.1016/j.tics.2013.11.005

Cattinelli, I., Borghese, N. A., Gallucci, M., and Paulesu, E. (2013). Reading the reading brain: A new meta-analysis of functional imaging data on reading. *J. Neuroling.* 26, 214–238. doi: 10.1016/j.jneuroling.2012.08.001

Catts, H. W., Bridges, M. S., Little, T. D., and Tomblin, J. B. (2009). Reading achievement growth in children with language impairments. *J. Speech, Lang. Hear. Res.* 42, 163–176. doi: 10.1177/0022219408326219

Chyl, K., Fraga-González, G., Brem, S., and Jednoróg, K. (2021). Brain dynamics of (a)typical reading development—a review of longitudinal studies. *Npj Sci. Learn.* 6:1. doi: 10.1038/s41539-020-00081-5

Danelli, L., Berlingeri, M., Bottini, G., Ferri, F., Vacchi, L., Sberna, M., et al. (2012). Neural intersections of the phonological, visual magnocellular and motor/cerebellar systems in normal readers: Implications for imaging studies on dyslexia. *Hum. Brain Mapp.* 34, 2669–2687. doi: 10.1002/hbm.22098

Davis, N., Fan, Q., Compton, D. L., Fuchs, D., Fuchs, L. S., Cutting, L. E., et al. (2010). Influences of neural pathway integrity on children's response to reading instruction. *Front. Syst. Neurosci.* 4:150. doi: 10.3389/fnsys.2010.00150

De Vos, A., Vanvooren, S., Ghesquière, P., and Wouters, J. (2020). Subcortical auditory neural synchronization is deficient in pre-reading children who develop dyslexia. *Devel. Sci.* 23:6. doi: 10.1111/desc.12945

Dehaene, S., Cohen, L., Morais, J., and Kolinsky, R. (2015). Illiterate to literate: behavioural and cerebral changes induced by reading acquisition. *Nat. Rev. Neurosci.* 16, 234–244. doi: 10.1038/nrn3924

Dolan, R. J., Fletcher, P., and Paulesu, E. (1997). "The functional anatomy of reading," in *Human brain function*, ed. R. S. J. Frackowiak (San Diego: Academic Press), 301–328.

Eckert, M. A., Berninger, V. W., Vaden, K. I., Gebregziabher, M., and Tsu, L. (2016). Gray Matter Features of Reading Disability: a Combined Meta-Analytic and Direct Analysis Approach. *Eneuro* 3:15. doi: 10.1523/eneuro.0103-15.2015

Fraga González, G., ŽArić, G., Tijms, J., Bonte, M., Blomert, L., and van der Molen, M. W. (2015). A Randomized Controlled Trial on The Beneficial Effects of Training Letter-Speech Sound Integration on Reading Fluency in Children with Dyslexia. *PLoS One* 10:e0143914. doi: 10.1371/journal.pone.0143914

Galuschka, K., Ise, E., Krick, K., and Schulte-Körne, G. (2014). Effectiveness of Treatment Approaches for Children and Adolescents with Reading Disabilities: a Meta-Analysis of Randomized Controlled Trials. *PLoS One* 9:e89900. doi: 10.1371/journal.pone.0089900

Gebauer, D., Fink, A., Kargl, R., Reishofer, G., Koschutnig, K., Purgstaller, C., et al. (2012). Differences in Brain Function and Changes with Intervention in Children with Poor Spelling and Reading Abilities. *PLoS One* 7:e38201. doi: 10.1371/journal.pone.0038201

Geschwind, N. (1965a). Disconnexion Syndromes in animals and man. I. *Brain* 88:237. doi: 10.1093/brain/88.2.237

Geschwind, N. (1965b). Disconnexion Syndromes in animals and man. II. *Brain* 88:585. doi: 10.1093/brain/88.3.585

Hancock, R., Richlan, F., and Hoeft, F. (2017). Possible roles for fronto-striatal circuits in reading disorder. *Neurosci. Biobehav. Rev.* 72, 243–260. doi: 10.1016/j.neubiorev.2016.10.025

Heim, S., Pape-Neumann, J., van Ermingen-Marbach, M., Brinkhaus, M., and Grande, M. (2015). Shared vs. specific brain activation changes in dyslexia after training of phonology, attention, or reading. *Brain Struct. Funct.* 220, 2191–2207. doi: 10.1007/s00429-014-0784-y

Hoeft, F., McCandliss, B. D., Black, J. M., Gantman, A., Zakerani, N., Hulme, C., et al. (2011). Neural systems predicting long-term outcome in dyslexia. *Proc. Natl. Acad. Sci.* 108, 361–366. doi: 10.1073/pnas.1008950108

Holloway, I. D., van Atteveldt, N., Blomert, L., and Ansari, D. (2013). Orthographic Dependency in the Neural Correlates of Reading: evidence from Audiovisual Integration in English Readers. *Cereb. Cortex* 25, 1544–1553. doi: 10.1093/cercor/bht347

Horowitz-Kraus, T., Hershey, A., Kay, B., and DiFrancesco, M. (2019). Differential effect of reading training on functional connectivity in children with reading difficulties with and without ADHD comorbidity. *J. Neuroling.* 49, 93–108. doi: 10.1016/j.jneuroling.2018.09.002

Horowitz-Kraus, T., Toro-Serey, C., and DiFrancesco, M. (2015). Increased Resting-State Functional Connectivity in the Cingulo-Opercular Cognitive-Control Network after Intervention in Children with Reading Difficulties. *PLoS One* 10:e0133762. doi: 10.1371/journal.pone.0133762

Horowitz-Kraus, T., Vannest, J. J., Kadis, D., Cicchino, N., Wang, Y. Y., and Holland, S. K. (2014). Reading acceleration training changes brain circuitry in children with reading difficulties. *Brain Behav.* 4, 886–902. doi: 10.1002/brb3.281

Huber, E., Donnelly, P. M., Rokem, A., and Yeatman, J. D. (2018). Rapid and widespread white matter plasticity during an intensive reading intervention. *Nat. Comm.* 9:1. doi: 10.1038/s41467-018-04627-5

Jednoróg, K., Marchewka, A., Altarelli, I., Monzalvo Lopez, A. K., van Ermingen-Marbach, M., Grande, M., et al. (2015). How reliable are gray matter disruptions in specific reading disability across multiple countries and languages? insights from a large-scale voxel-based morphometry study. *Hum. Brain Mapp.* 36, 1741–1754. doi: 10.1002/hbm.22734

Koyama, M. S., di Martino, A., Kelly, C., Jutagir, D. R., Sunshine, J., Schwartz, S. J., et al. (2013). Cortical Signatures of Dyslexia and Remediation: an Intrinsic Functional Connectivity Approach. *PLoS One* 8:e55454. doi: 10.1371/journal.pone.0055454

Krafnick, A. J., Flowers, D. L., Napoliello, E. M., and Eden, G. F. (2011). Gray matter volume changes following reading intervention in dyslexic children. *NeuroImage* 57, 733–741. doi: 10.1016/j.neuroimage.2010.10.062

Landerl, K., and Moll, K. (2010). Comorbidity of learning disorders: prevalence and familial transmission. *J. Child Psychol. Psychiatry* 51, 287–294. doi: 10.1111/j.1469-7610.2009.02164.x

Linkersdörfer, J., Lonnemann, J., Lindberg, S., Hasselhorn, M., and Fiebach, C. J. (2012). Grey Matter Alterations Co-Localize with Functional Abnormalities in Developmental Dyslexia: An ALE Meta-Analysis. *PLoS One* 7:e43122. doi: 10.1371/journal.pone.0043122

Lyon, G. R., Shaywitz, S. E., and Shaywitz, B. A. (2003). A definition of dyslexia. *Ann. Dysl.* 53, 1–14. doi: 10.1007/s11881-003-0001-9

Maisog, J. M., Einbinder, E. R., Flowers, D. L., Turkeltaub, P. E., and Eden, G. F. (2008). A Meta-analysis of Functional Neuroimaging Studies of Dyslexia. *Ann. N Y Acad. Sci.* 1145, 237–259. doi: 10.1196/annals.1416.024

Marek, S., Tervo-Clemmens, B., Calabro, F. J., Montez, D. F., Kay, B. P., Hatoum, A. S., et al. (2022). Reproducible brain-wide association studies require thousands of individuals. *Nature* 603, 654–660. doi: 10.1038/s41586-022-04492-9

Martin, A., Kronbichler, M., and Richlan, F. (2016). Dyslexic brain activation abnormalities in deep and shallow orthographies: a meta-analysis of 28 functional neuroimaging studies. *Hum. Brain Mapp.* 37, 2676–2699. doi: 10.1002/hbm.23202

Martin, A., Schurz, M., Kronbichler, M., and Richlan, F. (2015). Reading in the brain of children and adults: a meta-analysis of 40 functional magnetic resonance imaging studies. *Hum. Brain Mapp.* 36, 1963–1981. doi: 10.1002/hbm.22749

Mascheretti, S., de Luca, A., Trezzi, V., Peruzzo, D., Nordio, A., Marino, C., et al. (2017). Neurogenetics of developmental dyslexia: from genes to behavior through brain neuroimaging and cognitive and sensorial mechanisms. *Transl. Psychiatry* 7:e987. doi: 10.1038/tp.2016.240

Mechelli, A., Crinion, J. T., Long, S., Friston, K. J., Ralph, M. A. L., Patterson, K., et al. (2005). Dissociating Reading Processes on the Basis of Neuronal Interactions. *J. Cogn. Neurosci.* 17, 1753–1765. doi: 10.1162/089892905774589190

Meyler, A., Keller, T. A., Cherkassky, V. L., Gabrieli, J. D., and Just, M. A. (2008). Modifying the brain activation of poor readers during sentence comprehension with extended remedial instruction: a longitudinal study of neuroplasticity. *Neuropsychologia* 46, 2580–2592. doi: 10.1016/j.neuropsychologia.2008.03.012

Morken, F., Helland, T., Hugdahl, K., and Specht, K. (2017). Reading in dyslexia across literacy development: a longitudinal study of effective connectivity. *NeuroImage* 144, 92–100. doi: 10.1016/j.neuroimage.2016.09.060

Müller-Axt, C., Anwander, A., and von Kriegstein, K. (2017). Altered Structural Connectivity of the Left Visual Thalamus in Developmental Dyslexia. *Curr. Biol.* 27, 3692.e–3698.e. doi: 10.1016/j.cub.2017.10.034

Nugiel, T., Roe, M. A., Taylor, W. P., Cirino, P. T., Vaughn, S. R., Fletcher, J. M., et al. (2019). Brain activity in struggling readers before intervention relates to future reading gains. *Cortex* 111, 286–302. doi: 10.1016/j.cortex.2018.11.009

Odegard, T. N., Ring, J., Smith, S., Biggan, J., and Black, J. (2008). Differentiating the neural response to intervention in children with developmental dyslexia. *Ann. Dysl.* 58, 1–14. doi: 10.1007/s11881-008-0014-5

Olulade, O., Flowers, D., Napoliello, E., and Eden, G. (2015). Dyslexic children lack word selectivity gradients in occipito-temporal and inferior frontal cortex. *NeuroImage: Clin.* 7, 742–754. doi: 10.1016/j.nicl.2015.02.013

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021:n71. doi: 10.1136/bmj.n71

Partanen, M., Siegel, L. S., and Giaschi, D. E. (2019). Effect of reading intervention and task difficulty on orthographic and phonological reading systems in the brain. *Neuropsychologia* 130, 13–25. doi: 10.1016/j.neuropsychologia.2018.07.018

Paulesu, E., Danelli, L., and Berlingeri, M. (2014). Reading the dyslexic brain: multiple dysfunctional routes revealed by a new meta-analysis of PET and fMRI activation studies. *Front. Hum. Neurosci.* 8:830. doi: 10.3389/fnhum.2014.00830

Paulesu, E., Demonet, J. F., Fazio, F., McCrory, E., Chanoine, V., Brunswick, N., et al. (2001). Dyslexia: cultural Diversity and Biological Unity. *Science* 291, 2165–2167. doi: 10.1126/science.1057179

Paulesu, E., Frith, U., Snowling, M., Gallagher, A., Morton, J., Frackowiak, R. S. J., et al. (1996). Is developmental dyslexia a disconnection syndrome? *Brain* 119, 143–157. doi: 10.1093/brain/119.1.143

Paulesu, E., McCrory, E., Fazio, F., Menoncello, L., Brunswick, N., Cappa, S. F., et al. (2000). A cultural effect on brain function. *Nat. Neurosci.* 3, 91–96. doi: 10.1038/71163

Perdue, M. V., Mahaffy, K., Vlahcevic, K., Wolfman, E., Erbeli, F., Richlan, F., et al. (2022). Reading intervention and neuroplasticity: a systematic review and meta-analysis of brain changes associated with reading intervention. *Neurosci. Biobehav. Rev.* 132, 465–494. doi: 10.1016/j.neubiorev.2021.11.011

Peterson, R. L., Pennington, B. F., Shriberg, L. D., and Boada, R. (2009). What Influences Literacy Outcome in Children With Speech Sound Disorder? *J. Speech Lang. Hear. Res.* 52, 1175–1188. doi: 10.1044/1092-4388(2009/08-0024)

Ramus, F., Altarelli, I., Jednoróg, K., Zhao, J., Scotto, and Di Covella, L. (2018). Neuroanatomy of developmental dyslexia: Pitfalls and promise. *Neurosci. Biobehav. Rev.* 84, 434–452. doi: 10.1016/j.neubiorev.2017.08.001

Ramus, F., and Szenkovits, G. (2008). What Phonological Deficit? *Q. J. Exp. Psychol.* 61, 129–141. doi: 10.1080/17470210701508822

Richards, T., Peverly, S., Wolf, A., Abbott, R., Tanimoto, S., Thompson, R., et al. (2016). Idea units in notes and summaries for read texts by keyboard and pencil in middle childhood students with specific learning disabilities: cognitive and brain findings. *Trends Neurosci. Educ.* 5, 146–155. doi: 10.1016/j.tine.2016.07.005

Richards, T. L., Aylward, E. H., Field, K. M., Grimme, A. C., Raskind, W., Richards, A. L., et al. (2006b). Converging Evidence for Triple Word Form Theory in Children With Dyslexia. *Dev. Neuropsychol.* 30, 547–589. doi: 10.1207/s15326942dn3001_3

Richards, T. L., Aylward, E. H., Berninger, V. W., Field, K. M., Grimme, A. C., Richards, A. L., et al. (2006a). Individual fMRI activation in orthographic mapping and morpheme mapping after orthographic or morphological spelling treatment in child dyslexics. *J. Neuroling.* 19, 56–86. doi: 10.1016/j.jneuroling.2005.07.003

Richards, T. L., Berninger, V. W., Yagle, K., Abbott, R. D., and Peterson, D. (2018). Brain's functional network clustering coefficient changes in response to instruction (RTI) in students with and without reading disabilities: Multi-leveled reading brain's RTI. *Cog. Psychol.* 5:1424680. doi: 10.1080/23311908.2018.1424680

Richards, T. L., Berninger, V. W., Yagle, K. J., Abbott, R. D., and Peterson, D. J. (2017). Changes in DTI diffusivity and fMRI connectivity cluster coefficients for students with and without specific learning disabilities in written language: brain's response to writing instruction. *J. Nat. Sci.* 3, 1–26.

Richlan, F. (2012). Developmental dyslexia: dysfunction of a left hemisphere reading network. *Front. Hum. Neurosci.* 6:120. doi: 10.3389/fnhum.2012.00120

Richlan, F. (2014). Functional neuroanatomy of developmental dyslexia: the role of orthographic depth. *Front. Hum. Neurosci.* 8:347. doi: 10.3389/fnhum.2014.00347

Richlan, F. (2019). The Functional Neuroanatomy of Letter-Speech Sound Integration and Its Relation to Brain Abnormalities in Developmental Dyslexia. *Front. Hum. Neurosci.* 13:21. doi: 10.3389/fnhum.2019.00021

Richlan, F. (2020). The Functional Neuroanatomy of Developmental Dyslexia Across Languages and Writing Systems. *Front. Psychol.* 11:155. doi: 10.3389/fpsyg.2020.00155

Richlan, F., Kronbichler, M., and Wimmer, H. (2009). Functional abnormalities in the dyslexic brain: a quantitative meta-analysis of neuroimaging studies. *Hum. Brain Mapp.* 30, 3299–3308. doi: 10.1002/hbm.20752

Richlan, F., Kronbichler, M., and Wimmer, H. (2011). Meta-analyzing brain dysfunctions in dyslexic children and adults. *NeuroImage* 56, 1735–1742. doi: 10.1016/j.neuroimage.2011.02.040

Richlan, F., Kronbichler, M., and Wimmer, H. (2013). Structural abnormalities in the dyslexic brain: a meta-analysis of voxel-based morphometry studies. *Hum. Brain Mapp.* 34, 3055–3065. doi: 10.1002/hbm.22127

Richlan, F., Sturm, D., Schurz, M., Kronbichler, M., Ladurner, G., and Wimmer, H. (2010). A Common Left Occipito-Temporal Dysfunction in Developmental Dyslexia and Acquired Letter-By-Letter Reading? *PLoS One* 5:e12073. doi: 10.1371/journal.pone.0012073

Romeo, R. R., Christodoulou, J. A., Halverson, K. K., Murtagh, J., Cyr, A. B., Schimmel, C., et al. (2017). Socioeconomic Status and Reading Disability: neuroanatomy and Plasticity in Response to Intervention. *Cereb. Cort.* 28, 2297–2312. doi: 10.1093/cercor/bhx131

Schlaggar, B. L., and McCandliss, B. D. (2007). Development of Neural Systems for Reading. *Ann. Rev. Neurosci.* 30, 475–503. doi: 10.1146/annurev.neuro.28.061604.135645

Schurz, M., Kronbichler, M., Crone, J., Richlan, F., Klackl, J., and Wimmer, H. (2014). Top-down and bottom-up influences on the left ventral occipito-temporal cortex during visual word recognition: an analysis of effective connectivity. *Hum. Brain Mapp.* 35, 1668–1680. doi: 10.1002/hbm.22281

Schurz, M., Sturm, D., Richlan, F., Kronbichler, M., Ladurner, G., and Wimmer, H. (2010). A dual-route perspective on brain activation in response to visual words: evidence for a length by lexicality interaction in the visual word form area (VWFA). *NeuroImage* 49, 2649–2661. doi: 10.1016/j.neuroimage.2009.10.082

Schurz, M., Wimmer, H., Richlan, F., Ludersdorfer, P., Klackl, J., and Kronbichler, M. (2015). Resting-State and Task-Based Functional Brain Connectivity in Developmental Dyslexia. *Cereb. Cortex* 25, 3502–3514. doi: 10.1093/cercor/bhu184

Schuster, S., Hawelka, S., Hutzler, F., Kronbichler, M., and Richlan, F. (2016). Words in Context: the Effects of Length, Frequency, and Predictability on Brain Responses During Natural Reading. *Cereb. Cort.* 26, 3889.2–3904. doi: 10.1093/cercor/bhw184

Shaywitz, B. A., Shaywitz, S. E., Blachman, B. A., Pugh, K. R., Fulbright, R. K., Skudlarski, P., et al. (2004). Development of left occipitotemporal systems for skilled reading in children after a phonologically- based intervention. *Biolog. Psychiatry* 55, 926–933. doi: 10.1016/j.biopsych.2003.12.019

Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Mencl, W., Fulbright, R. K., Skudlarski, P., et al. (2002). Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biolog. Psychiatry* 52, 101–110. doi: 10.1016/s0006-3223(02)01365-3

Silani, G., Frith, U., Demonet, J. F., Fazio, F., Perani, D., Price, C., et al. (2005). Brain abnormalities underlying altered activation in dyslexia: a voxel based morphometry study. *Brain* 128, 2453–2461. doi: 10.1093/brain/awh579

Suggate, S. P. (2010). Why what we teach depends on when: grade and reading intervention modality moderate effect size. *Dev. Psychol.* 46, 1556–1579. doi: 10.1037/a0020612

Temple, E., Deutsch, G. K., Poldrack, R. A., Miller, S. L., Tallal, P., Merzenich, M. M., et al. (2003). Neural deficits in children with dyslexia ameliorated by behavioral remediation: Evidence from functional MRI. *Proc. Natl. Acad. Sci.* 100, 2860–2865. doi: 10.1073/pnas.0030098100

Thiebaut de Schotten, M., Cohen, L., Amemiya, E., Braga, L. W., and Dehaene, S. (2012). Learning to Read Improves the Structure of the Arcuate Fasciculus. *Cereb. Cort.* 24, 989–995. doi: 10.1093/cercor/bhs383

Van Atteveldt, N., Formisano, E., Goebel, R., and Blomert, L. (2004). Integration of Letters and Speech Sounds in the Human Brain. *Neuron* 43, 271–282. doi: 10.1016/j.neuron.2004.06.025

Van der Mark, S., Klaver, P., Bucher, K., Maurer, U., Schulz, E., Brem, S., et al. (2011). The left occipitotemporal system in reading: Disruption of focal fMRI connectivity to left inferior frontal and inferior parietal language areas in children with dyslexia. *NeuroImage* 54, 2426–2436. doi: 10.1016/j.neuroimage.2010.10.002

Vandermosten, M., Boets, B., Wouters, J., and Ghesquière, P. (2012). A qualitative and quantitative review of diffusion tensor imaging studies in reading and dyslexia. *Neurosci. Biobehav. Rev.* 36, 1532–1552. doi: 10.1016/j.neubiorev.2012.04.002

Vogel, A. C., Church, J. A., Power, J. D., Miezin, F. M., Petersen, S. E., and Schlaggar, B. L. (2013). Functional network architecture of reading-related regions across development. *Brain Lang.* 125, 231–243. doi: 10.1016/j.bandl.2012.12.016

Wanzek, J., Stevens, E. A., Williams, K. J., Scammacca, N., Vaughn, S., and Sargent, K. (2018). Current Evidence on the Effects of Intensive Early Reading Interventions. *J. Learn. Disab.* 51, 612–624. doi: 10.1177/0022219418775110

Wanzek, J., Vaughn, S., Scammacca, N., Gatlin, B., Walker, M. A., and Capin, P. (2016). Meta-Analyses of the Effects of Tier 2 Type Reading Interventions in Grades K-3. *Educ. Psychol. Rev.* 28, 551–576. doi: 10.1007/s10648-015-9321-7

Wanzek, J., Vaughn, S., Scammacca, N. K., Metz, K., Murray, C. S., Roberts, G., et al. (2013). Extensive Reading Interventions for Students With Reading Difficulties After Grade 3. *Rev. Educ. Res.* 83, 163–195. doi: 10.3102/0034654313477212

Willcutt, E. G., Betjemann, R. S., McGrath, L. M., Chhabildas, N. A., Olson, R. K., DeFries, J. C., et al. (2010). Etiology and neuropsychology of comorbidity between RD and ADHD: the case for multiple-deficit models. *Cortex* 46, 1345–1361. doi: 10.1016/j.cortex.2010.06.009

Wimmer, H. (1993). Characteristics of developmental dyslexia in a regular writing system. *Appl. Psycholing.* 14, 1–33. doi: 10.1017/s0142716400010122

Wimmer, H., Schurz, M., Sturm, D., Richlan, F., Klackl, J., Kronbichler, M., et al. (2010). A dual-route perspective on poor reading in a regular orthography: an fMRI study. *Cortex* 46, 1284–1298. doi: 10.1016/j.cortex.2010.06.004

World Health Organization (2016). *ICD-10 – International Statistical Classification of Diseases and Related Health Problems 10th Revision*, 5th Edn. Geneva: World Health Organization.

Yeatman, J. D., Dougherty, R. F., Ben-Shachar, M., and Wandell, B. A. (2012). Development of white matter and reading skills. *Proc. Natl. Acad. Sci.* 109:44. doi: 10.1073/pnas.1206792109

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Neuroscience

SYSTEMATIC REVIEW
published: 30 June 2022
doi: 10.3389/fnins.2022.898800

Typical and Atypical Development of Visual Expertise for Print as Indexed by the Visual Word N1 (N170w): A Systematic Review

Kathleen Kay Amora[1,2*†], **Ariane Tretow**[3†], **Cara Verwimp**[4,5†], **Jurgen Tijms**[4,5], **Paavo H. T. Leppänen**[3] and **Valéria Csépe**[1,6]

[1] Brain Imaging Centre, Research Centre for Natural Sciences, Budapest, Hungary, [2] Faculty of Modern Philology and Social Sciences, Multilingualism Doctoral School, University of Pannonia, Veszprém, Hungary, [3] Department of Psychology, University of Jyväskylä, Jyväskylä, Finland, [4] Department of Developmental Psychology, University of Amsterdam, Amsterdam, Netherlands, [5] Rudolf Berlin Center, Amsterdam, Netherlands, [6] Institute for Hungarian and Applied Linguistics, University of Pannonia, Veszprém, Hungary

The visual word N1 (N170w) is an early brain ERP component that has been found to be a neurophysiological marker for print expertise, which is a prelexical requirement associated with reading development. To date, no other review has assimilated existing research on reading difficulties and atypical development of processes reflected in the N170w response. Hence, this systematic review synthesized results and evaluated neurophysiological and experimental procedures across different studies about visual print expertise in reading development. Literature databases were examined for relevant studies from 1995 to 2020 investigating the N170w response in individuals with or without reading disorders. To capture the development of the N170w related to reading, results were compared between three different age groups: pre-literate children, school-aged children, and young adults. The majority of available N170w studies ($N = 69$) investigated adults ($n = 31$) followed by children (school-aged: $n = 21$; pre-literate: $n = 4$) and adolescents ($n = 1$) while some studies investigated a combination of these age groups ($n = 12$). Most studies were conducted with German-speaking populations ($n = 17$), followed by English ($n = 15$) and Chinese ($n = 14$) speaking participants. The N170w was primarily investigated using a combination of words, pseudowords, and symbols ($n = 20$) and mostly used repetition-detection ($n = 16$) or lexical-decision tasks ($n = 16$). Different studies posed huge variability in selecting electrode sites for analysis; however, most focused on P7, P8, and O1 sites of the international 10–20 system. Most of the studies in adults have found a more negative N170w in controls than poor readers, whereas in children, the results have been mixed. In typical readers, N170w ranged from having a bilateral distribution to a left-hemispheric dominance throughout development, whereas in young, poor readers, the response was mainly right-lateralized and then remained in a bilateral distribution. Moreover, the N170w latency has varied according to age group, with adults having an earlier onset yet with shorter latency than school-aged and pre-literate children. This systematic review provides a comprehensive picture of the development of

Edited by:
Li-Hai Tan,
Shenzhen Institute of Neuroscience, China

Reviewed by:
Anthony James Krafnick,
Dominican University, United States
Yang Zhang,
University of Minnesota Health Twin Cities, United States

***Correspondence:**
Kathleen Kay Amora
amora.kathleen@ttk.hu;
amora.kathleen@phd.uni-pannon.hu

† These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Neurodevelopment, a section of the journal Frontiers in Neuroscience

Received: *17 March 2022*
Accepted: *24 May 2022*
Published: *30 June 2022*

Citation:
Amora KK, Tretow A, Verwimp C, Tijms J, Leppänen PHT and Csépe V (2022) Typical and Atypical Development of Visual Expertise for Print as Indexed by the Visual Word N1 (N170w): A Systematic Review. Front. Neurosci. 16:898800. doi: 10.3389/fnins.2022.898800

print expertise as indexed by the N170w across age groups and reading abilities and discusses theoretical and methodological differences and challenges in the field, aiming to guide future research.

Systematic Review Registration: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42021228444.

Keywords: reading development, dyslexia, words, developmental reading disorder (DRD), event-related potentials (ERP), visual expertise, N170, systematic review

INTRODUCTION

Reading, which involves successfully and fluently linking letters to sounds, is one of the prerequisites to participate in today's society. Learning to read is commonly shaped through years of exposure to text and formal teaching. Although we are constantly exposed to text, some do not successfully develop fluent reading skills, with the poorest 3–10% of the children being considered to have developmental dyslexia or developmental reading disorder (DRD; Snowling, 2013).

Fast recognition of words is critical for attaining automatized reading in alphabetic orthographies (McCandliss et al., 2003) and is associated with a reorganization of the visual systems that are evolving to process the new word forms efficiently. Event-related potential (ERP) studies have associated the visual N170 component, which peaks around 170 milliseconds after stimulus onset, with the expertise for visual stimuli such as words. The visual word N170 (hereafter referred to as N170w) is a response with a negative deflection commonly largest over occipitotemporal regions, and its lateralization depends on maturation and reading experience (Maurer and McCandliss, 2008). The emergence of N170w is supposedly rooted in the visual word form area (VWFA) within the ventral occipitotemporal cortex (vOTC) of the left hemisphere, which has been known to show sensitivity to visual words throughout literacy (McCandliss et al., 2003; Rossion et al., 2003; Dehaene et al., 2010). Moreover, it has been considered as a neurophysiological marker for print expertise with prelexical sensitivity to letter/character strings (Maurer et al., 2006; Luck, 2012). Higher N1 amplitudes for words than low-level visual control stimuli such as meaningless symbol strings or shapes have been reported across languages (e.g., Dutch: Fraga González et al., 2014, German: Maurer et al., 2006, Portuguese: Araújo et al., 2012).

Several studies have explored the N170 component, which is reported as a category-specific visual expertise marker (Maurer et al., 2008b), and has been studied extensively in face perception studies (e.g., Bentin et al., 1996; Feuerriegel et al., 2015). Other studies have also associated the N170 with sensory processing related to auditory information (Leppänen and Lyytinen, 1997) and referred to the modulation of N170 by attention (Herrmann and Knight, 2001). However, the N170w associated with print tuning has become of particular interest in reading disorder studies in recent years. Aside from the mismatch negativity (MMN) which is commonly used to discuss the role of auditory processing in reading development, the N170w provides a more reading- specific insight related to visual processing for print, which is the primary visual stimuli for reading. Moreover, N170 is reported to possibly predict later reading outcomes as the N170w response is modulated by reading skills (Brem et al., 2013). Furthermore, the N170 has a role in attention, which could be taken into consideration in relation to the visual attention span deficit theory, referring to a higher attention level required in dyslexics for processing of words. Different investigations aiming at characterizing the N170w have identified two different processes; coarse and fine print tuning (e.g., Zhao et al., 2014; Tong et al., 2016a; Kemény et al., 2018). Coarse print tuning, which indicates sublexical processing, entails differential processing of words and non-orthographic symbol strings, whereas fine print tuning usually taps into lexical processes and is required for processing of differences between print and closely matched false font or pseudo-character strings (Maurer et al., 2005; Eberhard-Moscicka et al., 2016). Even though many studies have aspired to shed light on the main visual component with reading development, most of them performed in typical readers or reading disordered individuals have produced contradictory results. These could be due to variability in participant groups, stimuli, and task-specific factors.

Zhao et al. (2014) demonstrated that coarse and even fine-tuning of the N170w can be developed within 1 year of reading instruction. However, N170w print specialization has been found to occur later in children with DRD (Maurer et al., 2007, 2011), suggesting differences in the developmental trajectory of N170w specialization of individuals with DRD compared to their typically developing peers. Longitudinal studies have shown an inverted U-shape development curve of the N170w, with an increased response for orthographic stimuli in beginning readers followed by a slight decrease when readers become fluent (Maurer et al., 2006; Fraga González et al., 2021). However, some studies have shown evidence for a persistent N170w print tuning deficit in individuals with DRD, with no or small differences in the N170w responses to word-like stimuli and matched symbol strings in adults compared to their typically developing peers (Mahé et al., 2012). In addition, for print, it has been found that a bilateral, though somewhat right hemisphere-dominated N170w topography in children changes gradually into left-lateralized topography when reading becomes more automatized. This change occurs shortly after the start of formal reading instruction, contributing to letter-speech sound integration in the form of grapheme-phoneme correspondences (Maurer et al., 2006; Brem et al., 2013). However, for individuals with DRD, the response lateralization showed no consistent pattern: left (e.g., Araújo et al., 2012), bilateral (e.g., Fraga González et al., 2014), or right-lateralized (e.g., van Setten et al., 2019) distributions were reported.

Although numerous studies have demonstrated an atypical N170w response to words in individuals with DRD, the effects regarding amplitude, latency, and lateralization have been inconsistent. Moreover, the variation in experimental designs and setups could pose challenges in interpreting results for interested researchers in the field. Therefore, our systematic review assimilated existing research on typical and atypical development of visual reading processes as reflected by the N170w response. The main objective of this review was to give an overview of the status quo of the N170w literature related to reading development in terms of reading ability (typically vs. atypically developing readers) and age group (from pre-literate age until adulthood). For our secondary objectives, we examined differences in N170w in comparison with other word-like conditions (e.g., pseudowords, nonwords) and the potential impact of various linguistic factors (e.g., language, orthographic depth). In addition, we investigated theoretical and methodological differences applied in the N170w studies to guide future research using this component to investigate typical and atypical reading.

MATERIALS AND METHODS

Protocol and Registration

The protocol for this systematic review was pre-registered and uploaded to https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42021228444. All aspects of this review adhered to the Preferred Reporting Items in Systematic Reviews (PRISMA) guidelines (Moher et al., 2009).

Eligibility Criteria and Study Selection

Studies included in the current review satisfied the following criteria after the full-text review: (1) cross-sectional, longitudinal, and intervention studies on the visual word N1/N170 employing different stimulus conditions, i.e., letter/character strings vs. non-letter/non-character stimuli (case studies, reviews, theses or dissertations, and gray literature were excluded; as well as using single letters only as stimuli was excluded); (2) a sample involving participants with or/and without developmental reading disorders (DRD) (studies that focused only on other neurological/developmental conditions or comorbidities aside from DRD (e.g., ADHD) as well as with impaired hearing or a (severely) visual handicap were excluded); (3) participants that could be categorized into one of the following age groups: pre-literate children (3–6 years old), school-aged/literate children (7–11 years old) and young adults (18–35 years old); and (4) reported findings in an English-language, peer-reviewed journal between 1995 and 2020. The earlier year limitation (1995) was implemented to not have a bias toward earlier works, but also to have a clear limitation that helps in keeping the methodological considerations consistent and comparable (i.e., equipment, sample size), whereas the late year limitation (2020) served as a clear cut-off of the search date when the search terms were applied.

Systematic Review Procedure

Information Sources, Search, Data Collection Process

We searched Web of Science, PubMed (MEDLINE), PsychINFO, PubPsych, ProQuest, Scopus, PsycNET, and Cochrane for studies using the following search strings: (N1 OR N170) AND (EEG OR ERP OR event-related potential*) AND (visual OR word OR print) AND (expertise OR read* OR develop*) AND [read* AND (disorder* OR disab* OR dyslexi* OR difficult* OR problem* OR develop*)] AND (participant* OR child* OR adult*).

Final searches were conducted on the 11th and 19th of January 2021. The articles underwent four rounds of screening: removal of duplicates, abstract screening, full-text reading, and data extraction. Removal of duplicates, title, and abstract screening were performed using the Rayyan software for systematic reviews (Ouzzani et al., 2016). The evaluation process was conducted by three independent raters, with title and abstract screening being performed fully blinded.

Risk of Bias in Individual Studies

Included studies underwent a risk-of-bias assessment using the Newcastle—Ottawa Scale (NOS) adapted to cross-sectional studies (Modesti et al., 2016). Each rater judged every study based on seven quality items categorized into three sections: the study group selection (representativeness of the sample, sample size, non-respondents, measurement tool for assessment of reading skill), the comparability of the groups; and the outcome (assessment and statistics). Each rater awarded a star per item if the study fits the criteria. Obtained NOS scores ($M = 7$, $SD = 2$) were reported in **Supplementary Table S1**. Interrater reliability was assessed through percentage agreement of rater1, rater2, and rater3 of the NOS. For this, 10% of the reviewed studies ($n = 7$) were randomly selected and reassessed by the second and third rater. Interrater reliability between each rater pair was 71.24% (R1/R2, R1/R3, R3/R2).

Data Items

The following data were extracted from all selected papers: participant information (e.g., sample size, participant age, reading ability groups), EEG parameters (e.g., pre-processing steps and region/scalp areas of interest as defined by electrode set used in the analyses), stimuli and task characteristics (e.g., language, experimental design), and ERP results (i.e., amplitude, lateralization, latency). We based our ERP summary on the statistical results and the graphical representations present in the text. The full details of the extracted data can be found in **Supplementary Table S2**.

Synthesis of Results

We employed a narrative synthesis to compile the results regarding N170w, amplitude, latency, and lateralization of the selected studies and provided summary tables that included essential extracted features of the study (e.g., participants, age, task, results). In extracting the results for individual studies, we excluded ERP results using other forms of analyses (e.g., topographic analysis of variance, LORETA). The original scope of means and effect sizes extraction of the selected papers had to be reviewed due to the lack of reported means and effect sizes in the papers included in this review. For evaluation of lateralization and amplitude, variables were introduced, which enabled comparison across papers despite the missing mean and effect sizes (i.e., C>DRD, referring to the amplitude of control subjects being enhanced compared to subjects with reading difficulties). Effects of intervention studies on N170w were not assessed; thus, the pre-intervention EEG data only was used for data extraction on N170w for those studies involving training.

RESULTS

Study Selection

The initial database search identified 572 articles. Out of 282 non-duplicates, 146 articles were excluded after title and abstract screening using the Rayyan software for systematic reviews (Ouzzani et al., 2016), leaving a number of 136 articles in the full-text screening. All articles were reviewed by authors K. K. A., A. T., and C. V. with a two out of the three-majority decision for inclusion. Twelve conflicting articles were additionally reviewed by the remaining co-authors, leading to the inclusion of two out of twelve articles. After applying the inclusion and exclusion criteria, 59 articles were excluded during full-text screening and eight articles during data extraction, resulting in 69 articles included in the review. A flowchart of this selection process is displayed in **Figure 1**.

A normal distribution across publication years is significantly noticeable among the included articles (see **Figure 2**). Dense publication years were 2011 (n = 9) and 2013 (n = 8). Specific characteristics of each of the studies can be found in **Supplementary Table S2**.

Methodological Characteristics

Participants

Of the 69 studies included in this systematic review, eight examined the N170w in pre-literate children, 31 in school-aged children, three in adolescents, and 41 in young adults aged between 18 and 35. The total number exceeds 69 studies, as 12 of these included more than one age group. The results of the three studies that examined the N170w in adolescents, are combined with the young adult group, as the mean age of the adolescents (M_{age} = 17.24 years) was close to our lower edge of the young adults age range, and the reported results in terms of amplitude and lateralization were comparable to the results in adults. A substantial number of studies only included typical readers (n = 42), whereas 27 studies compared controls with people with dyslexia only (n = 23) and/or otherwise defined sample (i.e., poor readers or spellers, illiterate or at-risk individuals; n = 8). The number of participants included in each of the studies demonstrated a wide range from 11 to 72. The exact values for each of the reviewed studies together with participant, age, and gender distribution can be obtained *via* **Supplementary Table S2**. Criteria to consider participants as reading impaired or control varied widely across studies. Participants were considered reading impaired based on either a formal dyslexia diagnosis or the evaluation of reading scores below the 25th, 20th, 16th, and 10th percentile; or 1.5 or 1 standard deviation below the average. On the other hand, typical readers had percentile scores above 10 to >25 in reading tests. These lead to discrepancies across studies as DRD and TD readers overlap across studies reporting criteria (n = 23).

Language, Stimuli, and Procedure

Most of the studies were conducted in German-speaking populations (n = 17), followed by English (n = 15), and Chinese (n = 14). A minority of five studies investigated a second language. Paradigm types varied between repetition-detection-task (n = 16), lexical decision task (n = 16), N-back task (n = 6), and other paradigms (n = 31). All 69 studies used words as a condition, and either had it as the only condition (n = 7) or compared words to pseudowords (n = 10), pseudo-homophones (n = 2) or non-words (n = 1). Other comparisons were made to symbols (n = 13), faces (n = 5) or pictures (n = 2). Thirty studies used more than two conditions, mainly comprising words, pseudowords and symbols (n = 20). For a detailed overview of all stimuli per study we refer to **Supplementary Table S2**.

Words presented had an average character length of M = 6.62 (SD = 2.39, 3–13) letters or strokes. When reported, the word frequency of words commonly ranged in high (n = 23) or low to high (n = 10) frequency values.

Stimuli duration of words across studies varied between 100 and 5,250 ms, which differed across participants age groups: adults M = 489.22, SD = 317.49; school-aged children M = 845.77, SD = 724.62; pre-literate children M = 1,125, SD = 1683.96. Paradigm difference in stimulus duration was visible for the bigger clusters of detection tasks (M = 550.31 ms, SD = 460.19) and lexical decision tasks (M = 784.38 ms, SD = 903.09). The explicit word/symbol processing task (5,250 ms) and dual valence task (100 ms) were the most deviating paradigms. The number of presented trials was another dividing factor, ranging from 40 to 576 trials for the word conditions (Brem et al., 2013; Collins et al., 2017). Distance to screen for the word presentation ranged from 50 cm to 145 cm (M = 81.59 cm, SD = 23.82 cm) across studies. Interstimulus intervals (ISI) were composed of different components (e.g., fixation cross and blank screen) across studies. Common feedback, response screens, and blink screens were among the reported procedures for the composition of trials (see **Supplementary Table S2**).

EEG Analysis

The presented studies (N = 69) had a significant difference in the number of EEG channels recorded (19–128, M_{adult} = 64.57, SD_{adult} = 38.67; $M_{litchild}$ = 68.23, $SD_{litchild}$ = 41.27; $M_{prelitchild}$ = 44.13, $SD_{prelitchild}$ = 13.29). Across all studies, most common electrode setups were 64 (n = 17) and 128 (n = 16) electrodes, with one additional study having both setups. Electrodes were reported as Ag/AgCl (n = 53), TiN (n = 7), implemented in caps of various manufacturers (see **Supplementary Table S2**). For EGI systems, the common impedance threshold laid at 50 kΩ; for other systems, it varied between the 5–20 kΩ threshold, with a high distribution across systems and studies in general (5–100, M = 22.40, SD = 22.31). EEG data were recorded at various sampling rates, ranging from 200 to 2,048 Hz. Most studies did not report on downsampling procedures (n = 55); if reported, we recorded values between 256 and 500 Hz. While the reference

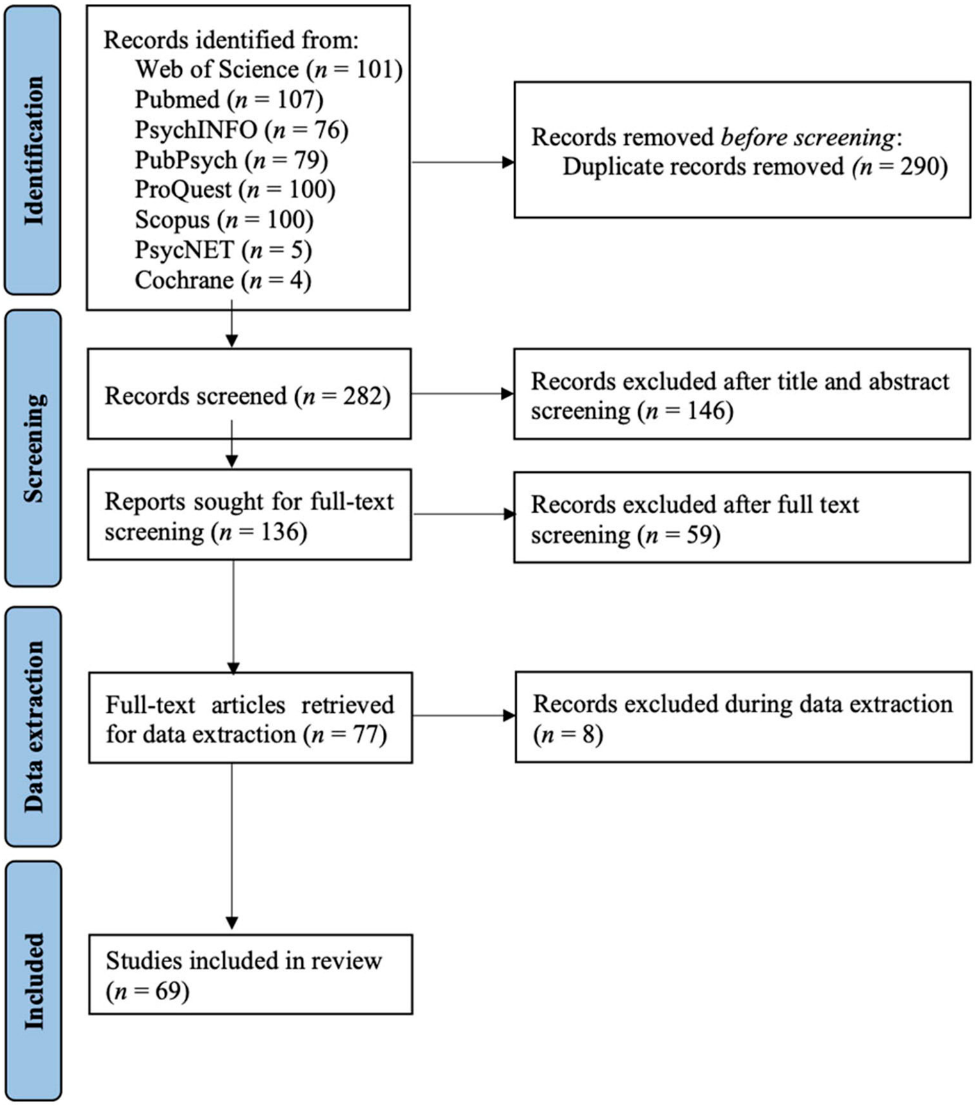


FIGURE 1 | PRISMA flow diagram of the article search, screening, and selection methods. Design adapted from Page et al. (2021).

electrodes used varied across studies (e.g., mastoid, nose tip, Cz, and Biosemi CMS/DRL), re-referencing to the average was a common practice ($n = 50$) as preprocessing step. Other re-referencing methods were reported as Cz, average of mastoids, and multi-electrode referencing (Simon et al., 2007: using 20 out of 32 electrodes, F7, F3, C3, T3, CP3, TP7, T5, P3, F8, F4, C4, T4, CP4, TP8, T6, P4, Fz, Cz, Cpz, and Pz). During recording, common online filtering ranged between 0.1 and 100 Hz. Further, low- (20–48 Hz) and high-pass (0.01–1 Hz) filters were applied. Common baseline windows ranged between 50 and 500 ms pre-stimulus, whereas the most used time frames for baseline were at 100 ms ($n = 28$) and 200 ms ($n = 18$) pre-stimulus onset. A difference between the applied baseline windows was visible between pre-literate and other age groups ($M_{prelit} = -112.5$, $SD = 13.36$, $M_{other} = -154.68$, $SD = 81.50$), possibly related to the small number of papers ($n = 8$) targeting pre-literate population. Independent Component Analysis (ICA) for ocular artifacts and automated artifact rejection with threshold (between ±80 and 125 μV) was commonly reported; if manual rejection was performed, it was commonly performed in combination with another approach. The number of trials included after artifact rejection was sparsely reported.

Regarding the further analysis, the epochs around the target word varied across studies, ranging in the length of the epochs from 250 ms to 2 s, $M = -158.98$ ms (−500–0) to $M = 860.03$ ms (250–1,550). The timeframe in which the peak of N170w was obtained in studies regarding the three age groups differed significantly between adult and child groups (pre-literate children: 175–238.5, $M = 216.56$, $SD = 20.53$; school-aged children: 175–238.5, $M = 215.25$, $SD = 16.43$; adults: 150–270, $M = 183.01$, $SD = 22.29$). These studies have mostly used either global field power (GFP) analyses ($n = 23$), visual peak detection

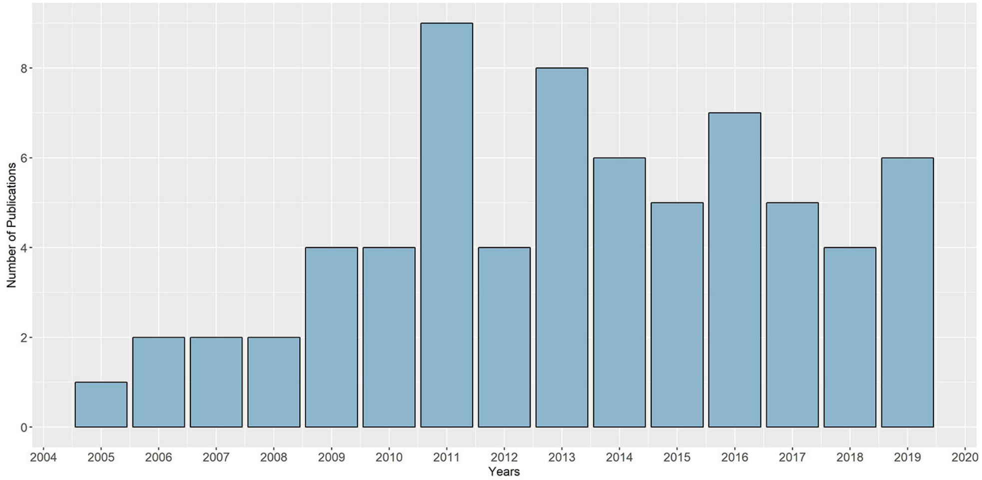


FIGURE 2 | Distribution of included studies across publication years.

(*n* = 14), or literature reviewing (*n* = 7) for selection of the N170w time window. The regions of interest (ROI) examined for N170w have varied across studies, though most studies focused on P7 (*n* = 47), P8 (*n* = 38), and O1 (*n* = 36). N170w amplitudes were obtained using the mean amplitude of the identified ERP time window (*n* = 37) or maximum peak amplitude within the ERP time window (*n* = 21). A lack of reported mean values of the N170w amplitudes to words was observed in most studies, with reliance on the presentation of the mean amplitudes in graphs and ERP waveforms. This form of presentation led to the analysis of N170w amplitude being limited to a qualitative approach of the presented graphs, as also presented statistical results did not include word condition only results.

Statistical Analysis

Forty of the reviewed studies obtained their statistical results by applying analysis of variance (ANOVA). Multivariate analysis of variance (MANOVA) was performed in six studies. Greenhouse-Geisser, Tukey HSD, or Bonferroni corrections were mentioned to be applied by nine studies. Linear models were in the minority, with three applications across studies. Meanwhile, a *t*-test as a lone standing evaluation of N170w specific values was reported by two studies. Between-subject factors across studies were group, age, gender, reading level, hearing level, and others. Within-subject factors mainly consisted of condition and stimuli features and hemispheres/electrode site. Commonly, the study design and statistical computations were not designed to be investigating the N170 response to words alone.

Results of Individual Studies

The full details of the extracted results can be found in **Supplementary Table S3**.

N170w in Typically Developing vs. Developmental Reading Disorder/Poor Readers

Results are reviewed by age group relative to the number of studies that compared different reading ability groups (typically developing: TD, and atypically developing such as developmental reading disorder/poor readers/low reading ability: DRD/PR). Some studies that used the term "Developmental Dyslexia/Dyslexia" are referred to as DRD in this paper. Amplitude, latency, and lateralization comparisons for each age group are displayed in **Tables 1**, **2**.

Amplitude

Forty studies investigated the N170w amplitude in TD and DRD/PR. A total of 29 studies compared the N170w amplitudes between TD and DRD/PR individuals in pre-literate children (*n = 3*), school-aged children (*n = 14*), or young adults (*n = 12*).

In pre-literate children, only three studies investigated the N170w between TD and at risk of DRD/PR. Studies revealed contradictory results, wherein one found larger N170w amplitudes in controls (Li et al., 2013), and two found no amplitude differences between TD and at-risk of DRD/PR groups (Maurer et al., 2007; Brem et al., 2013).

Thirty-one studies explored the N170w in school-aged children, of which 14 compared TD with DRD/PR. Five studies showed a larger N170w amplitude for DRD/PR as compared to controls (Brem et al., 2013; Fraga González et al., 2014, 2016b; Zhao et al., 2014; van Setten et al., 2019), five showed a larger N170w amplitude for controls than DRD/PR (Maurer et al., 2007, 2011; Jucla et al., 2010; Kast et al., 2010; Bakos et al., 2018), and four showed no difference (Araújo et al., 2012; Hasko et al., 2013; Kemény et al., 2018; Pleisch et al., 2019). One specific study further divided the TD and DRD children into young (M_{age} = 8.3) and old (M_{age} = 11.4) sub-groups and found that in younger groups, TD exhibited a more negative N170w

TABLE 1 | N170w Amplitude and Latency results in comparing TD and DRD/PR by age group.

Studies	Amplitude (*N* = 29)				Latency (*N* = 7)			
	C > DRD/PR	C < DRD/PR	C = DRD/PR	Total	C > DRD/PR	C < DRD/PR	C = DRD/PR	Total
Pre-literate	1	0	2	3	–	–	–	0
School-aged	5	5	4	14	1	–	4	5
Young adults	11	1 (right)*	1 (left)*	12	–	2	–	2

Counts in each column refer to the number of studies reporting that result.
**Different sub-groupings in one study (Dujardin et al., 2011).*

TABLE 2 | N170w Lateralization results in comparing TD and DRD/PR by age group.

Studies	Lateralization (*N* = 20)				
	C = left DRD/PR = equal	C = right DRD/PR = equal	C = left DRD/PR = right	C = equal DRD/PR = right	No difference
Pre-literate (*n* = 2)	1	0	0	1	0
School-aged (*n* = 11)	1	3	0	0	4 (bilateral), 2 (left), 1 (right)
Young adults (*n* = 7)	6*	0	1*	0	2 (left)*

Counts in each column refer to the number of studies reporting that result.
**Different sub-groupings in one study (Dujardin et al., 2011: C, left, DRD1, left at trend level, DRD2, bilateral; Mahé et al., 2013: C, left; PR, bilateral, DRD, right at trend level).*

than DRD/PR, whereas the opposite pattern was found for older children (Maurer et al., 2011).

Forty-two studies investigated the N170w amplitude in young adults, of which 12 compared TD and DRD/PR groups. Eleven studies showed that controls exhibited a larger, thus more negative, N170w than DRD/PR subgroups (Savill and Thierry, 2011a,b; Korinth et al., 2012; Mahé et al., 2012, 2013; Waldie et al., 2012; González-Garrido et al., 2014; Korinth and Breznitz, 2014; Araújo et al., 2015; van Setten et al., 2016; Collins et al., 2017). One specific study examined two subgroups of people with DRD based on the inspection of the ERPs; one that exhibited an N170 but no N320 and one with the two waves fused together (Dujardin et al., 2011). The authors found no difference on N170w amplitudes between TD and the first subgroup of DRD over the left hemisphere, but the latter showed more negativity than the former on the right hemisphere electrodes.

Latency

Only 20 out of 69 selected studies explored the latency of the N170w. Thirteen of these provided specific mean latency values. Reported latency results were mainly from the studies comparing different groups (TD vs. DRD/PR or age). Some studies also analyzed the N170w latency values regarding hemispheric distribution (left vs. right) within participant groups.

Eight studies compared the mean N170w latencies of TD and DRD/PR groups. No such studies were conducted in pre-literate TD and at-risk of DRD/PR children. In school-aged children, four studies showed similar mean latencies for both TD and DRD/PR groups (Kast et al., 2010; Maurer et al., 2011; Hasko et al., 2013; Zhao et al., 2014), whereas one study reported that controls had longer mean latencies than DRD (van Setten et al., 2019). In young adults, two studies reported longer mean latencies for DRD than controls (Savill and Thierry, 2011a; Waldie et al., 2012).

van Setten et al. (2016, 2019) were interested in the assumed interaction of mean N170w latency and hemispheric distribution in TD and DRD groups and found a significantly longer mean N170w latency in the right hemisphere compared to the left in both TD and DD/PR groups in young adults and school-aged children.

Lateralization

Out of all the selected 69 studies, 61 investigated the lateralization of the N170w. However, only 20 compared the lateralization between typically developing and reading impaired participants.

In pre-literate children, two studies compared the N170w lateralization between TD and at-risk of PR. Li et al. (2013) reported a left-lateralized N170w for controls, but bilateral activity in at-risk of PR. In contrast, Brem et al. (2013) found bilateral activity for pre-literate controls and a right-dominated N170w, although only at a trend level for pre-literate at-risk of PRs.

Twenty-seven studies investigated the lateralization of the N170w in school-aged children, of which 11 compared TD and DRD/PR groups. Out of the 11 studies, seven studies showed no difference in hemispheric dominance of the N170w between TD and DRD/PR children: four studies reported bilateral activation (Jucla et al., 2010; Hasko et al., 2013; Kemény et al., 2018; Pleisch et al., 2019), two reported left (Maurer et al., 2011; Araújo et al., 2012) and one reported right activation preponderance (van Setten et al., 2019) in both groups. The remaining four studies reported either left (Kast et al., 2010) or right-lateralization (Fraga González et al., 2014, 2016b; Zhao et al., 2014) for controls only, but found bilateral activity in DRD/PR children. To conclude, lateralization in DRD/PR school-aged children was mainly reported to be bilateral ($n = 8$) and control school-aged children appeared to show left, right and bilateral dominance ($n = 11$, $n = 7$, $n = 12$).

Thirty-six studies on N170w lateralization were found in young adults. Seven of these compared TD and DRD/PR groups, of which six studies showed left lateralization of the N170w for the controls and a bilateral activation for DRD/PR groups (Dujardin et al., 2011; Mahé et al., 2012, 2013; González-Garrido et al., 2014; Araújo et al., 2015; Collins et al., 2017). Aside from bilateral activation, the other DRD subtype in Dujardin and colleagues' (2011) study showed left lateralization of the N170w, though at trend level only. Moreover, one study showed left-lateralization for both TD and DRD (van Setten et al., 2016), and another study found left-lateralization for controls but investigated poor readers and adults with DRD separately and found that the former exhibited a bilateral activation of the N170w whereas the latter showed a right-lateralized N170w at trend level (Mahé et al., 2013). These results indicate a clear left-hemispheric distribution for typical reading adults ($n = 33$), with more bilateral distribution occurrences in reading impaired adults ($n = 6$).

N170w From Pre-literate Age to Adulthood

Eighteen studies gave additional insights on the development of N170w amplitude by including different age groups using a cross-sectional or longitudinal design. These studies mainly evaluated control subjects (Maurer et al., 2005, 2006, 2007, 2011; Brem et al., 2006, 2009, 2013; Spironelli and Angrilli, 2009; Van Strien et al., 2009; Cao and Zhang, 2011; Cao et al., 2011; Dundas et al., 2014; Coch and Meade, 2016; Eberhard-Moscicka et al., 2016; Tong et al., 2016a; Curzietti et al., 2017; van Setten et al., 2019; Zhao et al., 2019).

Amplitude

Only two studies compared the N170w across pre-literate age, school-aged, and adulthood in typically developing individuals (Maurer et al., 2006; Eberhard-Moscicka et al., 2016). Eberhard-Moscicka et al. (2016) investigated the development of the N170w in the context of foreign language learning (English). However, the results in this review only included N170w response to the stimuli in the native language, German. Both authors found that N170w amplitudes consistently decreased in adults. However, two studies showed a reversed effect in the children groups, wherein Eberhard-Moscicka et al. (2016) showed a decrease of N170w amplitude from pre-literate to school-children, and Maurer et al. (2006) found the opposite: school-aged children produced a larger N170w amplitude compared to pre-literate children. Two other studies included TD school-aged children, adolescents (M_{age} = 16.2 years), and adults: the adolescents exhibited a larger N170w compared to adults (Brem et al., 2006, 2009) but smaller when compared to school-aged children (Brem et al., 2009).

Eleven studies compared two TD age groups. Maurer et al. (2007) found that pre-literate children exhibited smaller N170w amplitudes than school-aged children. Five studies compared TD school-aged children and adults (Spironelli and Angrilli, 2009; Cao and Zhang, 2011; Cao et al., 2011; Coch and Meade, 2016; van Setten et al., 2019), whereas one study compared pre-literate children and TD adults (Maurer et al., 2005). All found similar results, i.e., larger N170w amplitudes in children compared to adults.

In addition, four papers compared young (M_{age}= 8) and old school-aged ($M_{age} = 11$) children and collectively corroborated the finding of Maurer et al. (2011), i.e., larger N170w amplitudes in younger children compared to the older group (Van Strien et al., 2009; Cao et al., 2011; Tong et al., 2016a; Zhao et al., 2019). One study divided the adults into young (20–30 years old) and old (>40 years old) groups, wherein the latter exhibited a larger N170w than the former (Curzietti et al., 2017). Lastly, one study compared gender differences, with boys showing larger N170w amplitude than girls (Spironelli et al., 2010).

Latency

Six studies compared the mean latencies of two or three age groups. Five of these showed that the N170w peaked earlier in adults than in pre-literate children (Maurer et al., 2005), school-aged children (Brem et al., 2009; Cao and Zhang, 2011; Cao et al., 2011), and adolescents (Brem et al., 2006). Only one study showed similar mean latencies in school-aged children and adults (Coch and Meade, 2016).

Five studies investigated the interaction of mean latency and hemispheric distribution of the N170w. Three studies were conducted on pre-literate children and revealed opposite results. (Zhao et al., 2015) reported in their training study that the N170w occurred later over the right than the left hemisphere for the visual learning group (visual identification of characters); however, they saw a reversed pattern in the writing condition group (manual tracing and copying of characters) at the pre-test phase before training. The same research group (Zhao et al., 2018) found, according to their earlier finding, that the N170w latency was only slightly delayed in the right hemisphere compared to the left (Zhao et al., 2018), and another study did not find any latency differences between the hemispheres (Maurer et al., 2005). To examine whether the reported latencies across studies differed significantly between the hemispheres, we conducted a two-tailed t-test, which did not reveal significant differences across the three studies presented for pre-literate children ($M_{left} = 215.5$, $SD_{left} = 7.5$; $M_{right} = 217.2$, $SD_{right} = 5.1$). Two studies divided their school-aged sample into a young and old subgroup (Maurer et al., 2011; Tong et al., 2016a). Maurer et al. (2011) found a longer latency for the younger children compared to the older ones, whereas Tong et al. (2016a) reported the opposite pattern. However, school-aged children generally showed nearly no differences in the mean N170w latencies between the left and right hemispheres ($M_{left} = 214.7$ ms, $M_{right} = 215.5$ ms). For young adults, controversial latency values have been reported, with a longer mean latency of the N170w over the right hemisphere than the left in one study (van Setten et al., 2016), and the opposite was observed in another one (Xue et al., 2019). Compared to pre-literate and school-aged children, the N170w occurred earlier in adults ($M_{left} = 161.2$ ms, $M_{right} = 161.9$ ms).

Lateralization

Nine studies compared the N170w lateralization across different age groups. Two studies found that pre-literate and school-aged children exhibited bilateral N170w, but this N170w became left-lateralized in adulthood (Maurer et al., 2006, 2007). However, one finding showed that pre-literate children exhibited right-lateralized N170w, which became left-lateralized in adulthood (Maurer et al., 2005). Other findings either showed a left-lateralized N170w (Cao and Zhang, 2011; Cao et al., 2011; Dundas et al., 2014), a bilateral distribution (Mercure et al., 2009), or a right-lateralized N170w (Spironelli and Angrilli, 2009) in childhood that became left-lateralized in adulthood (Mercure et al., 2009; Spironelli and Angrilli, 2009; Cao and Zhang, 2011; Dundas et al., 2014). Lastly, one study investigated school-aged children, adolescents, and adults and found no differences in the lateralization across the age groups measured (Brem et al., 2009).

Many studies investigated lateralization of the N170w in one age group only. For pre-literate children, bilateral activity was reported in two studies of the same research group (Zhou et al., 2009; Zhao et al., 2015). Nine studies investigated the N170w lateralization in school-aged children. Five reported left-lateralization of the N170w (Van Strien et al., 2009; Cao et al., 2011; Su et al., 2015; Sacchi and Laszlo, 2016; Bakos et al., 2018), and two reported no difference between the responses recorded over the two hemispheres (Eberhard-Moscicka et al., 2015; Tong et al., 2016a). Another study on TD school-aged children compared two types of orthographic scripts, alphabetic and logographic, and found a right-lateralized N170w for the former and a bilateral distribution for the latter (Tong et al., 2016b). Lastly, one study compared gender differences in TD school-aged children, wherein girls exhibited right dominance and boys showed bilateral activity (Spironelli et al., 2010). Twenty-two studies investigated TD young adults only, and the vast majority ($n = 20$) of these studies showed significant left-lateralization (Brem et al., 2006; Simon et al., 2007; Maurer et al., 2008a,b; Lin et al., 2011; Mercure et al., 2011; Yum et al., 2011; Korinth et al., 2012; Taha and Khateb, 2013; Taha et al., 2013; Okumura et al., 2015; Curzietti et al., 2017; Emmorey et al., 2017; Uno et al., 2017; Yang et al., 2017; Wei et al., 2018; Davis et al., 2019; Faísca et al., 2019; Xue et al., 2019) or at trend level (Maurer et al., 2010). In contrast, two studies reported either bilateral (Okumura et al., 2014) or right lateralization (Cao et al., 2013). An overviewof the lateralization results is displayed in **Table 3**.

TABLE 3 | N170w Lateralization results in comparing age groups in studies that only included a TD group.

Studies	Lateralization ($N = 33$)		
	Left > Right	Left < Right	Left = Right
Pre-literate ($n = 2$)	0	0	2
School-aged ($n = 9$)	5	2*,**	4*,**
Young Adults ($n = 22$)	20	1	1

Counts in each column refer to the number of studies reporting that result.

**Different scripts: Tong et al., 2016b- alphabetic: right-lateralized N170w and logographic: bilateral distribution.*

***Different gender: Spironelli et al., 2010 – girls: right-lateralized, boys: bilateral distribution.*

N170w vs. Word-Like Conditions

To investigate the N170 response related to early lexical effects, we included studies that, aside from words, used word-like stimuli. These word-like conditions consisted of pseudowords (PW, resembles the orthographic and phonological structure of a real word, thus pronounceable), pseudo-homophones (PH, sounds like a real word but incorrectly spelled), and non-words (NW, orthographically or phonologically illegal letter strings that are not pronounceable, excluding symbols and false fonts).

Amplitude

Within-subject manipulations of word-like stimuli such as comparing words vs. pseudowords, pseudo-homophones, or non-words were examined in 21 studies, mostly in typically developing individuals, and eight of these compared TD and DRD/PR groups. Three studies compared two different age groups, i.e., school-aged children and adults. Furthermore, one study investigated these word and word-like conditions in pre-literate children only, seven studies involved exclusively school-aged children and ten studies included only young adults.

Mean amplitudes between word and word-like comparisons in a Chinese study on pre-literate children revealed a higher amplitude for line and character conditions compared to the stroke (re-arrangement of stroke combinations in a radical) and radical (non-character stimulus due to illegal position of radicals) conditions in the left hemisphere (Zhao et al., 2018). The left hemisphere showed an overall, more robust N170 response. In school-aged children, results showed either a more negative N170 response for pseudowords in TD compared to DRD children (Kast et al., 2010) or no difference in TD vs. DRD (Hasko et al., 2013), particularly in the right hemisphere (Zhao et al., 2014). Zhao et al. (2014) showed that the processing of words, pseudowords, and non-words in the left hemisphere varied across TD and PR children. More specifically, the N170 responses to words were more negative than non-words in TD, whereas no difference was found for PR children. In addition, the responses elicited by words were more negative than to pseudowords in PR children, whereas no difference could be found in TD. Moreover, the responses on pseudoword vs. non-word comparisons on the left hemisphere showed a trend level in TD children, with pseudowords showing a more negative N170 than non-words, but no difference between pseudowords and non-words was found for PR children.

For school-aged children, most studies revealed no differences in the N170 response between word and pseudoword conditions (Eberhard-Moscicka et al., 2015; Tong et al., 2016a; Zhao et al., 2019). However, Zhao et al. (2019) corroborated these findings for the right hemisphere only for their older subgroup in the same study, but they found a more negative N170 amplitude for pseudowords compared to words in the left hemisphere. In comparing words and non-words, one study showed no difference (Pleisch et al., 2019), and another study found a more negative N170 response for words than non-words in the left hemisphere (Tong et al., 2016a).

In six studies on TD young adults, five studies found no difference between words, pseudowords, pseudo-characters, or non-words (Simon et al., 2007; Lin et al., 2011; Okumura et al., 2014; Wei et al., 2018) and one study showed a larger N170 in pseudo-homophones than words (Taha and Khateb, 2013). These results seemed to be moderated by the task design, as these results were confirmed in an explicit task but showed different results in an implicit task, i.e., a larger N170 for words than pseudowords (Faísca et al., 2019). In comparing TD and DRD/PR young adult groups, results showed diversity, wherein (a) words showed less negative N170 than pseudo-homophones and pseudowords in both TD and DRD adults (Araújo et al., 2015); (b) found no difference between words and non-words in the TD group but a more negative N170 to non-words than to words in DRD (Waldie et al., 2012); (c) recorded more negative N170 to words than to non-words and pseudowords in the TD group, but found no difference in the N170 between words and pseudowords in the DRD group over the left hemisphere (Mahé et al., 2012, 2013); or (d) more negative N170 was recorded over the left hemisphere to pseudowords in the TD group, with an opposite result for the DRD group (Dujardin et al., 2011).

Lastly, while Cao and Zhang (2011) found no difference between word and word-like conditions between school-aged children and adults, two studies reported more negative N170w than by pseudowords in TD adults, with no difference for school-aged children (Coch and Meade, 2016; Eberhard-Moscicka et al., 2016). Moreover, one study comparing the N170 to pseudowords and non-words showed no difference in school-aged and young adult groups (Coch and Meade, 2016).

Latency

Only six studies investigated the latency differences comparing the N170 recorded in word and word-like conditions. Coch and Meade (2016) compared the N170 latencies of pseudowords and non-words in school-aged children and adults and found longer latency of the N170 responses in children in both conditions compared to adults. In pre-literate children, the N170 for Chinese characters occurred earlier over the left hemisphere than for radical and stroke combinations, whereas in the right hemisphere, the N170 for radical combinations occurred first, followed by that for stroke and character combinations (Zhao et al., 2018). In school-aged children, Zhao et al. (2014) compared TD and PR and found that for the TD group, non-words elicited longer latency responses in both left and right hemispheres than in the PR group. However, the N170 latency to pseudowords and words differed based on the hemisphere, a later response to pseudowords than words over the left hemisphere, and a faster appearing response over the right in both TD and PR groups.

Meanwhile, pseudowords showed the longest latency for the PR group, followed by words and non-words in the left hemisphere, whereas in the right hemisphere, the PR group showed similar results with the TD group, i.e., shortest latency for pseudowords followed by words and non-words. Hasko et al. (2013) found that the N170 for pseudo-homophones and pseudowords had a shorter latency than for words, and this result did not differ between the TD and DRD children. However, Bakos et al. (2018) reported the opposite; the N170 for words exhibited a shorter latency than for pseudo-homophones, but again no differences were found between the TD and DRD groups. Coch and Meade (2016) found that while the N170 latencies to words and pseudowords did not differ in typically developing 3rd and 5th graders, pseudowords elicited longer latency responses than words in 4th graders. Lastly, studies in young adults showed no difference between the N170 response latencies to words and pseudo-homophones (Taha and Khateb, 2013), words, and pseudowords (Coch and Meade, 2016) or pseudowords and non-words (Coch and Meade, 2016).

Lateralization

Eighteen studies investigated the lateralization of the N170 response in word-like conditions. In a study comparing different reading ability groups of school-aged children, Hasko et al. (2013) found no differences between TD and DRD groups, as both groups showed a bilateral distribution of the N170 to words and non-words. Similar bilateral response distribution was reported for pseudowords in TD children (Jucla et al., 2010; Hasko et al., 2013) and in DRD children (Jucla et al., 2010; Kast et al., 2010). In studies on young adults, Araújo et al. (2015) found no differences between TD and DRD groups, wherein pseudo-homophones generated a larger N170 than words and pseudowords over the right hemisphere but showed no differences in the N170 between the word and word-like conditions over the left hemisphere. Meanwhile, the N170 response to pseudowords showed a left-lateralized distribution trend in TD adults (Mahé et al., 2012, 2013) but showed the reverse for DRD adults (Mahé et al., 2012).

The remaining studies only reported lateralization on TD readers in one age group: one study in pre-literate children, two in school-aged children, and nine in young adults. For pre-literate children, the N170 response to pseudowords showed bilateral distribution, similar to the word condition in pre-literate (Zhao et al., 2018) and school-aged children (Eberhard-Moscicka et al., 2015). However, non-words showed a left-lateralized distribution of the N170 (Pleisch et al., 2019). In young adults, most of the studies showed left-lateralization of the N170 for pseudowords or pseudo-characters (Simon et al., 2007; Cao and Zhang, 2011; Lin et al., 2011; Wei et al., 2018), pseudo-homophones (Taha and Khateb, 2013) and non-words (Okumura et al., 2015; Uno et al., 2017), whereas one showed bilateral activation of the N170 for non-words (Okumura et al., 2014). In addition, Faísca et al. (2019) compared words versus pseudowords and found that while the N170 responses to words were more negative over the left recording sites than to pseudowords, no difference in the N170 was present over the right (Faísca et al., 2019). Meanwhile, other studies showed no difference in the N170 between words and pseudowords over the left hemisphere (Simon et al., 2007; Wei et al., 2018).

DISCUSSION

This work is the first systematic review that assimilated existing research on the N170 response to words (N170w) in individuals with and without reading difficulties. Out of seven databases, 69 peer-reviewed studies were included, of which the majority was conducted in adults, followed by school-aged children and pre-readers, mainly in German-, English- or Chinese-speaking populations. Our main goal was to synthesize findings on the differences in the N170w amplitude, latency, and lateralization and to capture the typical and atypical development of the N170w by comparing different age groups, namely, pre-literate children (3–6 years old), school-aged/literate children (7–11 years old) and young adults (18–35 years old). Aside from this, we aimed to shed light on the assumed fine-tuning of the emerging print expertise shown by the N170 by comparing the N170w with those recorded in word-like conditions across different studies. Lastly, we compared the ERP methods used across studies. Here, our main objective was to provide an overview of various paradigms and recording systems used in N170 research in reading.

Comparison of N170w in Typical and Atypical Readers

Results on the N170w amplitude illustrate that most TD readers, particularly adults, show a larger, more negative N170w than impaired readers. This larger and more negative N170w in typical readers can be explained by a more efficient visual orthographic processing, e.g., expertise in print. It has been interpreted in the reviewed literature to indicate effective orthographic processing during the prelexical stage (Simon et al., 2007; Dujardin et al., 2011; González-Garrido et al., 2014) as well as an efficient specialization enhanced by exposure to print and successful reading acquisition *via* efficient learning and conversion of letter-sound correspondences (Brem et al., 2013; Zhao et al., 2015). Here, poorer reading performance of the DRD/PR groups compared to TD has been interpreted as a consequence of insufficient visual tuning or identification of print or word forms, which continued as persistent weakness in adulthood. Based on the reviewed literature, this can imply a slower, inconsistent orthographic processing (Savill and Thierry, 2011a; Waldie et al., 2012), lower general reactivity to orthographic stimuli (Maurer et al., 2005; Savill and Thierry, 2011a), impairment of visual plasticity exclusive to print at the beginning of reading acquisition (Maurer et al., 2007), deficient processing in visual recognition cortical areas (Kast et al., 2010) or unconventional specialization of the responsible brain networks (Mahé et al., 2012, 2013). Different cognitive domains have also been suggested as responsible modulators of the orthographic specialization reflected by the N170w. One is the inefficient attention allocation system, as shown by the P1 ERP component, suggesting the importance of domain-general functions related to visual processing (Korinth et al., 2012; Korinth and Breznitz, 2014).

The N170w latency findings suggest similar processing time, e.g., similar latencies in TD and DRD groups, or longer latency in controls compared to DRD in childhood. However, in adulthood, findings consistently report a longer latency for DRD than controls. This result is interpreted in the literature as less efficient processing of orthographic cues in dyslexic participants (Savill and Thierry, 2011a). Moreover, this delayed processing of words may be associated with neurobiological deviations reflected in the electrophysiological correlates, here the N170w, in impaired readers. Differences in the structural connectivity, atypical hemispheric asymmetry, or processing differences shown by EEG band power and coherence could possibly show these assumed neurobiological differences (e.g., Arns et al., 2007; Dhar et al., 2010; Fraga González et al., 2016a, 2018). In Waldie et al. (2012) study, event-related brain potentials and EEG coherence, measuring the neural synchrony, were investigated in late-proficient bilingual, dyslexic, and control adult participants performing a lexical task. While higher synchrony was found between hemispheres in the gamma range in the dyslexic group, the same was found in the theta range compared to the control group. In addition, the higher between-hemisphere synchrony was accompanied by lower amplitude N170w in the dyslexic group. The authors interpreted their findings as an asynchrony of neuronal activity at the crucial moment when word forms need to be distinguished. However, the EEG/MEG connectivity studies available on TD-DRD comparisons yield inconclusive results and should further be examined in future studies.

Specialization of print, part of the reading and language network (McCandliss et al., 2003; Dehaene et al., 2010), is typically reported to be left-lateralized in typical readers and bilateral in impaired readers. This left lateralization is thought to be driven by phonological processing, referred to as the phonological mapping hypothesis suggested by Maurer and McCandliss (2008). Evidence for this phonological mapping hypothesis has been specially found for languages that use grapheme-phoneme conversion rules (i.e., alphabetic languages) but has been challenged in studies that used logographic or syllabic languages using lexical morphemes (Maurer and McCandliss, 2008; more discussion, see Linguistic Factors Section). The core idea of this hypothesis was that print processing in the visual cortex underwent left lateralization due to the left-lateralized phonological processing (Price et al., 1997; Rumsey et al., 1997). Although beyond the scope of this review, this early theory might correspond to findings on a left-lateralized hemodynamic activity during visual word recognition tasks (Brem et al., 2009; Maurer et al., 2011; Pleisch et al., 2019). Several neuroimaging studies identified the left ventral occipitotemporal cortex, referred to as the Visual Word Form area (VWFA), as a critical structure for fluent and efficient word recognition (Cohen et al., 2000; McCandliss et al., 2003; Dehaene et al., 2010; Coch and Meade, 2016). This argument is beyond the scope of this review, however for more discussion, we refer to Cohen et al. (2000) and Chen et al. (2019).

The preferential activation to print in the left ventral occipitotemporal cortex has been attributed to successful grapheme-phoneme learning and mapping when formal reading instruction begins (Brem et al., 2010; Karipidis et al., 2017; Pleisch et al., 2019), as well as visual or script familiarity (Brem et al., 2013) or higher word knowledge in pre-literate children, thus highlighting the key role of reading exposure (Li et al., 2013). However, for school-aged children, some studies have reported a more right-lateralized N170w (or reduced left-lateralization of N170w) for typical readers (Fraga González et al., 2014, 2016b; Zhao et al., 2014), which may contradict the general assumption that successful grapheme-phoneme correspondences indicate left-lateralization. The authors indicated that this reduced left-lateralization of N170w for typical readers (Fraga González et al., 2014, 2016b; Zhao et al., 2014) could be due to (a) specialization of the visual word form area, implying a successful lexical access and whole-word level specialization (Fraga González et al., 2014), (b) more automatized reading in typical readers (Maurer et al., 2006; Fraga González et al., 2014) or (c) employment of attentional strategies in orthographic word decoding than processing phonology or semantic information (Fraga González et al., 2014, 2016b). This slight right-lateralization was also reported for pre-literate children that were later classified as poor readers, which can be attributed but not limited to visual familiarity to letters (Brem et al., 2013). This result is due to the non-reading preschoolers but with high letter knowledge as their sample, indicating that exposure may have helped it develop even before reading instruction starts (Maurer et al., 2005; Brem et al., 2013). Aside from this, Brem et al. (2010) found that in pre-literate children with eventual poor reading outcomes, this right-lateralized negativity can be attributed to possible differences in print processing strategies which can be modulated by attention (i.e., focusing on whole-word associations strategy than using letter-sound correspondences). Lastly, some studies on TD school-aged children also reported bilateral activation, which was interpreted as a delayed or missing automatization or an incomplete development of print sensitivity (Hasko et al., 2013). Likewise, this bilateral activation was found in impaired readers, in pre-literate age, childhood, and adulthood, which might indicate a failure to exhibit automatic grapheme-phoneme conversion needed for skilled reading, and is typically mastered through increased exposure to print and continuous reading experience (Brem et al., 2010; Karipidis et al., 2017; Pleisch et al., 2019).

Development of the N170w

Generally, results indicate an amplitude decrease of the N170w with age; thus, less negative N170w amplitudes have been reported in adults than children. This amplitude decrease has been suggested to be related to more reading experience (Brem et al., 2010; Karipidis et al., 2017; Pleisch et al., 2019) and fluency gains (Fraga González et al., 2016b). Studies that looked into young and old subgroups within school-aged children have consistently found that larger N170w was elicited in younger groups than in their older counterparts, reflecting a higher print tuning in the early phase of reading acquisition (Maurer et al., 2007). This developmental course across studies adheres to the suggested inverted U-shaped development of print tuning (Maurer et al., 2006; Fraga González et al., 2014; Pleisch et al., 2019) found as evident in the three age groups included in our review. Pre-literates showed low N170w amplitude due to non-exposure (Maurer et al., 2006), which increased upon the start of reading instruction mainly during the first two years of learning to read and then leveled off around the second to fifth grade (Maurer et al., 2011) continuing to decrease to adulthood as a result of increased exposure with a consequence of enhanced print expertise (Fraga González et al., 2014). Maurer et al. (2006) argued that such plastic reorganization of the brain for print could not be due to general maturation, as this would lead to an increased N170 for both words and matched symbols. N170w latencies showed a characteristic developmental trajectory, with adults having an earlier onset than school-aged and pre-literate children. This result was interpreted to reflect automatization after becoming an expert reader (Maurer et al., 2006). Lastly, the N170w of typical readers changed its bilateral distribution to a dominant left-hemispheric presence throughout development, whereas this response was mainly right-lateralized for younger poor readers and continuously remained in a bilateral distribution. This N170w left-lateralization throughout development can be attributed to the synchrony of orthographic and phonological correspondences as reading expertise improves, indicating a word reading automaticity (Maurer and McCandliss, 2008).

Word vs. Word-Like Conditions

Results looking into the differences in amplitude, latency, and lateralization of the N170 elicited in word, and word-like conditions (pseudowords, pseudo-homophones, non-words) report huge variability across studies comparing TD and DRD/PR groups. These comparisons investigated how lexicality effects might be involved in the processes giving rise to the N170 component. Some studies referred to the changes found as the result of fine-tuning, which involved different processing for words compared to word-like stimuli. These changes could possibly be influenced by early lexical activation (Mahé et al., 2013), usually occurring at the late interval of the N1 ERP component (Eberhard-Moscicka et al., 2016). Unlike the early maturation for print upon reading instruction in children (Maurer et al., 2007; Brem et al., 2013; Eberhard-Moscicka et al., 2015), the emergence of selective responses to word forms most likely follows a prolonged development since it would require higher reading abilities to delineate different types of word forms (Centanni et al., 2018; Pleisch et al., 2019). In this case, this could partly explain different results across studies described below when comparing word and word-like conditions.

Differences in the fine-tuning are thought to relate to processes reflected by the later N1 associated with orthographic regularity or pronunciability, thus expecting pseudowords to elicit a larger N170 than non-words for typically developing children (Zhao et al., 2014). Failure for DRD/PR individuals to catch these pseudoword-non-word differences can be attributed to impaired sublexical orthographic processing, which may entail less sensitivity to letter positioning and sequences (Araújo et al., 2015). Furthermore, differences between words and pseudowords are brought by non-automatized grapheme-phoneme mapping, thus reflecting non-generalization of N170w specialization to pseudowords (Maurer and McCandliss, 2008). Some studies showed a right-lateralized lexicality effect in the DRD/PR group, indicating a right hemisphere overactivation typical for the DRD population, which is negatively correlated with reading skills (Shaywitz et al., 1998, 2002; Mahé et al., 2012). However, aside from these possible scenarios, other factors that could explain the variations found in results across studies can be attributed to a lack of lexical access (Korinth et al., 2012), variations in stimulus material, i.e., linguistic and non-linguistic stimuli (Barber and Kutas, 2006; Hasko et al., 2013), task design and demands, i.e., implicit and explicit tasks (Bentin et al., 1999; Faísca et al., 2019), limited reading experience (Kast et al., 2010; Hasko et al., 2013) or linguistic variables (Bentin et al., 1999; Pegado et al., 2014; further discussed in Section Linguistic factors).

With regards to lateralization in TD readers, words and pseudowords showed a similar bilateral distribution of the N170 in pre-literate age and childhood, which could be affected by the degree as to which reading stage they were in, wherein this case, these children might not have enough exposure yet to have developed print expertise or automatized grapheme-phoneme mapping (Brem et al., 2010; Karipidis et al., 2017; Pleisch et al., 2019). An alternative explanation would be the task demands or linguistic variables (Eberhard-Moscicka et al., 2015). Most studies using word-like conditions in TD adults showed left-lateralized responses to pseudowords, arguing that print specialization generalizes from words to well-ordered letter strings (Maurer and McCandliss, 2008; Dujardin et al., 2011). For alphabetic scripts, this is probably the result of the recruitment of phonology in the successful activation of grapheme-phoneme mapping, known as the phonological mapping hypothesis (Maurer and McCandliss, 2008), whereas, for logographic scripts, it must be primarily based on orthographic processing rather than phonology due to arbitrary sound-graphic correspondences or the reliance on morpheme structures (Zhou et al., 2009; Lin et al., 2011). However, the interpretation for left-lateralization of pseudowords should be taken into caution depending on the orthographic depth of the language involved as those with inconsistent grapheme to phoneme mappings (i.e., opaque orthography) complicates the automaticity and might therefore lead to non-left-lateralization (Maurer and McCandliss, 2008).

Linguistic Factors

It has been long contested whether alphabetic vs. logographic languages are processed differently in the brain, and a few studies tried to investigate this in relation to the N170 (Wong et al., 2005; Maurer et al., 2008b; Cao et al., 2011; Lin et al., 2011; Qin et al., 2016). The authors aimed to answer whether the type of orthographic script modulates the lateralization of print specialization or whether it is entirely dependent on script familiarity of the participants involved in the experiments. From our systematic search, studies focusing on alphabetic scripts, mainly Latin scripts, mostly found evidence for a left-lateralized N170w (e.g., Maurer et al., 2005; Dujardin et al., 2011; Mahé et al., 2012, 2013; Dundas et al., 2014; González-Garrido et al., 2014; Araújo et al., 2015; Collins et al., 2017). Studies conducted in Chinese or Japanese found that logographic or syllabic scripts, revealed a left-lateralized N170 response to their characters as well (Maurer et al., 2008b; Cao et al., 2011; Yum et al., 2011; Qin et al., 2016; Wei et al., 2018; Xue et al., 2019). These findings collectively suggest that left-lateralization develops through reading expertise and visual form familiarity even in languages without direct grapheme-phoneme mapping.

Aside from the type of the scripts, the orthographic depth of a language (i.e., the consistency in which spelling is mapped onto sounds; Schmalz et al., 2015) has also been investigated as a potential modulator for the N170w, with transparent languages having more direct correspondences and opaque languages having less direct correspondences. According to the classification of European languages of Seymour et al. (2003) and the included papers using Asian languages, the reviewed studies showed some orthographic diversity with 52% of the sample investigating deep orthography languages (e.g., English, French, Chinese), 33% including shallow orthographies (e.g., German, Japanese) and the remaining 15% including languages that were medium transparent (e.g., Dutch, Portuguese, Arabic). Most studies in deep orthography showed a more negative N170w for controls than their reading-impaired counterparts (in English: Savill and Thierry, 2011a,b; Waldie et al., 2012; Collins et al., 2017; in French: Jucla et al., 2010; Mahé et al., 2012, 2013; in Hebrew: Korinth and Breznitz, 2014), whereas those categorized in the middle showed the opposite; a larger N170w in impaired readers compared to controls (Fraga González et al., 2014, 2016b; van Setten et al., 2019). Split results have also been noted in studies with shallow orthography depending on the age group. A more detailed investigation of this is beyond the scope of this review. However, it could pose another question for future studies on how N170w is affected by lexical vs. non-lexical reading routes, hence an area of further exploration regarding the differences between coarse and fine-tuning of N170 across different orthographies.

Methodological Considerations

The studies discussed in the current review have used a variety of experimental designs, yielding many different variables across experiments. Before the measured raw signal can be analyzed, it has to undergo a series of preprocessing operations, such as re-referencing, offline filtering, correcting or rejecting artifacts, which might in themselves influence ERP outcomes. Moreover, included studies have shown a considerable disparity on how and where ERPs are quantified. Most studies have used GFP or previous literature to determine the N170w time window, using the mean activity within this time window or detecting the most extreme amplitude value. Despite Picton et al. (2000) giving guidelines for reporting the results of ERP studies, a significant amount of our reviewed studies did not achieve a holistic reporting of all essential aspects necessary to compare the methodology of ERP research. Studies mostly focused on P7, P8, and O1 electrode sites but sometimes were even spread further for temporal sites such as T7 and T8. Moreover, a significant variety was found across studies in using single electrode activity or mean activity within an electrode cluster. Despite the choice of cluster or single electrode analyses, the recording location of the electrodes should be considered, as it diversifies the findings within/across studies as the scalp position of the electrode might differ from the template position. Digitization of electrode position or clear deviation description is advisable (Picton et al., 2000). Furthermore, the diversity of the EEG acquisition systems calls for an evaluation of the effect of the acquisition system as a contributing variable for N170w amplitude. More care should be given to the amplifier specifications and online filtering of the recordings. Even in offline filtering, the filter choice should be described in detail (i.e., backward, forward, zero phases), and the same is valid for the high pass and low pass values. Most articles delivered the voltage level applied for thresholds used in artifact rejections and stated other artifact rejection methods, though, were inconsistent in reporting the final number of trials used for averaging. The final number of trials per condition is an essential factor to report, even to evaluate the amplitude measure (peak, mean over time window) used to extract the N170w value (Picton et al., 2000). Most of the reviewed studies gave an illustration of ERP waveforms, although the labeling of graphs did not follow a convention. Therefore, it is advisable to pay attention to explicit labeling. Crucial to be aligned across studies is the reporting of statistical tests, their outcomes, and effect sizes, especially descriptive values of the N170w amplitude and latency were lacking in many studies. It should be considered as good practice to provide a satisfactory amount of statistical information in order to help the reader to understand the full scope of calculations, as well as enabling to compute, e.g., effect sizes if wanted. For a more comprehensive guide on statistical reporting in brain research, see Gross et al. (2013). For more comprehensive discussion and guidance on ANOVA application in ERP research, we refer to Dien (2017).

Substantial heterogeneity was found across experimental paradigms. Most studies used a repetition-detection task or a lexical decision task, with a combination of words, pseudowords, and symbols to examine the N170w. Different outcomes of the N170w, particularly its left-lateralization, among age and reading-ability groups might vary due to external factors such as stimuli and tasks/experimental designs (Maurer et al., 2007; van Setten et al., 2019). Faísca et al. (2019) mentioned that while early lexical effects (more negative N170w than pseudowords) were evident in implicit tasks (e.g., one-back repetition task) on adults, this result could not be replicated in explicit tasks (e.g., delayed reading aloud). In contrast, Maurer et al. (2005) noted that explicit linguistic tasks showed more sensitivity to lexical differences than implicit tasks; however, the study also highlighted exceptions, even though scarce, in which linguistic characteristics of the stimuli can affect N170w. Regardless of the two different points of view, the N170w can be seen as dependent on task demands representing a difference in susceptibility to top-down processes based on the tasks' goal, especially when investigating the fine-tuning component, which is related to early lexicality effects (Faísca et al., 2019). It is perceived that tasks that require low-level visual recognition (e.g., repetition detection tasks) may have elicited a much more automatic reading for words than pseudowords, whereas, for conscious linguistic tasks, a focus on the grapheme to phoneme decoding may have taken place (Maurer et al., 2005; Eberhard-Moscicka et al., 2016; Faísca et al., 2019). Moreover, left-lateralization of N170w is linked as well with the type of processing strategies or attentional allocation during the early phase of reading acquisition (Maurer et al., 2010; Faísca et al., 2019), but as reading expertise is enhanced with age, the left-lateralization becomes automatic and less susceptible to the attention and task demands (Strijkers et al., 2011; Faísca et al., 2019).

This flexibility in experimental designs and data analysis is a common target of criticism, as it inflates the chance of false positives and complicates the comparison of findings across studies. Several authors attempted to provide publication guidelines to facilitate methodological transparency (Picton et al., 2000; Keil et al., 2014; Clayson et al., 2019; Paul et al., 2021). It remains an educational process for researchers, as such information is crucial to assess the quality of research and ensure that enough information is available to undertake replication studies. One way to overcome these issues is study preregistration, a locked plan containing a study hypothesis, methodology, and data analysis plan (Paul et al., 2021). Enhancing the use of pre-registration in interaction with the common alignment of EEG data analyses as an approach to overcome the reproducibility and comparability of EEG analyses could be considered.

Limitations

Studies on pre-literate children in this review are generally scarcer than the other age groups. The reported results for pre-literate children are taken from three studies only, thus giving limited power in driving solid conclusions due to the limited sample size. Aside from this, different EEG acquisition systems, pre-processing steps, and experimental designs have been utilized in the studies included in the review, offering substantial heterogeneity. Variations in the EEG preprocessing steps, such as different filtering values, can also affect direct comparisons.

Reporting of results seemed incomplete as most of the papers did not report mean amplitude or latency values; hence, no claims can be made about the contribution of the proposed moderators. Our review synthesizes the results of the included studies qualitatively with the inclusion of descriptive statistics for some variables, as it was not possible to obtain enough effect sizes for the computation of funnel plots from the given data of the papers included in this systematic review. Most studies did not provide complete statistical information to calculate these effect sizes in their manuscripts. Thus, an argument can be made that this systematic review represents literature with a publication bias, as we only included peer-reviewed studies and did not access gray literature. To address these limitations, a meta-analysis is highly called upon to provide a more comprehensive picture of the N170w.

Due to our strictly focused search on N1/N170 in words and reading disorders in children or adults, it is unavoidable that we may have missed relevant studies on N170 (e.g., Qin et al., 2016) that did not use all of the combinations stated in our search strategy (e.g., no mention of keywords "word", AND "develop*"). In this case, these articles did not appear in our search and thus, were not included in this review. Alongside, this review only included studies published before mid-January of 2021; thus, all new publications after this period, even though they would fit the criteria, are not included and analyzed here. Another limitation in this domain is excluding combined data approaches, such as fixation-related potentials. A strict standpoint is taken on combined data studies, as one aspect of the presented systematic review was the methodological consideration of ERP research in word recognition, and combined data analyses commonly go beyond the methodological scope of conventional EEG research, thus are not comparable, especially about the perception of words. For a recent review of fixation-related potentials and reading, one can explore Degno and Liversedge (2020).

Lastly, dyslexia screening and assessment tools varied widely across the included studies, yielding different criteria to classify participants as reading impaired or typical reader. This variation might be important to consider in comparing results due to the possibility of different degrees of reading difficulties, as well the potential inclusion of different DRD subtypes. Previous studies successfully identified subtypes of DRD using learning algorithms such as mixed modeling (Torppa et al., 2007), latent profile analysis (Wolff, 2010) and confirmatory latent profile analysis (Niileksela and Templin, 2019). Although it would be interesting to see how DRD subtypes affect N170w development, this might be challenging in brain research due to lower sample sizes. Only two studies in the current review looked into subtypes; One study looked into specific difficulties in reading and spelling (Kemény et al., 2018) but did not find significant differences between the reading and spelling deficit groups, and Dujardin et al. (2011) identified dyslexia subgroups on the basis of N170 but not on the basis of their reading related skills as those did not yield a significant difference.

CONCLUSION

This review provides a more comprehensive overview of the development of the N170w across age groups (pre-literate age, school-aged and adulthood) and reading abilities (typically developing, developmental reading disorders/ developmental dyslexia/poor readers), as well as the response of N170 between word and word-like stimuli. Lastly, we discussed theoretical and methodological differences and challenges in the field to guide future research. Results showed that in adult studies, N170w amplitude is more negative in the controls than the poor readers, although mixed results were reported for children with varying reading ability. N170w lateralization is also in question, as left-lateralization is more straightforwardly reported in typical adults but still variable during childhood. Lastly, N170w vs. other word-like conditions gave mixed results across studies, depending on the investigated hemisphere, stimuli and tasks employed, as well as linguistic variables.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

KA and VC: conceptualization. KA, AT, CV, JT, PL, and VC: protocol writing, revisions, and writing—revision and editing. KA, AT, and CV: database search, synthesis of results, analysis, and writing—original draft. All authors contributed to the article and approved the submitted version.

FUNDING

This research was funded by the Neo-PRISM-C project (European Union Horizon 2020 Program, H2020-MSCA-ITN-2018) under the Marie Skłodowska-Curie Innovative Training Network (Grant Agreement No. 813546).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2022.898800/full#supplementary-material

REFERENCES

Araújo, S., Bramão, I., Faísca, L., Petersson, K. M., and Reis, A. (2012). Electrophysiological correlates of impaired reading in dyslexic pre-adolescent children. *Brain Cogn.* 79, 79–88. doi: 10.1016/j.bandc.2012.02.010

Araújo, S., Faísca, L., Bramão, I., Reis, A., and Petersson, K. M. (2015). Lexical and sublexical orthographic processing: an ERP study with skilled and dyslexic adult readers. *Brain Lang.* 141, 16–27. doi: 10.1016/j.bandl.2014.11.007

Arns, M., Peters, S., Breteler, R., and Verhoeven, L. (2007). Different brain activation patterns in dyslexic children: evidence from EEG power and coherence patterns for the double-deficit theory of dyslexia. *J. Integr. Neurosci.* 6, 175–190. doi: 10.1142/S0219635207001404

Bakos, S., Landerl, K., Bartling, J., Schulte-Körne, G., and Moll, K. (2018). Neurophysiological correlates of word processing deficits in isolated reading and isolated spelling disorders. *Clin. Neurophysiol.* 129, 526–540. doi: 10.1016/j.clinph.2017.12.010

Barber, H. A., and Kutas, M. (2006). Interplay between computational models and cognitive electrophysiology in visual word recognition. *Brain Res. Rev.* 53, 98–123. doi: 10.1016/j.brainresrev.2006.07.002

Bentin, S., Allison, T., Puce, A., Perez, E., and McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *J. Cogn. Neurosci.* 8, 551–565. doi: 10.1162/jocn.1996.8.6.551

Bentin, S., Mouchetant-Rostaing, Y., Giard, M. H., Echallier, J. F., and Pernier, J. (1999). ERP manifestations of processing printed words at different psycholinguistic levels: time course and scalp distribution. *J. Cogn. Neurosci.* 11, 235–260. doi: 10.1162/089892999563373

Brem, S., Bach, S., Kucian, K., Kujala, J. V., Guttorm, T. K., Martin, E., et al. (2010). Brain sensitivity to print emerges when children learn letter–speech sound correspondences. *Proc. Natl. Acad. Sci.* 107, 7939–7944. doi: 10.1073/pnas.0904402107

Brem, S., Bach, S., Kujala, J. V., Maurer, U., Lyytinen, H., Richardson, U., et al. (2013). An electrophysiological study of print processing in kindergarten: the contribution of the visual N1 as a predictor of reading outcome. *Dev. Neuropsychol.* 38, 567–594. doi: 10.1080/87565641.2013.828729

Brem, S., Bucher, K., Halder, P., Summers, P., Dietrich, T., Martin, E., et al. (2006). Evidence for developmental changes in the visual word processing network beyond adolescence. *NeuroImage* 29, 822–837. doi: 10.1016/j.neuroimage.2005.09.023

Brem, S., Halder, P., Bucher, K., Summers, P., Martin, E., and Brandeis, D. (2009). Tuning of the visual word processing system: distinct developmental ERP and fMRI effects. *Hum. Brain Mapp.* 30, 1833–1844. doi: 10.1002/hbm.20751

Cao, F., Rickles, B., Vu, M., Zhu, Z., Chan, D. H. L., Harris, L. N., et al. (2013). Early stage visual-orthographic processes predict long-term retention of word form and meaning: a visual encoding training study. *J. Neurolinguistics* 26, 440–461. doi: 10.1016/j.jneuroling.2013.01.003

Cao, X., Li, S., Zhao, J., Lin, S., and Weng, X. (2011). Left-lateralized early neurophysiological response for Chinese characters in young primary school children. *Neurosci. Lett.* 492, 165–169. doi: 10.1016/j.neulet.2011.02.002

Cao, X.-H., and Zhang, H.-T. (2011). Change in subtle N170 specialization in response to Chinese characters and pseudocharacters. *Percept. Mot. Skills* 113, 365–376. doi: 10.2466/04.22.24.28.PMS.113.5.365-376

Centanni, T. M., Norton, E. S., Park, A., Beach, S. D., Halverson, K., Ozernov-Palchik, O., et al. (2018). Early development of letter specialization in left fusiform is associated with better word reading and smaller fusiform face area. *Dev. Sci.* 21, e12658. doi: 10.1111/desc.12658

Chen, L., Wassermann, D., Abrams, D. A., Kochalka, J., Gallardo-Diez, G., and Menon, V. (2019). The visual word form area (VWFA) is part of both language and attention circuitry. *Nat Commun* 10, 5601. doi: 10.1038/s41467-019-13634-z

Clayson, P. E., Carbine, K. A., Baldwin, S. A., and Larson, M. J. (2019). Methodological reporting behavior, sample sizes, and statistical power in studies of event-related potentials: barriers to reproducibility and replicability. *Psychophysiology* 56. doi: 10.1111/psyp.13437

Coch, D., and Meade, G. (2016). N1 and P2 to words and wordlike stimuli in late elementary school children and adults. *Psychophysiology* 53, 115–128. doi: 10.1111/psyp.12567

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M.-A., et al. (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. doi: 10.1093/brain/123.2.291

Collins, E., Dundas, E., Gabay, Y., Plaut, D. C., and Behrmann, M. (2017). Hemispheric organization in disorders of development. *Vis. Cogn.* 25, 416–429. doi: 10.1080/13506285.2017.1370430

Curzietti, M., Bonnefond, A., Staub, B., Vidailhet, P., and Doignon-Camus, N. (2017). The effects of age on visual expertise for print. *Brain Lang.* 169, 48–56. doi: 10.1016/j.bandl.2017.03.001

Davis, C. P., Libben, G., and Segalowitz, S. J. (2019). Compounding matters: event-related potential evidence for early semantic access to compound words. *Cognition* 184, 44–52. doi: 10.1016/j.cognition.2018.12.006

Degno, F., and Liversedge, S. P. (2020). Eye movements and fixation-related potentials in reading: a review. *Vision* 4, 11. doi: 10.3390/vision4010011

Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Filho, G. N., Jobert, A., et al. (2010). How learning to read changes the cortical networks for vision and language. *Science* 330, 1359–1364. doi: 10.1126/science.1194140

Dhar, M., Been, P. H., Minderaa, R. B., and Althaus, M. (2010). Reduced interhemispheric coherence in dyslexic adults. *Cortex* 46, 794–798. doi: 10.1016/j.cortex.2009.09.006

Dien, J. (2017). Best practices for repeated measures ANOVAs of ERP data: reference, regional channels, and robust ANOVAs. *Int. J. Psychophysiol.* 111, 42–56. doi: 10.1016/j.ijpsycho.2016.09.006

Dujardin, T., Etienne, Y., Contentin, C., Bernard, C., Largy, P., Mellier, D., et al. (2011). Behavioral performances in participants with phonological dyslexia and different patterns on the N170 component. *Brain Cogn.* 75, 91–100. doi: 10.1016/j.bandc.2010.10.006

Dundas, E. M., Plaut, D. C., and Behrmann, M. (2014). An ERP investigation of the co-development of hemispheric lateralization of face and word recognition. *Neuropsychologia* 61, 315–323. doi: 10.1016/j.neuropsychologia.2014.05.006

Eberhard-Moscicka, A. K., Jost, L. B., Fehlbaum, L. V., Pfenninger, S. E., and Maurer, U. (2016). Temporal dynamics of early visual word processing – early versus late N1 sensitivity in children and adults. *Neuropsychologia* 91, 509–518. doi: 10.1016/j.neuropsychologia.2016.09.014

Eberhard-Moscicka, A. K., Jost, L. B., Raith, M., and Maurer, U. (2015). Neurocognitive mechanisms of learning to read: print tuning in beginning readers related to word-reading fluency and semantics but not phonology. *Dev. Sci.* 18, 106–118. doi: 10.1111/desc.12189

Emmorey, K., Midgley, K. J., Kohen, C. B., Sehyr, Z. S., and Holcomb, P. J. (2017). The N170 ERP component differs in laterality, distribution, and association with continuous reading measures for deaf and hearing readers. *Neuropsychologia* 106, 298–309. doi: 10.1016/j.neuropsychologia.2017.10.001

Faísca, L., Reis, A., and Araújo, S. (2019). Early brain sensitivity to word frequency and lexicality during reading aloud and implicit reading. *Front. Psychol.* 10, 830. doi: 10.3389/fpsyg.2019.00830

Feuerriegel, D., Churches, O., Hofmann, J., and Keage, H. A. D. (2015). The N170 and face perception in psychiatric and neurological disorders: A systematic review. *Clin. Neurophysiol.* 126, 1141–1158. doi: 10.1016/j.clinph.2014.09.015

Fraga González, G., Pleisch, G., Di Pietro, S. V., Neuenschwander, J., Walitza, S., Brandeis, D., et al. (2021). The rise and fall of rapid occipito-temporal sensitivity to letters: transient specialization through elementary school. *Dev. Cogn. Neurosci.* 49, 100958. doi: 10.1016/j.dcn.2021.100958

Fraga González, G., Smit, D. J. A., van der Molen, M. J. W., Tijms, J., Stam, C. J., de Geus, E. J. C., et al. (2018). EEG resting state functional connectivity in adult dyslexics using phase lag index and graph analysis. *Front. Hum. Neurosci.* 12, 341. doi: 10.3389/fnhum.2018.00341

Fraga González, G., Van der Molen, M. J. W., Žarić, G., Bonte, M., and Tijms, J., Blomert, L., et al. (2016a). Graph analysis of EEG resting state functional networks in dyslexic readers. *Clin. Neurophysiol.* 127, 3165–3175. doi: 10.1016/j.clinph.2016.06.023

Fraga González, G., Žarić, G., Tijms, J., Bonte, M., and Blomert, L., Leppänen, P., et al. (2016b). Responsivity to dyslexia training indexed by the N170 amplitude of the brain potential elicited by word reading. *Brain Cogn.* 106, 42–54. doi: 10.1016/j.bandc.2016.05.001

Fraga González, G., Žarić, G., Tijms, J., Bonte, M., Blomert, L., and van der Molen, M. W. (2014). Brain-potential analysis of visual word recognition in dyslexics and typically reading children. *Front. Hum. Neurosci.* 8, 474. doi: 10.3389/fnhum.2014.00474

González-Garrido, A. A., Gómez-Velázquez, F. R., and Rodríguez-Santillán, E. (2014). Orthographic recognition in late adolescents: an assessment through event-related brain potentials. *Clin. EEG Neurosci.* 45, 113–121. doi: 10.1177/1550059413489975

Gross, J., Baillet, S., Barnes, G. R., Henson, R. N., Hillebrand, A., Jensen, O., et al. (2013). Good practice for conducting and reporting MEG research. *Neuroimage* 65, 349–363. doi: 10.1016/j.neuroimage.2012.10.001

Hasko, S., Groth, K., Bruder, J., Bartling, J., and Schulte-Körne, G. (2013). The time course of reading processes in children with and without dyslexia: an ERP study. *Front. Hum. Neurosci.* 7, 570. doi: 10.3389/fnhum.2013.00570

Herrmann, C. S., and Knight, R. T. (2001). Mechanisms of human attention: event-related potentials and oscillations. *Neurosc. Biobehav. Rev.* 25, 465–476. doi: 10.1016/S0149-7634(01)00027-6

Jucla, M., Nenert, R., Chaix, Y., and Demonet, J.-F. (2010). Remediation effects on N170 and P300 in children with developmental dyslexia. *Behav. Neurol.* 22, 121–129. doi: 10.1155/2010/913692

Karipidis, I., Pleisch, G., Röthlisberger, M., Hofstetter, C., Dornbierer, D., Stämpfli, P., et al. (2017). Neural initialization of audiovisual integration in prereaders at varying risk for developmental dyslexia. *Hum. Brain Mapp.* 38, 1038–1055. doi: 10.1002/hbm.23437

Kast, M., Elmer, S., Jancke, L., and Meyer, M. (2010). ERP differences of pre-lexical processing between dyslexic and non-dyslexic children. *Int. J. Psychophysiol.* 77, 59–69. doi: 10.1016/j.ijpsycho.2010.04.003

Keil, A., Debener, S., Gratton, G., Junghöfer, M., Kappenman, E. S., Luck, S. J., et al. (2014). Committee report: publication guidelines and recommendations for studies using electroencephalography and magnetoencephalography: guidelines for EEG and MEG. *Psychophysiology* 51, 1–21. doi: 10.1111/psyp.12147

Kemény, F., Banfi, C., Gangl, M., Perchtold, C. M., Papousek, I., Moll, K., et al. (2018). Print-, sublexical and lexical processing in children with reading and/or spelling deficits: an ERP study. *Int. J. Psychophysiol.* 130, 53–62. doi: 10.1016/j.ijpsycho.2018.05.009

Korinth, S. P., and Breznitz, Z. (2014). Fast and slow readers of the Hebrew language show divergence in brain response ∼200 ms post stimulus: an ERP study. *PLoS ONE* 9, e103139. doi: 10.1371/journal.pone.0103139

Korinth, S. P., Sommer, W., and Breznitz, Z. (2012). Does silent reading speed in normal adult readers depend on early visual processes? Evidence from event-related brain potentials. *Brain Lang.* 120, 15–26. doi: 10.1016/j.bandl.2011.08.003

Leppänen, P. H., and Lyytinen, H. (1997). Auditory event-related potentials in the study of developmental language-related disorders. *Audiol. Neurootol.* 2, 308–340. doi: 10.1159/000259254

Li, S., Lee, K., Zhao, J., Yang, Z., He, S., and Weng, X. (2013). Neural competition as a developmental process: early hemispheric specialization for word processing delays specialization for face processing. *Neuropsychologia* 51, 950–959. doi: 10.1016/j.neuropsychologia.2013.02.006

Lin, S. E., Chen, H. C., Zhao, J., Li, S., He, S., and Weng, X. C. (2011). Left-lateralized N170 response to unpronounceable pseudo but not false

Chinese characters—the key role of orthography. *Neuroscience* 190, 200–206. doi: 10.1016/j.neuroscience.2011.05.071

Luck, S. J. (2012). "Event-related potentials," in *APA Handbook of Research Methods in Psychology, Vol 1: Foundations, Planning, Measures, and Psychometrics APA Handbooks in Psychology*®, eds H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, and K. J. Sher (Washington, DC: American Psychological Association), 523–546. doi: 10.1037/13619-028

Mahé, G., Bonnefond, A., and Doignon-Camus, N. (2013). Is the impaired N170 print tuning specific to developmental dyslexia? A matched reading-level study with poor readers and dyslexics. *Brain Lang.* 127, 539–544. doi: 10.1016/j.bandl.2013.09.012

Mahé, G., Bonnefond, A., Gavens, N., Dufour, A., and Doignon-Camus, N. (2012). Impaired visual expertise for print in French adults with dyslexia as shown by N170 tuning. *Neuropsychologia* 50, 3200–3206. doi: 10.1016/j.neuropsychologia.2012.10.013

Maurer, U., Blau, V. C., Yoncheva, Y. N., and McCandliss, B. D. (2010). Development of visual expertise for reading: rapid emergence of visual familiarity for an artificial script. *Dev. Neuropsychol.* 35, 404–422. doi: 10.1080/87565641.2010.480916

Maurer, U., Brem, S., Bucher, K., and Brandeis, D. (2005). Emerging neurophysiological specialization for letter strings. *J. Cogn. Neurosci.* 17, 1532–1552. doi: 10.1162/089892905774597218

Maurer, U., Brem, S., Bucher, K., Kranz, F., Benz, R., Steinhausen, H.-C., et al. (2007). Impaired tuning of a fast occipito-temporal response for print in dyslexic children learning to read. *Brain* 130, 3200–3210. doi: 10.1093/brain/awm193

Maurer, U., Brem, S., Kranz, F., Bucher, K., Benz, R., Halder, P., et al. (2006). Coarse neural tuning for print peaks when children learn to *read*. 10, 749–758. doi: 10.1016/j.neuroimage.2006.06.025

Maurer, U., and McCandliss, B. D. (2008). "The development of visual expertise for words: the contribution of electrophysiology," in *Single-Word Reading: Behavioral and Biological Perspectives*, eds E. L. Grigorenko and A. J. Naples (New York, NY: Lawrence Erlbaum Associates Publishers), 43–63.

Maurer, U., Rossion, B., and McCandliss, B. D. (2008a). Category specificity in early perception: face and word N170 responses differ in both lateralization and habituation properties. *Front. Hum. Neurosci.* 2, 18. doi: 10.3389/neuro.09.018.2008

Maurer, U., Schulz, E., Brem, S., der Mark, S., van, Bucher, K., Martin, E., et al. (2011). The development of print tuning in children with dyslexia: evidence from longitudinal ERP data supported by fMRI. *Neuroimage* 57, 714–722. doi: 10.1016/j.neuroimage.2010.10.055

Maurer, U., Zevin, J. D., and McCandliss, B. D. (2008b). Left-lateralized N170 effects of visual expertise in reading: evidence from Japanese syllabic and logographic scripts. *J. Cogn. Neurosci.* 20, 1878–1891. doi: 10.1162/jocn.2008.20125

McCandliss, B. D., Cohen, L., and Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends Cogn. Sci.* 7, 293–299. doi: 10.1016/S1364-6613(03)00134-7

Mercure, E., Ashwin, E., Dick, F., Halit, H., Auyeung, B., Baron-Cohen, S., et al. (2009). IQ, fetal testosterone and individual variability in children's functional lateralization. *Neuropsychologia* 47, 2537–2543. doi: 10.1016/j.neuropsychologia.2009.04.027

Mercure, E., Kadosh, K. C., and Johnson, M. H. (2011). The N170 shows differential repetition effects for faces, objects, and orthographic stimuli. *Front. Hum. Neurosci.* 5, 6. doi: 10.3389/fnhum.2011.00006

Modesti, P. A., Reboldi, G., Cappuccio, F. P., Agyemang, C., Remuzzi, G., Rapi, S., et al. (2016). Panethnic differences in blood pressure in europe: a systematic review and meta-analysis. *PLoS ONE* 11, e0147601. doi: 10.1371/journal.pone.0147601

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., and Group, T. P. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* 6, e1000097. doi: 10.1371/journal.pmed.1000097

Niileksela, C. R., and Templin, J. (2019). Identifying dyslexia with confirmatory latent profile analysis. *Psychol. Schools* 56, 335–359. doi: 10.1002/pits.22183

Okumura, Y., Kasai, T., and Murohashi, H. (2014). Early print-tuned ERP response with minimal involvement of linguistic processing in Japanese Hiragana strings. *Neuro Rep.* 25, 410–414. doi: 10.1097/WNR.0000000000000110

Okumura, Y., Kasai, T., and Murohashi, H. (2015). Attention that covers letters is necessary for the left-lateralization of an early print-tuned ERP in Japanese hiragana. *Neuropsychologia* 69, 22–30. doi: 10.1016/j.neuropsychologia.2015.01.026

Ouzzani, M., Hammady, H., Fedorowicz, Z., and Elmagarmid, A. (2016). Rayyan—a web and mobile app for systematic reviews. *Syst. Rev.* 5, 210. doi: 10.1186/s13643-016-0384-4

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The prisma 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 372, 71. doi: 10.31222/osf.io/v7gm2

Paul, M., Govaart, G. H., and Schettino, A. (2021). Making ERP research more transparent: guidelines for preregistration. *Int. J. Psychophysiol.* 164, 52–63. doi: 10.1016/j.ijpsycho.2021.02.016

Pegado, F., Comerlato, E., Ventura, F., Jobert, A., Nakamura, K., Buiatti, M., et al. (2014). Timing the impact of literacy on visual processing. *Proc. Natl. Acad. Sci. U. S. A.* 111, E5233–E5242. doi: 10.1073/pnas.1417347111

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., et al. (2000). Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. *Psychophysiology* 37, 127–152. doi: 10.1111/1469-8986.3720127

Pleisch, G., Karipidis, I. I., Brem, A., Röthlisberger, M., Roth, A., Brandeis, D., et al. (2019). Simultaneous EEG and fMRI reveals stronger sensitivity to orthographic strings in the left occipito-temporal cortex of typical versus poor beginning readers. *Dev. Cogn. Neurosci.* 40, 100717. doi: 10.1016/j.dcn.2019.100717

Price, C. J., Moore, C., Humphreys, G. W., and Wise, R. J. S. (1997). Segregating semantic from phonological processes during reading. *J. Cogn. Neurosci.* 9, 727–733. doi: 10.1162/jocn.1997.9.6.727

Qin, R., Maurits, N., and Maassen, B. (2016). N170 tuning in Chinese: logographic characters and phonetic Pinyin script. *Sci. Stud. Read.* 20, 363–374. doi: 10.1080/10888438.2016.1199554

Rossion, B., Joyce, C. A., Cottrell, G. W., and Tarr, M. J. (2003). Early lateralization and orientation tuning for face, word, and object processing in the visual cortex. *Neuroimage* 20, 1609–1624. doi: 10.1016/j.neuroimage.2003.07.010

Rumsey, J. M., Horwitz, B., Donohue, B. C., Nace, K., Maisog, J. M., and Andreason, P. (1997). Phonological and orthographic components of word recognition. A PET-rCBF study. *Brain* 120, 739–759. doi: 10.1093/brain/120.5.739

Sacchi, E., and Laszlo, S. (2016). An event-related potential study of the relationship between N170 lateralization and phonological awareness in developing readers. *Neuropsychologia* 91, 415–425. doi: 10.1016/j.neuropsychologia.2016.09.001

Savill, N. J., and Thierry, G. (2011a). Reading for sound with dyslexia: evidence for early orthographic and late phonological integration deficits. *Brain Res.* 1385, 192–205. doi: 10.1016/j.brainres.2011.02.012

Savill, N. J., and Thierry, G. (2011b). Electrophysiological evidence for impaired attentional engagement with phonologically acceptable misspellings in developmental dyslexia. *Front. Psychol.* 2, 139. doi: 10.3389/fpsyg.2011.00139

Schmalz, X., Marinus, E., Coltheart, M., and Castles, A. (2015). Getting to the bottom of orthographic depth. *Psychon. Bull. Rev.* 22, 1614–1629. doi: 10.3758/s13423-015-0835-2

Seymour, P. H. K., Aro, M., and Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *Br. J. Psychol.* 94, 143–174. doi: 10.1348/000712603321661859

Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Mencl, W. E., Fulbright, R. K., Skudlarski, P., et al. (2002). Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biol. Psychiatry* 52, 101–110. doi: 10.1016/S0006-3223(02)01365-3

Shaywitz, S. E., Shaywitz, B. A., Pugh, K. R., Fulbright, R. K., Constable, R. T., Mencl, W. E., et al. (1998). Functional disruption in the organization of the brain for reading in dyslexia. *Proc. Natl. Acad. Sci. U. S. A.* 95, 2636–2641. doi: 10.1073/pnas.95.5.2636

Simon, G., Petit, L., Bernard, C., and Rebaï, M. (2007). N170 ERPs could represent a logographic processing strategy in visual word recognition. *Behav. Brain Funct.* 3, 21. doi: 10.1186/1744-9081-3-21

Snowling, M. J. (2013). Early identification and interventions for dyslexia: a contemporary view. *J. Res. Spec. Educ. Needs* 13, 7–14. doi: 10.1111/j.1471-3802.2012.01262.x

Spironelli, C., and Angrilli, A. (2009). Developmental aspects of automatic word processing: language lateralization of early ERP components in children, young adults and middle-aged subjects. *Biol. Psychol.* 80, 35–45. doi: 10.1016/j.biopsycho.2008.01.012

Spironelli, C., Penolazzi, B., and Angrilli, A. (2010). Gender differences in reading in school-aged children: an early ERP study. *Dev. Neuropsychol.* 35, 357–375. doi: 10.1080/87565641.2010.480913

Strijkers, K., Yum, Y. N., Grainger, J., and Holcomb, P. J. (2011). Early Goal-directed top-down influences in the production of speech. *Front. Psychol.* 2, 371. doi: 10.3389/fpsyg.2011.00371

Su, M., Wang, J., Maurer, U., Zhang, Y., Li, J., McBride, C., et al. (2015). Gene–environment interaction on neural mechanisms of orthographic processing in Chinese children. *J. Neurolinguistics* 33, 172–186. doi: 10.1016/j.jneuroling.2014.09.007

Taha, H., Ibrahim, R., and Khateb, A. (2013). How does Arabic orthographic connectivity modulate brain activity during visual word recognition: an ERP study. *Brain Topogr.* 26, 292–302. doi: 10.1007/s10548-012-0241-2

Taha, H., and Khateb, A. (2013). Resolving the orthographic ambiguity during visual word recognition in Arabic: an event-related potential investigation. *Front. Hum. Neurosci.* 7, 821. doi: 10.3389/fnhum.2013.00821

Tong, X., Lo, J. C. M., McBride, C., Ho, C. S., Waye, M. M. Y., Chung, K. K. H., et al. (2016a). Coarse and fine N1 tuning for print in younger and older Chinese children: orthography, phonology, or semantics driven? *Neuropsychologia* 91, 109–119. doi: 10.1016/j.neuropsychologia.2016.08.006

Tong, X., Maurer, U., Chung, K. K. H., and McBride, C. (2016b). Neural specialization for print in Chinese-English language learners. *J. Neurolinguistics* 38, 42–55. doi: 10.1016/j.jneuroling.2015.10.001

Torppa, M., Tolvanen, A., Poikkeus, A.-M., Eklund, K., Lerkkanen, M.-K., Leskinen, E., et al. (2007). Reading development subtypes and their early characteristics. *Ann. Dyslexia* 57, 3–32. doi: 10.1007/s11881-007-0003-0

Uno, T., Okumura, Y., and Kasai, T. (2017). Print-specific N170 involves multiple subcomponents for Japanese Hiragana. *Neurosci. Lett.* 650, 77–81. doi: 10.1016/j.neulet.2017.04.020

van Setten, E. R. H., Martinez-Ferreiro, S., Maurits, N. M., and Maassen, B. A. M. (2016). Print-tuning lateralization and handedness: an event-related potential Study in dyslexic higher education students. *Dyslexia* 22, 64–82. doi: 10.1002/dys.1519

van Setten, E. R. H., Maurits, N. M., and Maassen, B. A. M. (2019). N1 lateralization and dyslexia: an event-related potential study in children with a familial risk of dyslexia. *Dyslexia* 25, 84–102. doi: 10.1002/dys.1604

Van Strien, J. W., Glimmerveen, J. C., Martens, V. E. G., and De Bruin, E. A. (2009). Age-related differences in brain activity during extended continuous word recognition in children. *Neuroimage* 47, 688–699. doi: 10.1016/j.neuroimage.2009.05.020

Waldie, K. E., Badzakova-Trajkov, G., Lim, V. K., and Kirk, I. J. (2012). Lexical decision making in adults with dyslexia: an event-related potential study. *Ilha Desterro J. Engl. Lang. Lit. Engl. Cult. Stud.* 63, 37–68. doi: 10.5007/2175-8026.2012n63p37

Wei, D., Gillon Dowens, M., and Guo, T. (2018). Early lexical processing of Chinese words indexed by Visual Mismatch Negativity effects. *Sci. Rep.* 8, 1289. doi: 10.1038/s41598-018-19394-y

Wolff, U. (2010). Subgrouping of readers based on performance measures: a latent profile analysis. *Read. Writ.* 23, 209–238. doi: 10.1007/s11145-008-9160-8

Wong, A. C. N., Gauthier, I., Woroch, B., Debuse, C., and Curran, T. (2005). An early electrophysiological response associated with expertise in letter perception. *Cogn. Affect. Behav. Neurosci.* 5, 306–318. doi: 10.3758/CABN.5.3.306

Xue, L., Maurer, U., Weng, X., and Zhao, J. (2019). Familiarity with visual forms contributes to a left-lateralized and increased N170 response for Chinese characters. *Neuropsychologia* 134, 107194. doi: 10.1016/j.neuropsychologia.2019.107194

Yang, H., Zhao, J., Gaspar, C. M., Chen, W., Tan, Y., and Weng, X. (2017). Selectivity of N170 for visual words in the right hemisphere: evidence from single-trial analysis. *Psychophysiology* 54, 1128–1137. doi: 10.1111/psyp.12867

Yum, Y. N., Holcomb, P. J., and Grainger, J. (2011). Words and pictures: an electrophysiological investigation of domain specific processing in native Chinese and English speakers. *Neuropsychologia* 49, 1910–1922. doi: 10.1016/j.neuropsychologia.2011.03.018

Zhao, J., Kipp, K., Gaspar, C., Maurer, U., Weng, X., Mecklinger, A., et al. (2014). Fine neural tuning for orthographic properties of words emerges early in children reading alphabetic script. *J. Cogn. Neurosci.* 26, 2431–2442. doi: 10.1162/jocn_a_00660

Zhao, J., Maurer, U., He, S., and Weng, X. (2019). Development of neural specialization for print: evidence for predictive coding in visual word recognition. *PLoS Biol.* 17, e3000474. doi: 10.1371/journal.pbio.3000474

Zhao, P., Li, S., Zhao, J., Gaspar, C. M., and Weng, X. (2015). Training by visual identification and writing leads to different visual word expertise N170 effects in preliterate Chinese children. *Dev. Cogn. Neurosci.* 15, 106–116. doi: 10.1016/j.dcn.2015.09.002

Zhao, P., Zhao, J., Weng, X., and Li, S. (2018). Event-related potential evidence in Chinese children: type of literacy training modulates neural orthographic sensitivity. *Int. J. Behav. Dev.* 42, 311–320. doi: 10.1177/0165025417708341

Zhou, X., Ye, Z., Cheung, H., and Chen, H.-C. (2009). Processing the Chinese language: an introduction. *Lang. Cogn. Process.* 24, 929–946. doi: 10.1080/01690960903201281

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Neuroscience

ORIGINAL RESEARCH
published: 18 July 2022
doi: 10.3389/fnins.2022.919440

Disruption of Functional Brain Networks Underlies the Handwriting Deficit in Children With Developmental Dyslexia

Zhengyan Liu[1,2], Junjun Li[1,2], Hong-Yan Bi[1,2], Min Xu[3] and Yang Yang[1,2]**

[1] CAS Key Laboratory of Behavioral Science, Center for Brain Science and Learning Difficulties, Institute of Psychology, Chinese Academy of Sciences, Beijing, China, [2] Department of Psychology, University of Chinese Academy of Sciences, Beijing, China, [3] Center for Brain Disorders and Cognitive Sciences, School of Psychology, Shenzhen University, Shenzhen, China

Developmental dyslexia (DD) is a neurological-based learning disorder that affects 5-17.5% of children. Handwriting difficulty is a prevailing symptom of dyslexia, but its neural mechanisms remain elusive. Using functional magnetic resonance imaging (fMRI), this study examined functional brain networks associated with handwriting in a copying task in Chinese children with DD (n = 17) and age-matched children (n = 36). We found that dyslexics showed reduced network connectivity between the sensory-motor network (SMN) and the visual network (VN), and between the default mode network (DMN) and the ventral attention network (VAN) during handwriting, but not during drawing geometric figures. Moreover, the connectivity strength of the networks showing group differences was correlated with handwriting speed, reading and working memory, suggesting that the handwriting deficit in DD is linked with disruption of a large-scale brain network supporting motoric, linguistic and executive control processes. Taken together, this study demonstrates the alternations of functional brain networks that underly the handwriting deficit in Chinese dyslexia, providing a new clue for the neural basis of DD.

Keywords: developmental dyslexia, handwriting, fMRI, functional brain network, children

Edited by:
Charles A. Perfetti,
University of Pittsburgh, United States

Reviewed by:
Fan Cao,
Sun Yat-sen University, China
Juliana Dushanova,
Institute of Neurobiology, Bulgarian Academy of Sciences (BAS), Bulgaria

****Correspondence:***
Min Xu
xumin@szu.edu.cn
Yang Yang
yangyang@psych.ac.cn

Specialty section:
This article was submitted to Neurodevelopment, a section of the journal Frontiers in Neuroscience

Received: *13 April 2022*
Accepted: *20 June 2022*
Published: *18 July 2022*

Citation:
Liu Z, Li J, Bi H-Y, Xu M and Yang Y (2022) Disruption of Functional Brain Networks Underlies the Handwriting Deficit in Children With Developmental Dyslexia. Front. Neurosci. 16:919440. doi: 10.3389/fnins.2022.919440

INTRODUCTION

Developmental Dyslexia (DD) is a learning disorder that is characterized by unexpected reading difficulties despite adequate intelligence and educational opportunities. It affects 5-17.5% of children (Shaywitz, 1998). Previous neuroimaging studies have demonstrated that dyslexia is associated with the abnormalities of regional activity and functional connectivity in multiple brain systems including the left hemisphere reading network (e.g., the temporoparietal cortex, inferior frontal gyrus and occipitotemporal cortex) (Shaywitz et al., 2002; Hoeft et al., 2007; van der Mark et al., 2009, 2011; Boets et al., 2013; Olulade et al., 2013; Finn et al., 2014) and the cerebellum (Nicolson et al., 1999; Menghini et al., 2006; Yang et al., 2013).

In addition to reading difficulties, dyslexics pervasively exhibit handwriting deficits (Graham et al., 2021). For example, children with DD show poorer writing legibility (Martlew, 1992) and larger size of written scripts than typically developed children (Lam et al., 2011). Moreover, relative to typical readers, dyslexic readers were found to show increased writing latency (Afonso et al., 2020), reduced motor speed (Pagliarini et al., 2015) and more pauses (Sumner et al., 2012, 2014) during handwriting. Chinese is a logographic/morphosyllabic writing system, in which a grapheme corresponds to a syllabic morpheme (Perfetti and Harris, 2013). There is a lack of one-to-one correspondence between phonology and orthography in Chinese, and what's more, there are many homophonic characters. At the script level, character is the basic written unit in Chinese that has a square configuration consisting of many radicals and strokes, resulting in a high level of visual complexity. Due to the linguistic and visual features, handwriting becomes a prevalent strategy for masting Chinese reading via the elaboration of orthographic representation and the formation of motor memory (Tan et al., 2005, 2013). Accordingly, the handwriting problem is expected to be more relevant with dyslexia in Chinese than that in alphabetic languages (e.g., English) (Kalindi et al., 2015).

Handwriting is a complex process involving linguistic, motor and executive control processes, and thus the underlying causes of the handwriting deficit in DD is likely multifactorial. One possibility is that the handwriting deficit is derived from orthographic processing impairments in dyslexics (Cao et al., 2006; Boros et al., 2016), which lead to an inability to extract and to operate orthographic information quickly and accurately during handwriting. This hypothesis is supported by existing empirical evidence showing that dyslexics are more impacted by orthographic complexity (Arfe et al., 2020) or spelling regularity (Sumner et al., 2014) than typical readers during handwriting. Another possibility is that motor skill impairment is the origin of the handwriting deficit in DD. For instance, kinematic measures of handwriting processing revealed that, relative to typically developed readers, dyslexics showed increased motor variability (Pagliarini et al., 2015) and greater vulnerability to motor complexity (Gosse and Van Reybroeck, 2020) in handwriting.

Despite the extensive research on behavioral manifestations, the neural bases of the handwriting deficit in DD remain largely unknown. An fMRI study has reported that French-speaking dyslexic children showed reduced brain activation in the right anterior cerebellum relative to age-matched typically developed children in a dictation task (Gosse et al., 2022). The cerebellum is a key brain locus of motor processing, and thus this result favors the motor impairment hypothesis. However, this study used a region of interest (ROI) analysis approach, which can hardly delineate the full map of brain dysfunction associated with the handwriting deficit in DD. Recently, another fMRI study investigated the neural basis of the handwriting deficit in Chinese dyslexic children. Whole brain analysis revealed that Chinese dyslexic children showed decreased brain activation in somatomotor regions (the supplementary motor area (SMA) and postcentral gyrus) and visual-orthographic regions (the bilateral precuneus and right cuneus), while showed hyperactivation in the left inferior frontal gyrus and anterior cingulate cortex. Moreover, using seed-to-voxel connectivity analysis, this study revealed aberrant functional connectivity within the neural circuits for cognitive control and sensory-motor processes involved in handwriting in dyslexics (Yang et al., 2022). These findings suggest that the handwriting difficulty in DD is linked with a malfunction of distributed brain systems involved in motor, visual-orthographic and executive control processes. However, local activity and seed-based functional connectivity analyses are still not sufficient to decipher the large-scale interaction between brain regions involved in handwriting.

According to the graph theory of brain, functional and structural brain systems are organized as graphs formed by highly connected hubs and modularity (Bullmore and Sporns, 2009). Brain network analysis provides an intuitive and powerful way for illustrating the topological principles of brain function underlying complex cognitive processes and neurological disorders (Sporns, 2011). Such network analysis method has been applied in the investigation of the neural signatures of dyslexia, revealing that dyslexia is related to altered functional connectivity in multiple brain networks during rest (Finn et al., 2014) and task states (Zhang J. et al., 2021).

This study used a network analysis approach to explore the alternations of functional brain networks underlying the handwriting deficit in Chinese children with dyslexia. A delayed copying task was used, which is thought to have the advantages in controlling high-order linguistic/cognitive processes (Yang et al., 2022) Furthermore, to examine the influence of linguistic factors to motor execution in handwriting (Kandel and Perret, 2015), the frequency of character was manipulated in this study. Character frequency is a lexical variable that has been found to impact orthographic access during handwriting (Yang et al., 2018). We hypothesized that dyslexics would show disruption in multiple functional networks related to handwriting including the sensorimotor network, visual network and cognitive control network. Moreover, functional disruption of the motor and visual networks in DD was expected to be independent of character frequency, whereas group differences in cognitive control networks are mediated by character frequency.

MATERIALS AND METHODS

Participants

Seventeen children with DD (11 males) and 36 age-matched controls (15 males) participated in this study. The dyslexic participants were screened according to the following criteria: (1) the score of the Character Recognition Measures and Assessment Scale (CRM) (Wang, 1986) was at least 1.25 standard deviations (SD) below the average score of children in the same grade. The CRM is a standardized reading test that has been widely used for screening dyslexia in Mandarin-speaking children (Amalric and Dehaene, 2016; Zhang et al., 2018; Feng et al., 2020; Yang et al., 2022); (2) having a normal non-verbal intelligence quotient (IQ) standardized score (above 85) as evaluated by Combined Raven's Progressive Matrices; (3) having normal hearing, normal or corrected-to-normal vision, and no ophthalmological or neurological abnormalities; and (4) not suffering from attention deficit/hyperactivity disorder (ADHD) evaluated by the Chinese Classification of Mental Disorder 3 (CCMD-3). All the children were native speakers of Mandarin, and were right-handed as assessed by the Handedness Inventory (Snyder and Harris, 1993).

The study was approved by the ethics committee of the Institute of Psychology at the Chinese Academy of Sciences. Prior to entering the study, written informed consent was obtained from the guardian of each child participant. Detailed participant information was listed in **Table 1**.

Behavioral Tests

A series of behavioral tests were administered to examine reading, handwriting and domain-general cognitive skills of the participants.

Reading Skills Tests

Reading-related skills including reading fluency, phonological awareness and orthographic awareness were assessed. The reading fluency test consisted of 160 Chinese characters of medium to high frequency. The participants were asked to read aloud these characters as fast and accurately as possible within 1 min. The number of correctly named characters was defined as the final score. Phonological awareness was assessed by using the oddity tests. In this test, the participants were required to listen carefully to three syllables, and were then asked to orally report the odd syllable that differed in initial sound, final sound, or tone with the other two syllables. The final score was defined as the total number of items correctly answered. There were 10 items for each type of stimuli, and thus the maximum score was 30. Finally, orthographic awareness was evaluated in a character judgment task. This test consists of 40 real Chinese characters, 20 pseudo-characters and 20 non-characters. Participants were asked to judge whether the stimuli were real Chinese characters or not. The mean accuracy (ACC) and reaction time (RT) of real characters, pseudo-characters and non-characters were defined as the final score.

TABLE 1 | Demographic information of the participants and behavioral performance.

	Dyslexics (*n* = 17)	Controls (*n* = 36)	*P*-value
Sex (male/female)	11/6	15/21	0.117
Age	10.28 (0.57)	10.40 (0.54)	0.479
Raven IQ	105.76 (9.43)	111.81 (15.56)	0.087
CRM	1978.1 (315.65)	2908.46 (261.6)	<0.001
Reading-related skills			
Reading fluency (character/min)	65.29 (19.06)	100.58 (19.57)	<0.001
Phonological awareness	25.24 (3.99)	28.11 (2.04)	0.001
Orthographic awareness			
Mean ACC	0.75 (0.14)	0.84 (0.08)	0.004
Mean RT (in ms)	994.41 (143.2)	865.09 (137.46)	0.003
Handwriting skill			
Copying tasks			
Speed (characters/s)	0.46 (0.09)	0.50 (0.10)	0.226
Quality	24.75 (4.24)	25.55 (6.07)	0.628
Handwriting fluency			
Characters	24.35 (6.24)	27.64 (4.82)	0.040
Digits	51.88 (8.91)	58.50 (12.20)	0.051
Cognitive skill			
Phonological working memory	4.76 (1.03)	6.33 (1.64)	<0.001
Sustained attention	29.06 (7.84)	31.59 (5.98)	0.253

IQ = intelligence quotient, CRM = the Character Recognition Measures and Assessment Scale, ACC = accuracy, RT = response time, ms = millisecond, min = minute and s = second.

Handwriting Skills Tests

Handwriting skills were assessed in a copying task and a handwriting fluency task. In the copying task, participants were required to copy 48 Chinese characters using habitual writing styles (Yang et al., 2022). Writing quality and speed were evaluated. Writing quality was evaluated by two independent (one male) examiners using a 7-point Likert scale (1 = very bad and 7 = very good) based on six dimensions, including stroke form, slant, organization of radicals, neatness, average size, and overall appearance (Gimenez et al., 2014; Yang et al., 2020). The score was the sum of the sub-scores across all dimensions. The inter-rater reliability of the assessment was high (intra-class correlation coefficient (ICC) = 0.92). In the handwriting fluency test, participants were asked to continuously handwrite a Chinese sentence "妈妈永远爱我" ("Mommy loves me for forever") or digits ranging from 1 to 10 in Chinese as fast and as legibly as possible within 1 min. The score was the number of legible characters or digits.

Cognitive Skills Tests

Working memory and sustained attention were assessed because they are necessarily involved in handwriting and reading processes. First, phonological working memory was measured by using a backward digit span task, in which participants were asked to orally reproduce digits (3 to 10 digits) in the reverse order as they were presented (Zhang et al., 2018; Yang et al., 2022). The test was terminated when the participants failed in two consecutive trials of the same length, and the score was the maximum length of digit span with a correct response. Second, sustained attention was assessed using a digit cancellation test (Yang et al., 2022). Participants were required to search the target number ("3") from a list of numbers as quickly and accurately as possible within 3 min. The score was calculated according to the following formulas: score = attack - (false alarms + 0.5*omission), where attack was the number of correctly marked items, false alarms was the number of incorrectly marked items, and omission was the number of items missed.

Stimuli and Task Procedure During Functional Magnetic Resonance Imaging

The participants performed a delayed copying task during fMRI scanning. The stimuli included thirty-two Chinese characters, including 16 high-frequency characters (HFCs) (mean frequency = 2486 times per million) and 16 low-frequency characters (LFCs) (mean frequency = 91 times per million), according to the Wang (1986). In addition, a drawing condition was included as a control condition for excluding low-level visual and motor processes, in which participants were asked to draw geometric figures (line, dot, circle, and triangle) as instructed by presentation of the appropriate Chinese characters. A direct copying task was also included as part of a large study. However, as this condition rarely taps the processing of orthographic working memory that we were interested in, it was not analyzed in the present study. Participants were instructed to start handwritten or drawn responses when the cursor appeared (a pencil symbol).

Handwriting data were recorded using a tablet system that includes a touch-sensitive surface, a force-sensitive stylus and an adjustable support frame, which is MRI-safe without significantly degrading fMRI data quality (Tam et al., 2011). Participants used the stylus to write on the surface. The support frame was adjusted carefully for each participant to ensure that handwriting and drawing could be undertaken comfortably throughout the imaging session, and to enable tablet interaction with the forearm or wrist resting on the support such that there was no fatigue from handwriting against gravity. To approximate real handwriting, immediate visual ("ink") feedback was provided via a mirror installed in the scanner that can reflect the writing traces displayed on the computer screen during writing responses. Participants were trained to write and draw with matched duration and size, while minimizing movements of their upper arm and forearm to minimize task-related head motion during fMRI scanning.

A block design was employed in this study, with four blocks for each condition. Each block consisted of an instruction presenting for 2 s and subsequent four trials. In each trial, a fixation cross ('+') was first presented centrally for 0.5 s, followed by the presentation of a character for 1.2 s. Then, a blank screen was displayed during a delay period of 0.5 s; afterwards, the cursor appeared to allow participants to write or draw within the response period of 5.3 s. Eight blocks of central fixation with 12 s duration were interspersed between each of the two task blocks as a "rest" condition. Each participant underwent two fMRI runs, and each run consisted of two blocks of task condition and 8 rest blocks. Detailed information about the experimental design and fMRI scanning procedures have been reported previously (Yang et al., 2022).

Imaging Acquisition

Imaging was performed using a 3T MRI system (MAGNETOM Prismafit, Siemens, Erlangen, Germany) at the Beijing MRI Center for Brain Research of the Chinese Academy of Sciences. Functional MRI time series data with blood oxygenation level-dependent (BOLD) contrast were acquired using a two-dimensional, T2*-weighted, multiband gradient-echo echo planar imaging sequence (Moeller et al., 2010): four-fold acceleration, repetition time (TR) = 1000 ms, echo time (TE) = 30 ms, slice thickness = 2.2 mm, in-plane resolution = 2.2 × 2.2 mm, flip angle (θ) = 45°, 64 axial slices. High spatial resolution anatomical images were acquired using a three-dimensional, T1-weighted, magnetization-prepared rapid acquisition gradient echo sequence: TR = 2200 ms, TE = 3.49 ms, slice thickness = 1 mm, inversion time (TI) = 1000 ms, in-plane resolution = 1.0 × 1.0 mm, and θ = 8°.

Data Analysis

Behavior Data

Handwriting latency and duration were analyzed for the delayed copying task and the drawing task during fMRI scanning. Writing latency was defined as the time period between the appearance of the response screen and the start of the response (first contact with the tablet), while writing duration was defined as the length of time from the start of the response to the end of the last written or drawn stroke of the response. A 2 (group: dyslexics vs. controls) by 3 (stimulus type: HFC vs. LFC vs. figures) analysis of variance (ANOVA) was conducted for writing latency and duration, respectively. The statistical significance was set at $p < 0.05$.

Image Data

Preprocessing

Image preprocessing and statistical analyses were performed using SPM12 freeware ([1]Wellcome Department of Cognitive Neurology, University College London, London). The fMRI time series data for each participant were first corrected for head motion, and the corrected images were coregistered to the associated anatomical imaging data. The anatomical images were then transformed into Montreal Neurological Institute (MNI) stereotactic space, and the resulting transformation parameters were applied to yield fMRI time series data normalized in MNI space with cubic voxels at a spatial resolution of 2 mm × 2 mm × 2 mm. These images were then spatially smoothed using an isotropic Gaussian kernel template with 6 mm full-width at half-maximum. Three dyslexic children were excluded from the data analysis because of excessive head motion (>3 mm translation or >3° rotation) during fMRI scanning, and a dyslexic child was excluded due to poor quality of T1-weighted images. For age-matched controls, six children were excluded because of excessive head motion (>3 mm translation or >3° rotation) during fMRI scanning, and seven controls were excluded due to poor quality of T1-weighted images. Accordingly, thirteen children with dyslexia and twenty-three age-matched controls were included in further statistical analysis. The head motion was quantified by calculating the mean framewise displacement (FD) (Power et al., 2012) based on the estimates of the six head movement parameters. Independent two-sample t-tests indicated that dyslexics and controls did not differ in FD ($t(34) = -0.08$, $p = 0.933$).

Creation of Functional Connectivity Matrices

Functional connectivity (FC) matrices were created using the CONN functional connectivity toolbox (Whitfield-Gabrieli and Nieto-Castanon, 2012). First, 264 ROIs in spheres with 10-mm diameter were defined as nodes based on a validated parcellation template (Power et al., 2011; Cole et al., 2013). Then, BOLD time series signals corresponding to the conditions of HFC, LFC and figure were separately extracted from each ROI, and were then concatenated over blocks. Nuisance BOLD signal fluctuations from cerebrospinal fluid and white matter were estimated and removed using the anatomical component correction (CompCor) strategy (Behzadi et al., 2007). In addition, head motion (Six motion parameters and six first-order temporal derivatives) as well as the main effect of task were also regressed out. The data were high-pass filtered at 0.008 Hz to preserve

---

[1] http://www.fil.ion.ucl.ac.uk/spm/

task-relevant high-frequency signals. Next, Pearson's correlation coefficients between each pair of regional time series signals were computed and transformed into Fisher's z scores. Following this procedure, undirected and weighted 264 × 264 FC matrices were constructed for each condition and for each participant. Finally, the significant non-zero connections in FC matrices were defined as the significant edges for each condition by performing one-sample t-tests (false discovery rate (FDR) corrected $p < 0.05$) using the GRETNA toolbox (Wang et al., 2015).

Network-Based Statistical Analysis

The network-based statistic (NBS) analysis was applied to identify differences in the functional networks involved in handwriting and drawing between dyslexics and controls. NBS is a non-parametric method that can detect the specific connections within brain networks for the differences between psychological contexts. This approach rejects the null hypothesis on a component-basis controlling for family-wise error (FWE) rate, and thus achieves substantially greater statistical power than mass-univariate testing performed at the edge level (Zalesky et al., 2010). At the group level, independent two-sample t tests were applied for the HFC, LFC and figure condition respectively, because we were particularly interested in the group differences in functional network reconfiguration during handwriting. Factors including sex, age, Raven IQ and FD were included as covariates to account for the potential confounding effect. In addition, to account for the differences in behavioral performance during fMRI scanning, writing duration and writing latency during fMRI were also included as covariates. A mask containing the significant edges across the groups and conditions was applied to the group analysis, ensuring that statistical comparisons were restricted within a same network space (Jiang et al., 2013). A set of supra-threshold connections were first defined ($p < 0.01$), which was used to determine the topological components and their intensity-based sizes (the sum of test statistic values across all connections within a component) (Zalesky et al., 2012; Cao Q. et al., 2013; Spies et al., 2019). Then, non-parametric permutation tests were performed to estimate the significance of each component (permutation times = 5000, family-wise error (FWE) rate corrected $p < 0.05$). For each permutation, the labels of participants were randomized under the null hypothesis without affecting the test statistic (Zalesky et al., 2012). Finally, the corrected p value for a component of a given size was calculated as the proportion of permutations for which the largest component was the same size or greater. To further specify the function of networks, the identified networks were assigned to a well-established brain network partition consisting of 10 well-defined brain systems (frontoparietal, cingulo-opercular, default mode, dorsal attention, ventral attention, auditory, visual, salience, somatomotor and subcortical networks) (Power et al., 2011). Given an edge connects two nodes belonging to a same network, this edge was defined as a within-network functional connectivity. While, when an edge connects two nodes belonging to two different networks, this edge was defined as between-network functional connectivity.

In addition, the hubs were defined as the nodes with a connectivity strength of SD greater than the mean strength across all nodes in the network (Sporns et al., 2007; Liu et al., 2018). Node strength is analogous to node degree in weighted networks, which is defined as the sum of edge weights (i.e., Fisher's z scores) attached to a node (Fornito et al., 2016). The results were visualized using the BrainNet Viewer toolbox (Xia et al., 2013).

Correlation Between Network Connectivity and Behavioral Performance

Partial correlation analysis was conducted between the connectivity strength of the networks showing between-group differences and the performance of handwriting, reading and cognitive tests, controlling for age, Raven IQ and FD. The connectivity strength was defined as the average of connectivity weights (Fisher's z scores) of all edges of the networks. The statistical significance was set at $p < 0.05$, uncorrected for the multiple comparisons.

Validation Analysis

To evaluate the robustness of our results, two validation procedures were performed. First, we repeated the whole data analysis by using the FC matrices that were survived at a less stringent threshold of $p < 0.05$, uncorrected for multiple comparisons. Second, we reanalyzed the NBS analysis using NBS extent, in which the size of a network component is defined as the total number of connections it comprises.

RESULTS

Behavioral Results

Out-Scanner Behavioral Performance

The results of reading, handwriting and cognitive skills tests are presented in **Table 1**. Results indicated that dyslexics compared to controls showed inferior performance in reading fluency, phonological awareness and orthographic awareness. In addition, dyslexics showed poorer handwriting fluency (both characters and digits) than controls. However, dyslexics and controls did not differ in handwriting speed and quality in the pen-and-paper copying task. Finally, we found that, compared to controls, dyslexics exhibited reduced phonological working memory span, but exhibited intact sustain attention ability.

In-Scanner Behavioral Performance

The average writing duration and latency during fMRI scanning are presented in **Figure 1**. For writing duration, the interaction between group and stimulus type ($F(2, 68) = 1.18$, $p = 0.314$) and the main effect of group ($F(1, 34) = 0.01$, $p = 0.930$) were not significant. The main effect of condition was marginally significant ($F(2,68) = 3.01$, $p = 0.056$). *Post hoc* pairwise comparisons showed that the duration of copying HFCs was shorter than that of copying LFCs ($p = 0.008$), but there was no significant difference between copying HFCs and drawing figures ($p = 0.119$), or between copying LFCs and drawing figures ($p = 0.658$) (**Figure 1A**). For writing latency, the interaction between group and stimulus type ($F(2, 68) = 0.80$,

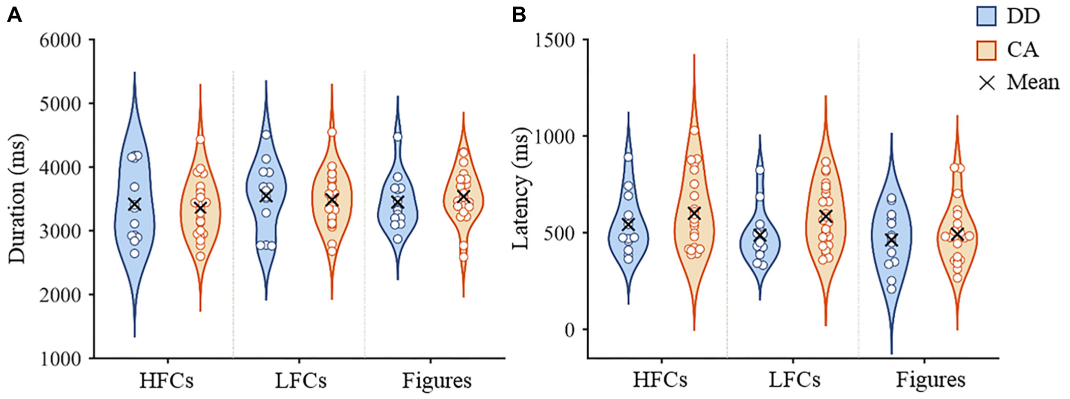


FIGURE 1 | In-scanner performance in the copying and drawing tasks. Violin plots of writing duration **(A)** and writing latency **(B)** of dyslexic children and controls during copying HFCs and LFCs and drawing figures. HFCs = high-frequency characters, LFCs = low-frequency characters, DD = developmental dyslexia and CA = chronological-age matched control group.

p = 0.455) and the main effect of group (F(1,34) = 1.82, p = 0.186) were not significant. The main effect of condition was significant (F(2,68) = 6.23, p = 0.003). *Post hoc* pairwise comparisons showed that the latency of copying HFCs was longer than that of drawing figures (p = 0.001), but there was no significant difference between copying HFCs and copying LFCs (p = 0.126) and between copying LFCs and drawing figures (p = 0.064) (**Figure 1B**).

Network-Based Statistic Analysis Results

The NBS analysis revealed that controls showed greater connectivity than dyslexic children in a functional brain network during copying HFCs, consisting of 66 nodes and 68 edges (**Figure 2A**). According to the functional network division (Power et al., 2011), this network can be grouped as internetwork connectivity between the sensory-motor network (SMN) and visual network (VN), between the default mode network (DMN) and ventral attention network (VAN), between the DMN and frontal-parietal network (FPN), and between the SMN and salience network (SAN) (**Figure 2C** and **Supplementary Figure 1**). Three nodes in the SMN (two nodes in the right medial frontal gyrus and the left precuneus), a node in the DMN (the right middle temporal gyrus), a node in the VAN (the right superior temporal gyrus) and a node in the FPN (the right inferior temporal gyrus) were identified as hubs (**Supplementary Table 1**).

During copying LFCs, controls also showed greater connectivity than dyslexic children in a functional brain network consisting of 48 nodes and 48 edges (**Figure 2B**). This network mainly encompassed internetwork connectivity between the SMN and VN, between the DMN and VAN, between the VN and DMN and between the VN and cingulo-opercular network (CON) (**Figure 2D** and **Supplementary Figure 1**). A node in the DMN (the right precuneus), a node in the VAN (the right superior temporal gurus) and a node in the SMN (the left precentral gyrus) were identified as hubs (**Supplementary Table 1**).

However, no significant differences in functional brain networks were detected between the two groups during drawing figures.

Correlation Between Network Connectivity and Behavioral Performance

Correlation analysis revealed that connectivity strength of the functional networks showing group differences was positively correlated with writing speed of digits (HFCs: $r = 0.37$, $p = 0.032$; LFCs: $r = 0.38$, $p = 0.027$), reading fluency (HFCs: $r = 0.53$, $p = 0.002$; LFCs: $r = 0.60$, $p < 0.001$), orthographic awareness (ACC: HFCs: $r = 0.30$, $p = 0.098$; LFCs: $r = 0.39$, $p = 0.026$; RT: HFCs: $r = -0.41$, $p = 0.021$; LFCs: $r = -0.39$, $p = 0.026$) and phonological working memory (HFCs: $r = 0.40$, $p = 0.023$; LFCs: $r = 0.37$, $p = 0.033$).

Validation Results

We found that when using a less stringent threshold for determining the FC metrics, dyslexics showed weaker functional connectivity than controls in a network involving in the VN, SMN, DMN, VAN and FPN in both the HFC and LFC conditions. This result was similar to the reported findings, despite the slight differences in connectivity strength (**Supplementary Figure 2**). Similarly, in the context of using an alternative NBS estimation approach, between-group differences were also identified in a similar functional network as the reported findings in the two handwriting conditions (**Supplementary Figure 3**). Collectively, these results indicated that the differences in functional networks between dyslexic and control children during copying HFCs and LFCs were largely reproducible.

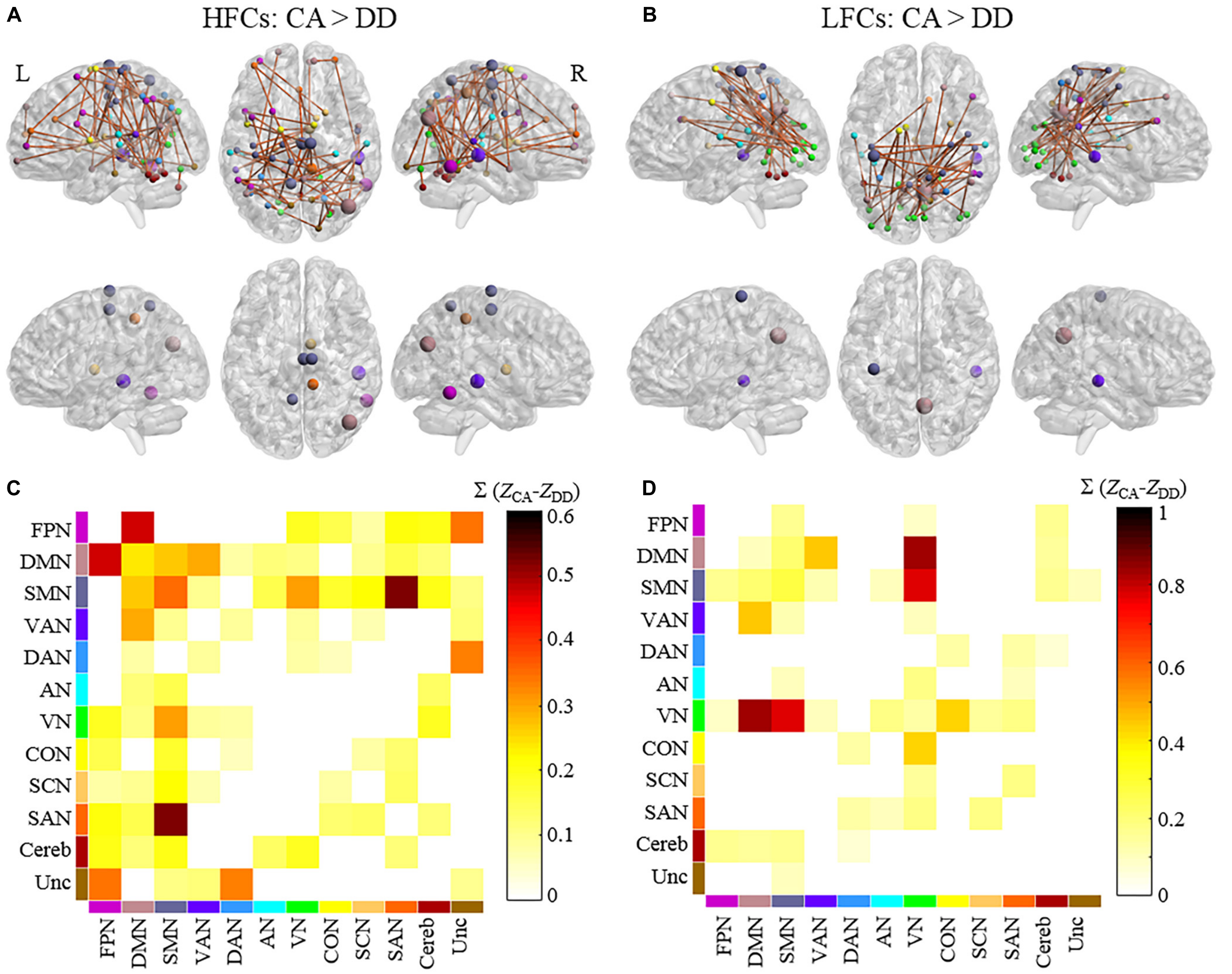


FIGURE 2 | The differences in functional brain networks between dyslexics and controls during handwriting. Brain networks and hubs of the network showing stronger connectivity in controls than in dyslexics during copying HFCs **(A)** and LFCs **(B)**. The colors of the nodes indicate the network to which they belong. The size of hubs is proportional to their node strength. The matrix plots showing the connectivity patterns within/between each pair of networks during copying HFCs **(C)** and LFCs **(D)**. The color of each element in the matrices represents the sum of the differences in connectivity strength of all the edges for the connected networks. L = left and R = right. HFCs = high-frequency characters, LFCs = low-frequency characters, DD = developmental dyslexia, CA = chronological-age matched control group. FPN = frontal-parietal network, DMN = default mode network, SMN = somatomotor network, VAN = ventral attention network, DAN = dorsal attention network, AN = auditory network, VN = visual network, CON = cingulo-opercular network, SCN = subcortical network, SAN = salience network, Cereb = cerebellum and Unc = Uncertain.

DISCUSSION

Using a network analysis approach, the study identified the aberrant functional brain networks associated with the handwriting deficit in Chinese children with dyslexia. We found that dyslexics showed reduced functional connectivity in large-scale brain networks during handwriting involving the VN, SMN, DMN, VAN and FPN, suggesting that task-relevant sensor-motor networks and domain-general executive control networks convergently contribute to the handwriting deficit in DD. Moreover, we found that the between-group differences in functional networks varied between the high-frequency and low-frequency conditions, suggesting that dyslexics' handwriting deficit was mediated by linguistical variables.

Behaviorally, we found that dyslexics showed reduced handwriting fluency relative to controls. This result is in line with previous research indicating that dyslexics showed impaired handwriting fluency (Arfe et al., 2020). However, dyslexics and controls showed no statistically significant differences in writing speed and quality in the pen-and-paper copying task. This result is inconsistent with previous findings showing reduced handwriting speed during copying tasks in dyslexics relative to controls (Lam et al., 2011; Meng et al., 2019). These findings suggest that the fluency task paradigm may be more sensitive to capture the insufficient automaticity of handwriting in dyslexics than the copying tasks in behavioral measures. Alternatively, because dyslexics showed the trend of decreasing handwriting speed and quality in the copying tasks, we speculated that the failure to reach statistical significances may be associated with the small sample size.

In line with our hypothesis, the brain network analysis revealed that dyslexics and controls differed in functional connectivity in a distributed brain network supporting visual, motoric and cognitive executive processes. Furthermore, behavioral recordings during fMRI scanning showed no differences in task-performance between the two groups of participants, excluding the possibility that the observed differences in functional brain networks are just derived from task difficulty. These findings suggest that the handwriting problem in dyslexics is not derived from a low-level perception and motor dysfunction, but instead from a failure of the integration of cognitive, sensory and motor systems. This argument is supported by the brain-behavior correlation analysis showing that the brain networks showing between-group differences are related to the skills of handwriting speed, reading and working memory.

Another critical finding of this study is that we did not observe between-group differences in the drawing condition, suggesting that the observed functional network abnormalities in dyslexics are specific to handwriting processing. Handwriting and drawing skills share several basic sensory and motor processes, which are supported by an overlapped brain circuit (Yuan and Brown, 2015). However, we found the brain basis specific to the handwriting deficit in DD, suggesting that the brain systems for handwriting and drawing have been dissociation in the middle age of children. This view is supported by a developmental study that reported that children around 10 years old have established the brain system of handwriting (Palmis et al., 2021).

The connectivity profiles of the networks showing group differences were characterized by referring a functional network template (Power et al., 2011). First, we found that dyslexics showed altered internetwork connectivity between the VN (including the nodes of the fusiform gyrus and lingual gyrus) and the SMN (including the nodes of the precentral gyrus and postcentral gyrus), irrespective of character frequency. This result is in accordance with previous findings of reduced brain activation in the visual-orthographic regions (Cao et al., 2018; Yang et al., 2022) and the visual perception regions (Yang et al., 2021) in Chinese dyslexic children. The VN is thought to support visual analysis of Chinese characters during handwriting (Wu et al., 2012; Cao Q. et al., 2013). Consistent with this interpretation, correlation analysis showed that network connectivity was positively correlated with orthographic awareness. Moreover, the SMN has been widely identified to be engaged in handwriting. Functionally, the bilateral primary motor regions are involved in motor control (Planton et al., 2013), while the medial frontal gyrus (including the SMA) serves the process of Chinese writing sequence (Zhang Z. et al., 2021) or motor response preparation (Planton et al., 2013). Consequently, the coupling between the SMN and VN is recruited to support the coordination of visual and motor controls necessary for handwriting. Alternatively, the SMN and VN may be recruited to serve the sensory feedbacks during handwriting, which plays an important role in optimizing motor output (Peterka, 2002). Because handwriting has not yet become fully automatic in children, an attentional controlled movement pattern is engaged in handwriting, which highly relies on visual and somatomotor feedbacks (Marquardt et al., 1999). Thus, the reduced connectivity between the VN and SMN may affect the functional integration engaged in the sensory feedback processing, thus slowing down handwriting speed or wrecking handwriting quality in dyslexics.

Another important finding of the present study is the disrupted connectivity of the DMN with multiple functional networks during handwriting in Chinese dyslexics. This result is consistent with the view that the DMN serves as an “integrative hub” for the cross talk between functional brain networks (Braga et al., 2013). First, the connectivity between the DMN and VAN was decreased in dyslexics relative to controls, which was observed in both high-frequency and low-frequency conditions. Although the altered connectivity of the DMN has been repeatedly identified in DD during resting and task states (Finn et al., 2014; Schurz et al., 2015), the specific role of the DMN in dyslexia is still unclear. In this sense, the identified association between the disruption of the DMN and handwriting deficit hints on a possible role of the DMN in dyslexia. The DMN is traditionally regarded as a task-negative functional network, whose activity is increased in internally oriented cognitive states (Raichle et al., 2001; Fox et al., 2005). According to this account of the DMN, dyslexia has been postulated to be associated with the failure of disengaging the DMN from reading-related networks (Boros et al., 2016; Cao et al., 2017). However, a growing body of empirical evidence suggests that the DMN is actively involved in goal-directed cognitive processing, such as task shift (Crittenden et al., 2015) and working memory (Spreng et al., 2014). Specifically, the DMN has been found to be involved in the application of learned information to make predictions in decision-making (Vatansever et al., 2017) and in the integration of external goals and internal representation (Spreng et al., 2014). Moreover, a recent study has demonstrated that the DMN encodes information associated with ongoing cognition for the memory-based guide in automated processing (Sormaz et al., 2018). Based on these findings, we posited that the DMN may play a role in high-level executive control for the integration of different brain systems involved in handwriting. In addition, the DMN may encodes the long-term representation of handwriting rules resulted from learning and practice. In line with this view, a prior study has demonstrated that the DMN is involved in visual-motor learning (Eryurek et al., 2022). On the other hand, the VAN is an attentional control network that serves the processing of unexpected stimuli, reflecting the bottom-up control processing, consistent with previous findings of the dysfunction of ventral and dorsal attentional networks in dyslexics (Meri et al., 2020). Previous studies have also found that the VAN was positively correlated with DMN during childhood (Barber et al., 2013). Thus, the disruption of the internetwork connectivity between the DMN and VAN may impede the cross talk between the internal representation of handwriting rules and external task contexts during handwriting in DD.

In addition to the common brain network abnormalities across the HFCs and LFCs conditions, we also found some differences in network connectivity between the two conditions. The frequency effect is a typical lexical variable that has been found to influence orthographic access during handwriting, and thus this result suggests the impact of linguistic factors to the handwriting deficit in DD. First, we found that the decreased connectivity between the DMN and the FPN in dyslexics was more evident in the high-frequency condition relative to the low-frequency condition. The FPN is a high-order control network for cognitive processes that flexibly interacts with other networks adapted to task demands (Cole et al., 2013). Coupled with the integrative role of the DMN in cognitive tasks (Braga et al., 2013; Wang et al., 2021), the internetwork connectivity between the FPN and DMN represents a neural circuit for executive control (Wang et al., 2021). Consistent with this interpretation, it has been reported that the connectivity between the FPN and DMN increases under the context of intentional speed-control processing in handwriting (Li et al., 2021). The linguistic information of the characters (phonologic or semantic) is more likely to be activated for the HFCs relative to LFCs, which may play an interferential role in the copying task because the orthographic forms have already been presented and thus the phonological or semantic information is not necessary. The disrupted connectivity between the FPN and DMN may hinder the adaptive control process during handwriting familiar HFCs in dyslexic children. In accordance with this explanation, functional connectivity between the DMN and executive control regions has been found to support the goal-directed semantic retrieval (Krieger-Redwood et al., 2016).

In addition, we found that dyslexics showed more pronounced abnormality in functional connectivity of the VN with the DMN and the CON in the LFCs condition. This result agrees with a resting-state functional connectivity study reporting the abnormalities of functional connectivity between the visual networks and prefrontal attention areas and the connectivity between the DMN and VN (Finn et al., 2014). The CON is a vital network hub of executive control that is thought to support the maintenance of task goals, the adjustments for feedback control and error monitor (Power et al., 2011). The reorganization of the CON has been found to be associated with reading improvement in DD (Horowitz-Kraus et al., 2015). Similarly, the DMN has been reported to generate top-down predictions by integrated memory-based information for automated cognitive processing (Vatansever et al., 2017). Thus, the reduced functional connectivity of the VN with the DMN and CON may reflect the problematic regulation from the DMN and CON to unimodal visual processing during handwriting in dyslexics. This explanation is favored by previous studies reporting the visual attention deficit in DD (Taran et al., 2022). Because the low-frequency characters were less familiar to the participants, the visual-orthographic processing may be more demanding in the LFCs condition relative to the HFCs condition, requiring more top-down control from the executive control regions to visual regions. Consequently, we found that the specific impairments in functional connectivity of the visual networks with the executive control networks during writing infrequent characters in dyslexics.

CONCLUSION

Using a network analysis method, this study revealed that the handwriting deficit in DD was associated with the abnormalities of network connectivity in multiple brain networks involved in visual-orthographic, motor and executive control processes. Our findings advance our understanding of the brain basis of DD.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

ETHICS STATEMENT

The study was approved by the Ethics Committee of the Institute of Psychology at the Chinese Academy of Sciences. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

ZL: conceptualization, methodology, visualization, validation, formal analysis, investigation and writing—original draft. JL: methodology, investigation and formal analysis, and writing—original draft. H-YB: conceptualization, project administration and writing—review, and editing. MX: conceptualization, formal analysis, writing—review and editing, and supervision. YY: conceptualization, methodology, funding acquisition, resources, writing—review and editing, and supervision. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (Nos. 31800954 and 32171054) and the Beijing Natural Science Foundation (No. 5222027).

ACKNOWLEDGMENTS

We thank all the participants and their families who participated in this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2022.919440/full#supplementary-material

REFERENCES

Afonso, O., Suarez-Coalla, P., and Cuetos, F. (2020). Writing impairments in Spanish Children with developmental dyslexia. *J. Learn. Disabil.* 53, 109–119.

Amalric, M., and Dehaene, S. (2016). Origins of the brain networks for advanced mathematics in expert mathematicians'. *Proc. Natl. Acad. Sci. U.S.A.* 113, 4909–4917. doi: 10.1073/pnas.1603205113

Arfe, B., Corato, F., Pizzocaro, E., and Merella, A. (2020). The effects of script and orthographic complexity on the handwriting and spelling performance of children with dyslexia. *J. Learn. Disabil.* 53, 96–108. doi: 10.1177/0022219419892845

Barber, A. D., Caffo, B. S., Pekar, J. J., and Mostofsky, S. H. (2013). Developmental changes in within- and between-network connectivity between late childhood and adulthood. *Neuropsychologia* 51, 156–167. doi: 10.1016/j.neuropsychologia.2012.11.011

Behzadi, Y., Restom, K., Liau, J., and Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage* 37, 90–101. doi: 10.1016/j.neuroimage.2007.04.042

Boets, B., Op de Beeck, H. P., Vandermosten, M., Scott, S. K., Gillebert, C. R., Mantini, D., et al. (2013). Intact but less accessible phonetic representations in adults with dyslexia. *Science* 342, 1251–1254. doi: 10.1126/science.1244333

Boros, M., Anton, J. L., Pech-Georgel, C., Grainger, J., Szwed, M., and Ziegler, J. C. (2016). Orthographic processing deficits in developmental dyslexia: beyond the ventral visual stream. *Neuroimage* 128, 316–327. doi: 10.1016/j.neuroimage.2016.01.014

Braga, R. M., Sharp, D. J., Leeson, C., Wise, R. J. S., and Leech, R. (2013). Echoes of the brain within default mode, association, and heteromodal cortices. *J. Neurosci.* 33, 14031–14039. doi: 10.1523/JNEUROSCI.0570-13.2013

Bullmore, Ed, and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 186–198.

Cao, F., Bitan, T., Chou, T. L., Burman, D. D., and Booth, J. R. (2006). Deficient orthographic and phonological representations in children with dyslexia revealed by brain activation patterns. *J. Child Psychol. Psychiatry* 47, 1041–1050. doi: 10.1111/j.1469-7610.2006.01684.x

Cao, F., Vu, M., Chan, D. H. L., Lawrence, J. M., Harris, L. N., Guan, Q., et al. (2013). Writing affects the brain network of reading in Chinese: a functional magnetic resonance imaging study. *Hum. Brain Mapp.* 34, 1670–1684.

Cao, F., Yan, X., Spray, G. J., Liu, Y., and Deng, Y. (2018). Brain mechanisms underlying visuo-orthographic deficits in children with developmental dyslexia. *Front. Hum. Neurosci.* 12:490. doi: 10.3389/fnhum.2018.00490

Cao, F., Yan, X., Wang, Z., Liu, Y., Wang, J., Spray, G. J., et al. (2017). Neural signatures of phonological deficits in Chinese developmental dyslexia. *NeuroImage (Orlando, Fla.)* 146, 301–311.

Cao, Q., Shu, N., An, L., Wang, P., Sun, L., Xia, M. R., et al. (2013). Probabilistic diffusion tractography and graph theory analysis reveal abnormal white matter structural connectivity networks in drug-naive boys with attention deficit/hyperactivity disorder. *J. Neurosci.* 33, 10676–10687. doi: 10.1523/JNEUROSCI.4793-12.2013

Cole, M. W., Reynolds, J. R., Power, J. D., Repovs, G., Anticevic, A., and Braver, T. S. (2013). Multi-task connectivity reveals flexible hubs for adaptive task control. *Nat. Neurosci.* 16, 1348–1355. doi: 10.1038/nn.3470

Crittenden, B. M., Mitchell, D. J., and Duncan, J. (2015). Recruitment of the default mode network during a demanding act of executive control. *eLife* 2015, e6481–e6481.

Eryurek, K., Ulasoglu-Yildiz, C., Matur, Z., Emre Öge, A., Gürvit, H., and Demiralp, T. (2022). Default mode and dorsal attention network involvement in visually guided motor sequence learning. *Cortex* 146, 89–105. doi: 10.1016/j.cortex.2021.10.006

Feng, X., Altarelli, I., Monzalvo, K., Ding, G., Ramus, F., Shu, H., et al. (2020). A universal reading network and its modulation by writing system and reading ability in French and Chinese children. *eLife* 9:e54591. doi: 10.7554/eLife.54591

Finn, E. S., Shen, X., Holahan, J. M., Scheinost, D., Lacadie, C., Papademetris, X., et al. (2014). Disruption of functional networks in dyslexia: a whole-brain, data-driven analysis of connectivity. *Biol. Psychiatry* 76, 397–404. doi: 10.1016/j.biopsych.2013.08.031

Fornito, A., Zalesky, A., and Bullmore, E. T. (2016). *Fundamentals of Brain Network Analysis*. Cambridge, MA: Academic Press.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., and Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl. Acad. Sci. U.S.A.* 102, 9673–9678.

Gimenez, P., Bugescu, N., Black, J. M., Hancock, R., Pugh, K., Nagamine, M., et al. (2014). Neuroimaging correlates of handwriting quality as children learn to read and write. *Front. Hum. Neurosci.* 8:155. doi: 10.3389/fnhum.2014.00155 doi: 10.3389/fnhum.2014.00155

Gosse, C., Dricot, L., and Van Reybroeck, M. (2022). Evidence of graphomotor dysfunction in children with dyslexia: a combined behavioural and fMRI experiment. *Cortex* 148, 68–88. doi: 10.1016/j.cortex.2021.11.021

Gosse, C., and Van Reybroeck, M. (2020). Do children with dyslexia present a handwriting deficit? Impact of word orthographic and graphic complexity on handwriting and spelling performance. *Res. Dev. Disabil.* 97:103553. doi: 10.1016/j.ridd.2019.103553

Graham, S., Aitken, A. A., Hebert, M., Camping, A., Santangelo, T., Harris, K. R., et al. (2021). Do children with reading difficulties experience writing difficulties? A meta-analysis. *J. Educ. Psychol.* 113, 1481–1506.

Hoeft, F., Meyler, A., Hernandez, A., Juel, C., Taylor-Hill, H., Martindale, J. L., et al. (2007). Functional and morphometric brain dissociation between dyslexia and reading ability. *Proc. Natl. Acad. Sci. U.S.A.* 104, 4234–4239. doi: 10.1073/pnas.0609399104

Horowitz-Kraus, T., Toro-Serey, C., and DiFrancesco, M. (2015). Increased resting-state functional connectivity in the cingulo-opercular cognitive-control network after intervention in children with reading difficulties. *PLoS One* 10:e0133762. doi: 10.1371/journal.pone.0133762

Jiang, G., Wen, X., Qiu, Y., Zhang, R., Wang, J., Li, M., et al. (2013). Disrupted topological organization in whole-brain functional networks of heroin-dependent individuals: a resting-state FMRI study. *PLoS One* 8:e82715. doi: 10.1371/journal.pone.0082715

Kalindi, S. C., McBride, C., Tong, X., Wong, N. L. Y., Chung, K. H. K., and Lee, C. Y. (2015). Beyond phonological and morphological processing: pure copying as a marker of dyslexia in Chinese but not poor reading of English. *Ann. Dyslexia* 65, 53–68. doi: 10.1007/s11881-015-0097-8

Kandel, S., and Perret, C. (2015). How does the interaction between spelling and motor processes build up during writing acquisition?. *Cognition* 136, 325–336. doi: 10.1016/j.cognition.2014.11.014

Krieger-Redwood, K., Jefferies, E., Karapanagiotidis, T., Seymour, R., Nunes, A., Ang, J. W. A., et al. (2016). Down but not out in posterior cingulate cortex: deactivation yet functional coupling with prefrontal cortex during demanding semantic cognition. *Neuroimage* 141, 366–377.

Lam, S. S., Au, R. K., Leung, H. W., and Li-Tsang, C. W. (2011). Chinese handwriting performance of primary school children with dyslexia. *Res. Dev. Disabil.* 32, 1745–1756.

Li, J., Hong, L., Bi, H. Y., and Yang, Y. (2021). Functional brain networks underlying automatic and controlled handwriting in Chinese. *Brain Lang.* 219:104962.

Liu, X., Gao, Y., Di, Q., Hu, J., Lu, C., Nan, Y., et al. (2018). Differences between child and adult large-scale functional brain networks for reading tasks. *Hum. Brain Mapp.* 39, 662–679.

Marquardt, C., Gentz, W., and Mai, N. (1999). 'Visual control of automated handwriting movements. *Exp. Brain Res.* 128, 224–228.

Martlew, M. (1992). Handwriting and spelling: dyslexic children's abilities compared with children of the same chronological age and younger children of the same spelling level. *Br. J. Educ. Psychol.* 62, 375–390. doi: 10.1111/j.2044-8279.1992.tb01030.x

Meng, Z. L., Wydell, T. N., and Bi, H. Y. (2019). Visual-motor integration and reading Chinese in children with/without dyslexia. *Read. Writ.* 32, 493–510.

Menghini, D., Hagberg, G. E., Caltagirone, C., Petrosini, L., and Vicari, S. (2006). Implicit learning deficits in dyslexic adults: an fMRI study. *Neuroimage* 33, 1218–1226.

Meri, R., Farah, R., and Horowitz-Kraus, T. (2020). Children with dyslexia utilize both top-down and bottom-up networks equally in contextual and isolated word reading. *Neuropsychologia* 147:107574. doi: 10.1016/j.neuropsychologia.2020.107574

Moeller, S., Yacoub, E., Olman, C. A., Auerbach, E., Strupp, J., Harel, N., et al. (2010). Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. *Magn. Reson. Med.* 63, 1144–1153. doi: 10.1002/mrm.22361

Nicolson, R. I., Fawcett, A. J., Berry, E. L. I, Jenkins, H., Dean, P., and Brooks, D. J. (1999). Association of abnormal cerebellar activation with motor learning difficulties in dyslexic adults. *Lancet* 353, 1662–1667. doi: 10.1016/S0140-6736(98)09165-X

Olulade, O. A., Napoliello, E. M., and Eden, G. F. (2013). 'Abnormal visual motion processing is not a cause of dyslexia. *Neuron* 79, 180–190.

Pagliarini, E., Guasti, M. T., Toneatto, C., Granocchio, E., Riva, F., Sarti, D., et al. (2015). Dyslexic children fail to comply with the rhythmic constraints of handwriting. *Hum. Mov. Sci.* 42, 161–182. doi: 10.1016/j.humov.2015.04.012

Palmis, S., Velay, J. L., Habib, M., Anton, J. L., Nazarian, B., Sein, J., et al. (2021). The handwriting brain in middle childhood. *Dev. Sci.* 24, e13046.

Perfetti, C. A., and Harris, L. N. (2013). Universal reading processes are modulated by language and writing system. *Lang. Learn. Dev.* 9, 296–316.

Peterka, R. J. (2002). Sensorimotor integration in human postural control. *J. Neurophysiol.* 88, 1097–1118.

Planton, S., Jucla, M., Roux, F. E., and Démonet, J. F. (2013). The "handwriting brain": a meta-analysis of neuroimaging studies of motor versus orthographic processes. *Cortex* 49, 2772–2787. doi: 10.1016/j.cortex.2013.05.011

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., and Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59, 2142–2154. doi: 10.1016/j.neuroimage.2011.10.018

Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., et al. (2011). Functional network organization of the human brain. *Neuron* 72, 665–678.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., and Shulman, G. L. (2001). A default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98, 676–682.

Schurz, M., Wimmer, H., Richlan, F., Ludersdorfer, P., Klackl, J., and Kronbichler, M. (2015). Resting-state and task-based functional brain connectivity in developmental dyslexia. *Cereb. Cortex* 25, 3502–3514.

Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Mencl, W. E., Fulbright, R. K., Skudlarski, P., et al. (2002). Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biol. Psychiatry* 52, 101–110.

Shaywitz, S. E. (1998). Dyslexia. *N. Engl. J. Med.* 338, 307–312.

Snyder, P. J., and Harris, L. J. (1993). Handedness, sex, and familial sinistrality effects on spatial tasks. *Cortex* 29, 115–134.

Sormaz, M., Murphy, C., Wang, H. T., Hymers, M., Karapanagiotidis, T., Poerio, G., et al. (2018). Default mode network can support the level of detail in experience during active task states. *Proc. Nat. Acad. Sci. U.S.A.* 115, 9318–9323.

Spies, M., Klobl, M., Hoflich, A., Hummer, A., Vanicek, T., Michenthaler, P., et al. (2019). Association between dynamic resting-state functional connectivity and ketamine plasma levels in visual processing networks. *Sci. Rep.* 9:11484. doi: 10.1038/s41598-019-46702-x

Sporns, O. (2011). The human connectome: a complex network. *Ann. N. Y. Acad. Sci.* 1224, 109–125.

Sporns, O., Honey, C. J., and Kotter, R. (2007). Identification and classification of hubs in brain networks. *PLoS One* 2:e1049. doi: 10.1371/journal.pone.0001049

Spreng, R. N., DuPre, E., Selarka, D., Garcia, J., Gojkovic, S., Mildner, J., et al. (2014). Goal-congruent default network activity facilitates cognitive control. *J. Neurosci.* 34, 14108–14114. doi: 10.1523/JNEUROSCI.2815-14.2014

Sumner, E., Connelly, V., and Barnett, A. L. (2012). Children with dyslexia are slow writers because they pause more often and not because they are slow at handwriting execution. *Read. Writ.* 26, 991–1008.

Sumner, E., Connelly, V., and Barnett, A. L. (2014). The influence of spelling ability on handwriting production: children with and without dyslexia. *J. Exp. Psychol. Learn. Mem. Cogn.* 40, 1441–1447.

Tam, F., Churchill, N. W., Strother, S. C., and Graham, S. J. (2011). A new tablet for writing and drawing during functional MRI. *Hum. Brain Mapp.* 32, 240–248.

Tan, L. H., Spinks, J. A., Eden, G. F., Perfetti, C. A., and Siok, W. T. (2005). Reading depends on writing, in Chinese. *Proc. Natl. Acad. Sci. U.S.A.* 102, 8781–8785.

Tan, L. H., Xu, M., Chang, C. Q., and Siok, W. T. (2013). China's language input system in the digital age affects children's reading development. *Proc. Natl. Acad. Sci.* 110, 1119–1123. doi: 10.1073/pnas.1213586110

Taran, N., Farah, R., DiFrancesco, M., Altaye, M., Vannest, J., Holland, S., et al. (2022). The role of visual attention in dyslexia: behavioral and neurobiological evidence. *Hum. Brain Mapp.* 43, 1720–1737. doi: 10.1002/hbm.25753

van der Mark, S., Bucher, K., Maurer, U., Schulz, E., Brem, S., Buckelmuller, J., et al. (2009). Children with dyslexia lack multiple specializations along the visual word-form (VWF) system. *Neuroimage* 47, 1940–1949. doi: 10.1016/j.neuroimage.2009.05.021

van der Mark, S., Klaver, P., Bucher, K., Maurer, U., Schulz, E., Brem, S., et al. (2011). The left occipitotemporal system in reading: disruption of focal fMRI connectivity to left inferior frontal and inferior parietal language areas in children with dyslexia. *Neuroimage* 54, 2426–2436. doi: 10.1016/j.neuroimage.2010.10.002

Vatansever, D., Menon, D. K., and Stamatakis, E. A. (2017). Default mode contributions to automated information processing. *Proc. Natl. Acad. Sci. PNAS* 114, 12821–12826.

Wang, H. (1986). *Modern Chinese Frequency Dictionary*. Beijing: Beijing Language Institute Press.

Wang, J., Wang, X., Xia, M., Liao, X., Evans, A., and He, Y. (2015). Corrigendum: GRETNA: a graph theoretical network analysis toolbox for imaging connectomics. *Front. Hum. Neurosci.* 9:458. doi: 10.3389/fnhum.2015.00458

Wang, X., Gao, Z., Smallwood, J., and Jefferies, E. (2021). Both default and multiple-demand regions represent semantic goal information. *J. Neurosci.* 41, 3679–3691. doi: 10.1523/JNEUROSCI.1782-20.2021

Whitfield-Gabrieli, S., and Nieto-Castanon, A. (2012). Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connect* 2, 125–141. doi: 10.1089/brain.2012.0073

Wu, C.-Y., Ho, M. H. R., and Chen, S. H. A. (2012). A meta-analysis of fMRI studies on Chinese orthographic, phonological, and semantic processing. *NeuroImage (Orlando, Fla.)* 63, 381–391.

Xia, M., Wang, J., and He, Y. (2013). BrainNet viewer: a network visualization tool for human brain connectomics. *PLoS One* 8:e68910. doi: 10.1371/journal.pone.0068910

Yang, X., Zhang, J., Lv, Y., Wang, F., Ding, G., Zhang, M., et al. (2021). Failure of resting-state frontal–occipital connectivity in linking visual perception with reading fluency in Chinese children with developmental dyslexia. *Neuroimage* 233:117911. doi: 10.1016/j.neuroimage.2021.117911

Yang, Y., Bi, H. Y., Long, Z. Y., and Tao, S. (2013). Evidence for cerebellar dysfunction in Chinese children with developmental dyslexia: an fMRI study. *Int. J. Neurosci.* 123, 300–310. doi: 10.3109/00207454.2012.756484

Yang, Y., Tam, F., Graham, S. J., Sun, G., Li, J., Gu, C., et al. (2020). Men and women differ in the neural basis of handwriting. *Hum. Brain Mapp.* 41, 2642–2655.

Yang, Y., Zhang, J., Meng, Z. L., Qin, L., Liu, Y., and Bi, H. Y. (2018). Neural correlates of orthographic access in Mandarin Chinese writing: an fMRI study of the word-frequency effect. *Front. Behav. Neurosci.* 12:288. doi: 10.3389/fnbeh.2018.00288

Yang, Y., Zuo, Z., Tam, F., Graham, S. J., Li, J., Ji, Y., et al. (2022). The brain basis of handwriting deficits in Chinese children with developmental dyslexia. *Dev. Sci.* 25:e13161. doi: 10.1111/desc.13161

Yuan, Y., and Brown, S. (2015). Drawing and writing: an ALE meta-analysis of sensorimotor activations. *Brain Cogn.* 98, 15–26. doi: 10.1016/j.bandc.2015.05.004

Zalesky, A., Cocchi, L., Fornito, A., Murray, M. M., and Bullmore, E. (2012). Connectivity differences in brain networks'. *Neuroimage* 60, 1055–1062.

Zalesky, A., Fornito, A., and Bullmore, E. T. (2010). Network-based statistic: identifying differences in brain networks. *Neuroimage* 53, 1197–1207.

Zhang, J., Liu, L., Li, H., Feng, X., Zhang, M., Liu, L., et al. (2021). Large-scale network topology reveals brain functional abnormality in Chinese dyslexic children. *Neuropsychologia* 157:107886. doi: 10.1016/j.neuropsychologia.2021.107886

Zhang, M., Xie, W., Xu, Y., and Meng, X. (2018). Auditory temporal perceptual learning and transfer in Chinese-speaking children with developmental dyslexia. *Res. Dev. Disabil.* 74, 146–159. doi: 10.1016/j.ridd.2018.01.005

Zhang, Z., Yuan, Q., Liu, Z., Zhang, M., Wu, J., Lu, C., et al. (2021). The cortical organization of writing sequence: evidence from observing Chinese characters in motion. *Brain Struc. Funct.* 226, 1627–1639.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Neuroscience

ORIGINAL RESEARCH
published: 19 July 2022
doi: 10.3389/fnins.2022.921977

Brain Source Correlates of Speech Perception and Reading Processes in Children With and Without Reading Difficulties

Najla Azaiez[1]*, Otto Loberg[2], Jarmo A. Hämäläinen[1,3] and Paavo H. T. Leppänen[1,3]

[1] *Department of Psychology, Faculty of Education and Psychology, University of Jyväskylä, Jyväskylä, Finland,* [2] *Department of Psychology, Faculty of Science and Technology, Bournemouth University, Bournemouth, United Kingdom,* [3] *Department of Psychology, Jyväskylä Center for Interdisciplinary Brain Research, University of Jyväskylä, Jyväskylä, Finland*

Neural correlates in reading and speech processing have been addressed extensively in the literature. While reading skills and speech perception have been shown to be associated with each other, their relationship remains debatable. In this study, we investigated reading skills, speech perception, reading, and their correlates with brain source activity in auditory and visual modalities. We used high-density event-related potentials (ERPs), fixation-related potentials (FRPs), and the source reconstruction method. The analysis was conducted on 12–13-year-old schoolchildren who had different reading levels. Brain ERP source indices were computed from frequently repeated Finnish speech stimuli presented in an auditory oddball paradigm. Brain FRP source indices were also computed for words within sentences presented in a reading task. The results showed significant correlations between speech ERP sources and reading scores at the P100 (P1) time range in the left hemisphere and the N250 time range in both hemispheres, and a weaker correlation for visual word processing N170 FRP source(s) in the posterior occipital areas, in the vicinity of the visual word form areas (VWFA). Furthermore, significant brain-to-brain correlations were found between the two modalities, where the speech brain sources of the P1 and N250 responses correlated with the reading N170 response. The results suggest that speech processes are linked to reading fluency and that brain activations to speech are linked to visual brain processes of reading. These results indicate that a relationship between language and reading systems is present even after several years of exposure to print.

Keywords: reading, ERPs, FRPs, auditory P1, auditory N250, visual N170, source reconstruction, brain correlates

Edited by:
Charles A. Perfetti,
University of Pittsburgh, United States

Reviewed by:
Yan Wu,
Northeast Normal University, China
McNeel Gordon Jantzen,
Western Washington University,
United States

****Correspondence:***
Najla Azaiez
najla.n.azaiez-zammit-chatti@jyu.fi
orcid.org/0000-0002-7525-3745

Specialty section:
This article was submitted to
Neurodevelopment,
a section of the journal
Frontiers in Neuroscience

Received: *17 April 2022*
Accepted: *20 June 2022*
Published: *19 July 2022*

Citation:
Azaiez N, Loberg O, Hämäläinen JA and Leppänen PHT (2022) Brain Source Correlates of Speech Perception and Reading Processes in Children With and Without Reading Difficulties.
Front. Neurosci. 16:921977.
doi: 10.3389/fnins.2022.921977

INTRODUCTION

Learning to read is a complex multi-step process that requires both visual and auditory processes (Kavale and Forness, 2000; Norton et al., 2015; Vernon, 2016; LaBerge and Samuels, 2017). The question of whether speech processing and visual processing deficits are linked to reading disorders has been extensively addressed in the literature (Breznitz and Meyler, 2003; Breznitz, 2006; Wright and Conlon, 2009; Georgiou et al., 2012; Kronschnabel et al., 2014; Francisco et al., 2017; Karipidis et al., 2017; Ye et al., 2017). However, the nature of the link between the two modalities remains unclear (Gibson et al., 2006; Wright and Conlon, 2009; Blau et al., 2010; Georgiou et al., 2012; Ye et al., 2017; Rüsseler et al., 2018; Stein, 2018). Several studies have investigated this relationship using simultaneous auditory and visual stimuli in dyslexic vs. typical readers using behavioral and brain measures (Aravena et al., 2018; Karipidis et al., 2018; Fraga-González et al., 2021). In the present study, we investigated the extent to which speech processing at the brain level is associated with reading fluency and brain activity during reading. We examined these associations in a group of children with different levels of reading skills, ranging from poor to good.

Reading difficulty (RD), or dyslexia, is a frequent neurodevelopmental impairment that is commonly reported among school-age children. It involves a failure to acquire a satisfactory level of reading and spelling performance, despite normal intelligence and typical linguistic performance, in the absence of any organic, psychiatric, or neurological disorders, and despite adequate pedagogical opportunities (Démonet et al., 2004; Peterson and Pennington, 2015; Snowling et al., 2020). Dyslexia has been commonly linked to deficits in speech processing (Schulte-Körne et al., 1998; Kujala et al., 2000; Bishop, 2007; Abrams et al., 2009; Hämäläinen et al., 2013; Christmann et al., 2015; Lizarazu et al., 2015; Gu and Bi, 2020) and phonological processing (Snowling, 1998; Richardson et al., 2004; Vellutino et al., 2004; Christmann et al., 2015; Smith-Spark et al., 2017; Goswami, 2019).

A frequently reported problem in dyslexia is word decoding, which is mainly described as a deficit in reading speed, accuracy, or spelling difficulties (Snowling, 2001; Vellutino et al., 2004; Siegel, 2006; Hulme and Snowling, 2014). According to phonological theory, RD is caused by alterations in brain functions, such as a deficit in phonological representations, an information storing dysfunction, or information inaccessibility (Ramus and Szenkovits, 2008; Hoeft et al., 2011; Boets et al., 2013; Hornickel and Kraus, 2013; Prestes and Feitosa, 2017). Based on this theory, one of the main hypotheses underlying the mechanism of reading disability is the creation of phoneme-grapheme neural connections or inadequate representations when processing speech signals. This deficit could result from an alteration of the process of decoding grapheme-phoneme correspondences when decoding single letters, letter clusters, or words while reading (Goswami, 2000; Prestes and Feitosa, 2017). Weakness in building a stable network by binding letters and sounds eventually leads to reading problems (Goswami, 2002; Noordenbos et al., 2012; Vogel et al., 2013). Several studies of brain responses in children with reading difficulties have reported deficits in speech and phonological processing (Snowling, 1998; Castles and Friedmann, 2014; Ramus, 2014; Catts et al., 2017), with atypical phonological or phonetic representations of speech sounds shown to alter normal phoneme and word identification. Alternatively, an impairment in letter-speech sound mapping has also been suggested to be the origin of reading problems (Ehri, 2005; Maurer et al., 2010; Žarić et al., 2014; Fraga-González et al., 2015). Several studies have shown that speech processing is tightly linked to reading processes and reading skills (Pennington and Bishop, 2009; Zhang and McBride-Chang, 2010; Price, 2012; Duncan, 2018). The early ERP response, P1/N1-P2/N2 complex, is known to reflect basic phonological processes such as sound detection and identification and complexity processing (Maurer et al., 2002; Alain and Tremblay, 2007; Durante et al., 2014; Hämäläinen et al., 2015). Another response, the N2/N250, which is also part of the early complex, has been described in the context of syllable processing and interpreted to reflect the building of neural representation with repeated auditory stimuli (Karhu et al., 1997; Ceponiene et al., 2005; Vidal et al., 2005; Hommet et al., 2009; Hämäläinen et al., 2018; Wass et al., 2019). Studies have shown that basic speech processing was a strong predictor of infants' and young children's reading skills development as early as 6 months of age (Leppänen et al., 2002; Meng et al., 2005; Boets et al., 2011; Hayiou-Thomas et al., 2017; Lohvansuu et al., 2018).

Using the electroencephalography (EEG) technique, letter-sound mapping was investigated in typical (CTR) and dyslexic readers, and the quality of letter-speech sound processing was shown to be related to reading fluency, with evidence of a relationship between the auditory and visual modalities (González et al., 2016; Moll et al., 2016; Karipidis et al., 2018). This grapheme-phoneme bind created during cross modalities network coactivation, has been described as a key step for developing fluent reading (Chyl et al., 2018; He et al., 2021) by enhancing the specialized visual areas related to print when presented with the corresponding letter-speech sound. This process typically occurs in the early learning stages of reading (Ehri, 2005; Fraga-González et al., 2021). As an example of this effect in EEG studies, it has been shown that ERP amplitudes (for the mismatch responses MMN and LDN, for example) were enhanced when speech sounds were presented to typical readers with letters—an effect that was absent in dyslexic readers (Froyen et al., 2009)—suggesting that in atypical reading development, this letter-speech bind is absent or very weak. Similar results were reported in adults. Blau et al. (2009) investigated whether phonological deficits impaired the mapping of speech sounds into equivalent letters. The authors showed reduced audiovisual integration among dyslexics compared to controls, which was linked to reduced activation of the superior temporal cortex, reflecting a deficit in auditory speech processing. The importance of the auditory cortex in the integration of letter-speech sounds has also been addressed in functional magnetic resonance imaging (fMRI) studies, both in adults (Van Atteveldt et al., 2004; Holloway et al., 2015; Yang et al., 2020) and in children. Yang et al. (2020) studied the neural basis of audiovisual integration deficits in dyslexic children *via* fMRI. Based on brain activation analysis, the authors reported a less developed correspondence of orthographic and phonological information matching in dyslexic children. They also reported reduced functional connectivity of important brain structures involved in integration processes, such as the left angular gyrus and the left lingual gyrus. This difference in the left superior temporal gyrus (STG) between the two groups of children was supported by previous findings in literature, whereas the angular gyrus (AG) activity was mainly related to task demand and attentional processes.

Visual processing deficits in reading have also been reported for dyslexia and reading problems (Eden et al., 1996; Lobier et al., 2012, 2014; Giofrè et al., 2019; Archer et al., 2020). Visual deficits related to reading have previously been reported at different levels, such as in the sensory, temporal, attentional, and memory processes (Farmer and Klein, 1995; Snowling, 2001; Facoetti et al., 2006; Boets et al., 2008; Wright and Conlon, 2009; Conlon et al., 2011; Goswami, 2015). For example, low-level visual processing in letter-speech sound integration was addressed using a mismatched paradigm to investigate the influence of speech sounds on letter processing. Despite previous evidence of the systematic modulation effect of speech sound processing on letter processing, the reverse effect was not found (Froyen et al., 2010). The emergence of letter-speech sound correspondence has been studied in children *via* different neuroimaging techniques. Brem et al. (2010) studied the establishment of a reading network *via* speech processing in beginning readers *via* ERP and fMRI. That study focused on the left occipitotemporal cortex underlying the VWFA. The authors showed that print sensitivity in this area emerged in the early phases of reading acquisition, highlighting the critical role of VWFA in sound-print mapping. The results of Brem et al.'s investigation of fMRI and EEG data clearly indicated brain activity enhancement in the occipitotemporal area after the establishment of speech-print mapping through training. The authors reported that the auditory network involved was not the only one, but that a visual network was clearly co-activated during the coding-decoding phases, which highlighted the importance of the VWFA in this learning process. Brem et al. also associated the activation of this brain area with the visual N1 response of the ERP component sensitive to print, more commonly named N170. Pleisch et al. (2019) studied differences in reading processes between typical and dyslexic first-grade children by measuring the neural activation of the N1 response to print *via* combined EEG–fMRI methods. A differential modulation reflecting sensitivity to print was found only in typical readers in the ventral occipitotemporal cortex. The authors concluded that functional brain alterations in the language network play a role in dysfluent reading development.

The role of speech and language as the basis for reading is well established, where most dyslexics show difficulties in phonological processing (Siegel, 2006; Navas et al., 2014; Giofrè et al., 2019). Sensory or orthographic visual processing deficits have only been observed in a subgroup of the dyslexics (Wright and Conlon, 2009; Giofrè et al., 2019). Visual processing in RD remains an important processing aspect to study in reading research, which has already been a focus of investigation in the literature (Salmelin et al., 1996; Lobier et al., 2014; Archer et al., 2020). However, the ties between visual and auditory information processes in the context of reading vs. speech processing remain unclear. The processing of several letters in a short timeframe is an important skill for developing fluent reading. It has been shown that RD is characterized by slow word recognition and a higher error rate compared to typical reading (Ozeri-Rotstain et al., 2020). Efficient word processing depends on parallel visual processing of multiple letters (Lobier et al., 2012), where a problem in letter pattern perception leads to a problem in orthographic processing, inducing reading problems (Georgiou et al., 2012).

Monzalvo et al. (2012) used fMRI to investigate cortical networks for vision and language by comparing cortical activity in minimally demanding visual tasks and speech-processing tasks. In the visual paradigm, objects, faces, words, and a checkboard were used as stimuli presented in different blocks, and short sentences in native and foreign (unfamiliar) languages were used in the speech processing paradigm. Both visual and spoken language systems have been reported to be impaired in dyslexics. Monzalvo et al. found that dyslexics had reduced activation of words in the VWFA in the visual task and reduced responses in different brain areas, including the posterior temporal cortex, left insula, planum temporal, and left basal language area, extending to the VWFA, in the speech tasks, and the authors concluded that there was hypoactivation in the VWFA for written words and speech listening. These results highlight the role of the VWFA as an associative area in the processing of both types of stimuli: visual information in reading and auditory information in speech processing. A more recent fMRI study by Malins et al. (2018) used a task of matching printed and spoken words to pictures and found a significant correlation between the neural activity of both print and speech and reading skills in children. The authors studied trial-by-trial neural activation of different brain areas and their relationship to reading. They showed that the variability of the neural activation to print was positively correlated with the activation variability of the inferior frontal gyrus providing an additional evidence on the relationship between reading skills and sound processing. The authors discussed the common neural activations for print and speech and highlighted individual differences.

When studying visual processing, the eye-tracking technique is frequently used to examine visual processes and eye movements to investigate reading and reading disorders (Jainta and Kapoula, 2011; Tiffin-Richards and Schroeder, 2015; Kim and Lombardino, 2016; Nilsson Benfatto et al., 2016; Jarodzka and Brand-Gruwel, 2017; Breadmore and Carroll, 2018; Robertson and Gallant, 2019; Christoforou et al., 2021). FRPs are a specific type of ERP that rely on eye fixations and their corresponding brain activity (Baccino, 2011). This combined technique is commonly used to investigate reading (Baccino, 2011; Wenzel et al., 2016; Loberg et al., 2019; Degno and Liversedge, 2020). The FRP is based on EEG measurements of brain activity in response to visual fixations obtained by extracting the signal-averaged time-locked to the onset of eye fixations (Baccino, 2011). Fixations in reading are known to reflect the online cognitive process of several factors, such as the duration and location of a word, word frequency, and predictability. This process occurs in a series of events, starting with the transmission of the visual signal of the printed word from the retina to the visual cortex, visual encoding, initiation of word identification, and programming the next eye movement (Degno and Liversedge, 2020). A commonly used measure for studying individual differences in reading is first-pass fixation duration. This measure reflects the cognitive components of early visual processing, word identification, attention shifts, and oculomotor control (Zhang et al., 2021a). Jainta and Kapoula's (2011) study of eye fixations in reading showed a large fixation disparity that caused unstable fixations in dyslexic children when reading sentences. The authors concluded that visual perturbation may cause letter/word recognition and processing difficulties that lead to reading disorders. Zhang et al. (2021a) used first-pass fixation in sentence reading to investigate the brain network in natural reading. They showed that seed regions in the early visual cortex, VWFA, and eye-movement control network were associated with individual reading performance and brain connectivity in a resting state.

Interestingly, this visual dysfunction was not found systematically, since some studies did not report any differences between RD and typical readers and not all children with RD show a visual deficit.

In the context of RD, both speech and visual processes have only rarely been investigated *via* the ERP method. For example, Bonte and Blomert (2004a) investigated dyslexic readers' phonological processing in spoken word recognition using a priming paradigm. The authors examined the general ERP response and reading skills of beginning readers and reported deficits in N1 and N2 speech processes in dyslexics compared to controls. They interpreted these results as a phonological processing deficit reflecting the recruitment of different neural sources (Bonte and Blomert, 2004a). The N250 response, which is known to be part of the obligatory response (P1-N250), was also investigated in dyslexia, and previous studies showed a different response in this component in the RD group compared to the control group (Lachmann et al., 2005; Lohvansuu et al., 2014). The N250 is thought to represent low-level auditory processing, such as sound detection or feature extraction, but it is also part of a critical processing stage, which is the formation of the neuronal representation of sound/speech stimuli (Karhu et al., 1997; Hämäläinen et al., 2015). As reading involves the ability to convert print into sound, it is critical to further investigate the N250 response and its relationship to reading, as previous evidence has shown differences in this component between good readers and dyslexics. However, the relationship between N250 and reading remains unclear. In addition to the N1-N2 findings, later ERP responses were also found deficient among RD participants, such as the P3a, the N400, and the LDN (Hämäläinen et al., 2008, 2013; Jednoróg et al., 2010; Desroches et al., 2013; Leppänen et al., 2019). These findings provide evidence that speech processing may be altered in dyslexics at different stages of processing and at different latencies.

The brain potential of interest in reading is the N170, an ERP component that peaks between 150 and 200 ms, with a peak around 170 ms and a temporo-occipital negative topography (Rossion et al., 2002; Maurer et al., 2005b; Sánchez-Vincitore et al., 2018). The N170 has been identified as reflecting facial recognition and being sensitive to facial expressions (Blau et al., 2009; Hinojosa et al., 2015; Wang et al., 2019). This component is known to be sensitive to orthographic processing (Rossion et al., 2003) and to letters strings/words in reading. When left lateralized, the N170 has been shown to be a reliable physiological marker of reading and reading skills (Maurer et al., 2005b, 2008; Lin et al., 2011; Hasko et al., 2013; Zhao et al., 2014; Lochy et al., 2016). For example, it was studied in dyslexic children compared to controls, where the N170 was shown to have a larger response in the dyslexic group compared to controls (Fraga González et al., 2014; González et al., 2016). Time-locked to the visual response, this ERP response becomes a strong indicator for studying the dynamics of the visual cognitive processes (labeled FRP N170) of reading and reading disorders (Dimigen et al., 2011, 2012; Kornrumpf et al., 2016; Loberg et al., 2019; Dimigen and Ehinger, 2021).

In the present study, we investigated how the basic speech ERP responses—the P1-N250—are related to reading process, and how the visual FRP response in reading—the N170, which is known as a reliable marker of reading processes (Maurer et al., 2005b; Hasko et al., 2013)—are associated with reading skills in the same children. Previous evidence has shown a link between speech perception and reading, with speech processing being a predictor of reading development at an early age, but the temporal-brain dynamics remain unclear. Moreover, the question of whether this relationship remains present after the development of reading skills has scarcely been investigated. Here, we aim to investigate whether the basic processes of speech remain associated with basic processing of reading in school-aged children who have established a reading network, and how their reading skills may reflect their neuronal activity. This study represents a new approach to investigate how visual reading and auditory speech processes may be interlinked and linked to reading skills by combining different methods (ERP, FRP, and CLARA) for high temporo-spatial analysis.

Both auditory and visual modalities were tested in two separate tasks: a speech perception task and a sentence-reading task. We used source reconstruction with correlation analyses to identify the link(s) among reading skills and auditory processes, reading skills and visual processes, and the neuronal activity of the two modalities. This enabled us to study the brain dynamics of these processes by examining the neuronal origin of brain activity at the source level and to explore its relationship to reading skills. Based on previous evidence, we hypothesized that speech perception basic responses (P1-N250) would show correlations with reading skills (Bonte and Blomert, 2004a; Lohvansuu et al., 2018) and that the visual N170 response would also correlate with reading skills (Maurer et al., 2008; Mahé et al., 2013; Fraga González et al., 2014). Furthermore, we expect to observe a relationship between the speech processes P1 and N250, and the visual reading processes over the VWFA within the same subjects in these two independent tasks.

MATERIALS AND METHODS

Participants

A total of 440 children from eight schools in the area of Jyväskylä, Finland, participated in three test cohorts. The study included a subsample of 112 children, all Finnish native speakers aged between 11 and 13. These children were invited to participate in the eSeek project (Internet and Learning Difficulties: A Multidisciplinary Approach for Understanding Reading in New Media). The participants were grouped based on their reading fluency scores derived from three different reading tasks. The latent score was computed for reading fluency using principal factor analysis (PAF) with PROMAX rotation in the IBM SPSS 24 statistical program (IBM Inc.). This score was based on the following three tests: The Word Identification Test, a subtest of standardized Finnish reading test ALLU (Lindeman, 1998) (factor loading 0.683); the Word Chain Test (Nevala and Lyytinen, 2000) (factor loading 0.683); and the Oral Pseudoword Text reading (Eklund et al., 2015) (factor loading 0.653).

The word identification test included 80 items, each consisting of a picture and four alternative written words. The task was to identify and connect correct picture–word pairs. The score was the number of correctly connected pairs within the 2 min. The word chain test consisted of 25 chains of four words written without spaces between them. The task was to draw a line at the word boundaries. The score was the number of correctly separated words within the 90 s time limit. The oral pseudoword text-reading test consisted of 38 pseudowords (277 letters). These pseudowords were presented in the form of a short passage, which children were instructed to read aloud as quickly and accurately as possible. The score was the number of correctly read pseudowords divided by the time, in seconds, spent on reading (for details, Kanniainen et al., 2019).

This reading score was computed for the whole sample for each subject. Children who scored below the 10th percentile were identified as poor readers (RD) and those who scored above the 10th percentile were identified as good readers (CTR).

All participants scoring equal to or below 15 points (10th percentile) in the cognitive non-verbal assessment testing were excluded. This assessment included a 30-item version of Raven's progressive matrices test (Raven and Court, 1998). Attentional problems were screened *via* the ATTention and EXecutive function rating teacher inventory (ATTEX in English and KESKY in Finnish) (Klenberg et al., 2010). To be included in the analyses, the participants had to score below 30 points on this test. Children with clear attentional problems were excluded from the study.

The brain response analyses were conducted on 112 participants: auditory data: 86 CTR participants (43 females and 43 males; age range = 11.78–12.84 years; mean age 12.36 years, SD: 0.27) and 26 RD participants (eight females and 18 males; age range = 11.84–12.94; mean age 12.31 years, SD: 0.34). Preprocessing and source modeling were performed on 92 participants' reading data: 65 CTR participants and 27 RD participants.

The correlation analysis only included participants with valid auditory and visual data. Sixty of these participants comprised the final CTR group (30 females and 30 males; age range = 11.88–12.84 years; mean age 12.37 years, SD: 0.28) and 20 participants were in the RD group (six females and 14 males; age range = 11.84–12.94 years; mean age 12.34 years, SD: 0.36). The final group, which included both samples from CTR and RD (labeled CTRD), comprised 80 subjects and was tested for normality and skewness. The tests showed a normal distribution and no skewness. For details, see the **Supplementary Material**.

None of the participants declared any auditory problems, and they all had normal or corrected vision with no history of neurological problems or head injuries. The current study was conducted in compliance with the Declaration of Helsinki, and the study protocols were approved by the Ethics Committee of the University of Jyväskylä, Finland. All of the methods used were performed in accordance with university guidelines and regulations. The participants and their parents provided signed informed consent prior to the study.

Materials and Procedures

Auditory Materials and Stimulus Presentation

The auditory stimulus used for this study was originally presented in a passive oddball paradigm designed for another study, comprising a standard stimulus and two deviant stimuli presented over a duration of 10 min. The target stimulus (standard) was presented 800 times in the paradigm, but only 200 trials, which were the pre-deviant standard stimulus responses, were used for the analysis. These trials are believed to have the strongest representations of stimuli due to repetition. The stimulus consisted of a Finnish monosyllabic word *suu* (which means "mouth" in English), a basic, frequent, short, and easy word that is commonly used by itself in the Finnish language but could also be part of other words like [o**suu**s ("a portion or contribution") or a**suu** ("lives")]. This is also one of the first words learned by Finnish children at a very early age and is therefore expected to have a strong neural representation among Finnish participants. The stimuli were recorded by a male native speaker and were pronounced in a neutral manner. The recording was equalized and normalized in segmental durations, pitch contours, and amplitude envelopes using Praat software (Boersma and Weenink, 2010) for a more detailed description of stimulus preparation (Ylinen et al., 2019). The stimuli were presented *via* a loudspeaker placed on the ceiling ∼100 cm above the participants' ear position and were presented at ∼65 dB. The stimulus volume level was calibrated before each recording with a sound level meter (Brüel and Kjaer) placed on a pedestal device at the participant's head position (with the following settings: sound incidence = frontal; time weighting = fast; ext filter = out; frequency weighting = A, range = 40–110 dB; display = max).

Reading Materials

Two hundred sentences, each with between five and nine words, and a median length of six words, were used as visual stimuli. The sentences were presented in 20-point Times New Roman font on the screen in a free-reading task. Each letter was subtended at an average visual angle of 0.4 degrees on the screen, where the distance of the participants was ∼60 cm from the monitor. A total of 912 words, with lengths varying from 5 to 13 letters, were included in the FRP analysis. The materials for this paradigm were part of a previous study. For a detailed description, see Loberg et al. (2019).

Data Measurements

EEG recordings were performed in a sound-attenuated and electrically shielded EEG laboratory room located at the University of Jyväskylä facilities. There was no task for the auditory paradigm. Each child was instructed to minimize movement while listening passively to auditory stimuli. To maintain the child's interest in the experiment, he/she watched a muted cartoon movie playing on a computer screen. In the reading paradigm, the measurement was performed in the same room using a dim light. The child was instructed to freely read different sentences that appeared on the screen. During the recordings, the experimenters observed the participant *via* live video camera streaming and monitoring from a separate control room to ensure the wellbeing of the participant and that the experiment proceeded as expected.

Both EEG datasets were recorded with 128 Ag-AgCl electrode nets (Electrical Geodesic, Inc.) with Cz as the online reference, using NeurOne software and a NeurOne amplifier (MegaElectronics Ltd., new designation Bittium). The data were sampled online at 1,000 Hz, high-pass filtered at 0.16 Hz, and low-pass filtered at 250 Hz during the recording. The experimenter aimed to keep impedances below 50 kΩ and the data quality was checked continuously. All necessary adjustments or corrections were performed during short breaks and between the experiments' blocks to maintain good quality throughout the measurements.

The Eyelink 1,000 with 2,000 Hz upgrade (SR research) version was used for the eye-movement data acquisition of the reading task using a 1,000 Hz sampling rate. The sentences were presented on a Dell Precision T5500 workstation with an Asus VG-236 monitor (1,920 × 1,080, 120 Hz, 52 × 29 cm). At the beginning and the end of each trial, the synchrony between the two measures was ensured with a mixture of transistor-to-transistor logic pulses (to EEG) and Ethernet messages [to eye tracking (ET)]. The participants held their heads in a chinrest during the measurements. The calibration routine consisted of a 13-point run of fixation dots performed before each block and before each trial. This reading task was divided into four blocks. If the fixation diverged from the calibration by more than one degree, the calibration was redone. The experiment's trial started only upon the experimenter's approval of the calibration. Once the task started, the participants were instructed to press a button to move to the next trial (for details, see Loberg et al., 2019). The participants were instructed to read as quickly as possible. The quality of the EEG and the ET was maintained throughout the experiment, and corrections and recalibrations were performed as required. Short breaks were taken when needed or upon the participant's request.

In both experiments, the participants were informed that they were allowed to terminate the experiment at any time in the case of discomfort.

Auditory Data Preprocessing

BESA Research 6.0 and 6.1 were used for offline data processing. Bad channels were identified from the data (number of bad channels: mean: 5.6, range: 1–13). Independent component analysis (Infomax applied to a 60-s segment of the EEG) (Bell and Sejnowski, 1995) was used to correct the blinks from each subject's data. Epoch length was set from −100 ms (pre-stimulus baseline) to 850 ms. The artifact detection criterion was set to a maximum of 175 μV for amplitude fluctuations within the total duration of the epoch. A high-pass filter of 0.5 Hz was set before averaging. Bad channels showing noisy data were interpolated using the spherical spline interpolation method (Ferree, 2006). The data were re-referenced offline to average the reference and averaged individually and separately for the standard stimulus.

Reading Data Preprocessing

The co-registered EEG-ET data were processed in MATLAB using EEGLAB (v14.1.2) with an EYE-EEG (0.85) add-on. A high-pass filter at 0.5 Hz and a low-pass filter at 30 Hz were applied. Synchronization between the raw gaze position data and the EEG data was performed using shared messages in both data streams at the beginning and the end of each trial. Gaze positions outside the screen were automatically discarded. Discarded trials included all zero gaze positions resulting from blinks and between trial gaps in the recordings. All fixations corresponding to all the words within the sentences, except for the last word, during a first-pass reading were used to compute the FRP estimate. The responses were locked to the first fixation of each word, mean word length of 8, and saccade amplitude of 1,8798'. A time window of 100 ms was also considered bad data before and after these values. A binocular median velocity algorithm for detecting fixations (and saccades) was applied to the remaining gaze positions.

Deconvolution Modeling of FRPs

The UNFOLD toolbox (Ehinger and Dimigen, 2019) was used for the FRPs estimation. The FRPs were estimated *via* a generalized linear model that was used for response estimation and the correction of overlaps between the responses with a generalized additive model for non-linear predictors (Loberg et al., 2019). The modeled response ranged from −700 to 500 ms from fixation onset. All blink time points, eye movements outside the screen, and segments with large fluctuations were removed from the response estimates. Fixations on the target word during re-readings were excluded from the FRP estimation.

Source Reconstruction and Spatial Filtering

Source analyses were conducted using BESA Research 6.1 and 7.0 to estimate the active sources in the speech processing and reading tasks. The neuronal sources were estimated *via* an inverse approach with a distributed source model in the brain volume: classical LORETA analysis recursively applied (CLARA) restricted to the cortex. For accurate forward head modeling, an appropriate FEM head model for 12-year-olds was implemented. Model solutions were created based on the group ERP brain source reconstructions for each brain component for the CTRD group combined in a unique model. For the auditory data, source locations were calculated for P1, P1-$_2$, N250, and N250-$_2$ (see an illustration of the ERP auditory responses in **Figure 1**). Model solutions were similarly computed for the reading data based on the group FRP estimates, where the target component was N170. The source analysis was performed ∼10 ms before the peak for all components. This time point was chosen after inspection and after searching for the best solutions for the different responses. This time showed the best modeling solution for the source activity, with the clearest sources and the best residual variance. These group-based solutions were used to create a standard model to filter cortical sources, and only sources that were found to be activated in the common group (CTRD) were included in the final model. For each CLARA source identified, a regional dipole was fixed to combine the power sum of the three

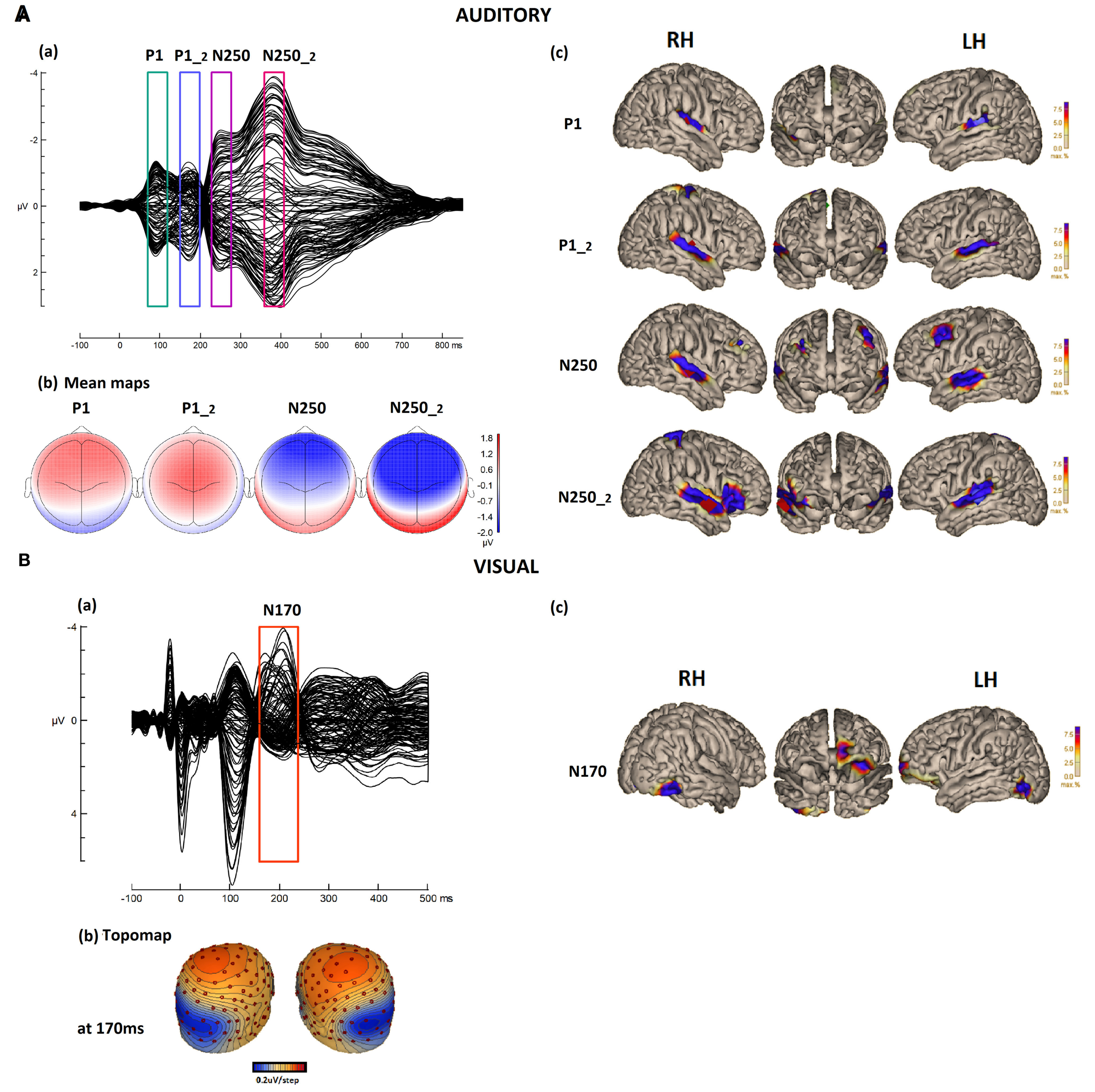


FIGURE 1 | (A) (a) Auditory/speech ERPs in the CTRD group (*N* = 80) grand average. Butterfly plots for the responses to the standard stimulus "suu" over 129 electrodes. The boxes around the peaks indicate P1, P1-2, N250, and N250-2 responses. (b) The corresponding mean topographic maps for the time windows of 70–120 ms (P1), 150–200 ms (P1-2), 230–280 ms (N250), and 360–410 ms (N250-2), respectively. (c) Cortical CLARA reconstruction for each component. **(B)** (a) Visual/reading FRPs in the CTRD group (*N* = 80) grand average. Butterfly plots for the responses to word stimuli over 129 electrodes. (b) The topographic map of N170 at 170 ms and (c) its cortical source CLARA reconstruction.

orthogonal orientations of the regional sources. The regional sources were computed for each component. They were then used as spatial source filters and applied to individual data. The source filter generated individual solution waveforms for each participant. A mean scalar value for each subject was computed as the sum of the source activity measures at all time points over a time window between ∼20 and 30 ms around the peak, specified for each component (a detailed description of the time windows is provided below).

Correlations

Correlations between source activations were converted into scalar values for each modality, and the reading scores (PAF) were examined across the CTRD group using Pearson's correlation coefficients. For each source activity, the mean value was calculated around the peak using MATLAB R2019b (Mathworks®), as described above. For the auditory data, the time windows for the averages were 80–110 ms for P1, 150–180 ms for P1-₂, 230–250 ms for N250, and 360–390 ms for N250-₂. For the visual data, the time window 180–210 ms was used for N170. These time windows were chosen based on visual inspection of the group ERP and FRP grand averages. The time windows were fixed so that the peak was always located in the middle of the window.

Pearson's correlation coefficients were calculated between the average source activity and the reading score of the participants using IBM SPSS statistics 26 (IBM corp), version 26.0.0.1, and applying a false discovery rates (FDR) correction of $q = 0.05$ (Benjamini and Hochberg, 1995) for the brain-to-behavior correlations and the brain-to-brain correlations. Correlations within brain activity between auditory and visual source activities were computed. A partial correlation (controlling for reading skills/PAF) between the source activity in the reading and speech processes was also performed.

RESULTS

Brain Responses and Source Reconstructions

Brain Responses to Auditory ERP and Visual FRP Data

The auditory grand average ERP and the different auditory components are illustrated in **Figure 1A**. The ERP waveform (**Figure 1A**a) shows four components that emerged in response to the auditory stimulus. The first component peaked at around 90 ms, with a clear fronto-central positive polarity, and reflected the P1 response to the stimulus onset. This was followed by a second positive component peaking at around 170 ms, reflecting a second P1 response (P1-₂) in response to the onset of the vowel or to the consonant-vowel transition. This response had a somewhat more central topography. The third component peaked at around 250 ms and reflected the N250 response to the stimulus onset, followed by a fourth component peaking at around 370 ms, most likely reflecting a second 250 (N250-₂) response to the consonant-vowel transition or the onset of the vowel in the stimulus. Both responses showed clear negativity in the fronto-central area, with a larger amplitude for the second N250 response (**Figure 1A**b).

The grand average of the FRPs during reading is illustrated in **Figure 1B**. The component peaking around 200 ms reflects the visual N170 response, with topography (**Figure 1B**b) showing a typical N170 response. The polarity was positive over the central area and negativity in the occipital areas, with a preponderance toward the left occipital hemisphere.

Cortical Sources in Speech Processing

The group-based cortical source reconstruction (applying CLARA) of the auditory responses is illustrated in **Figure 1A**c. For auditory P1, the source reconstruction at 80 ms, shows a bilateral focal activation of the primary auditory cortices (A1) [with a total residual variance (RV) of 1.78%]. The source reconstruction of the second component P1-₂ performed at 160 ms shows the activation of similar bilateral sources over the auditory cortices. This second response shows slightly larger activity covering a larger area than the first P1, with an additional small activation over the central region (total RV = 5.12%). The third source reconstruction performed at 230 ms for the first N250 response revealed four sources. Two sources were active bilaterally in the left and right temporal lobes at the level of the superior temporal area (STA). In addition, the inferior frontal area (IFA) in the left hemisphere and the middle frontal area in the right hemisphere were activated (total RV = 2.83%). The fourth reconstruction was performed for the N250-₂ response at 370 ms. The source reconstruction showed four sources: bilateral activation of the left and right STA, the third source in the right IFA, and the fourth in the center-right area of the cortex (total RV = 2.19%). Only the bilateral auditory sources across the different components were used to run the correlation analysis to investigate the relationship between the auditory speech perception processes and the reading processes at both the behavioral and neuronal levels. The other sources were discarded because they are believed to reflect additional processes that are related to attentional or semantic processes.

Cortical Sources in Reading Processing

The group-based cortical source reconstruction of the visual response is illustrated in **Figure 1B**c. For reading N170, the reconstruction was performed at 190 ms and showed five main sources (with an RV of 6.07%). Two sources were located in the left and right occipital areas: one over the middle temporal area and one over the right visual cortex. Two additional activations were also found over the left frontal area: one source located in the left orbitofrontal area and the second in the left prefrontal area. Only the visual reading sources of the occipital areas were kept for the correlation analysis to investigate the reading processes, as the frontal sources are believed to reflect other processes that are mainly related to attentional processes.

Correlations

Cortical Source Correlations With Reading Scores

Table 1 presents the correlations between the scalar values of the cortical source activity in the speech paradigm and reading scores, and in the cortical source activity in the reading paradigm and reading scores.

A significant negative correlation was found between the P1 source activity of the left auditory cortex (A1) and the reading score (PAF). The correlation analysis with the right source activity did not reveal any significant results. Neither the right nor the left brain activity of the P1-₂ or N250 sources correlated with PAF. At the time window of the N250-₂ response, source activities in both the left and right temporal areas (STA) correlated negatively with PAF. The correlations indicated that the larger the response, the poorer the reading score. The correlations between the scalar values of the visual sources and the PAF are illustrated in **Table 1**. Only the left occipital source activity located over the left occipital area (L VWFA) correlated negatively with the PAF score. However, this correlation became non-significant after multiple comparison corrections.

TABLE 1 | Brain-to-behavior correlation analysis between reading fluency and brain source activity in auditory and visual sources.

	Components									
	Auditory P1		**Auditory P1_2**		**Auditory N250**		**Auditory N250_2**		**Visual N170**	
Sources	**R AC**	**L AC**	**R STA**	**L STA**	**R STA**	**L STA**	**R STA**	**L STA**	**R VWFA**	**L VWFA**
Correlation	−0.141	**−0.337**	−0.034	−0.192	−0.204	−0.096	**−0.304**	**−0.273**	−0.210	**−0.224**
Significance	0.212	**0.002**[a]	0.762	0.880	0.690	0.396	**0.006**[a]	**0.014**[a]	0.062	**0.046**

AC, auditory cortex; STA, superior temporal area; VWFA, visual word form area; R, right hemisphere; L, left hemisphere.
The correlations significant before the FDR correlation are shown in bold.
[a] *indicates that the correlations remained significant after the FDR multiple comparison corrections.*

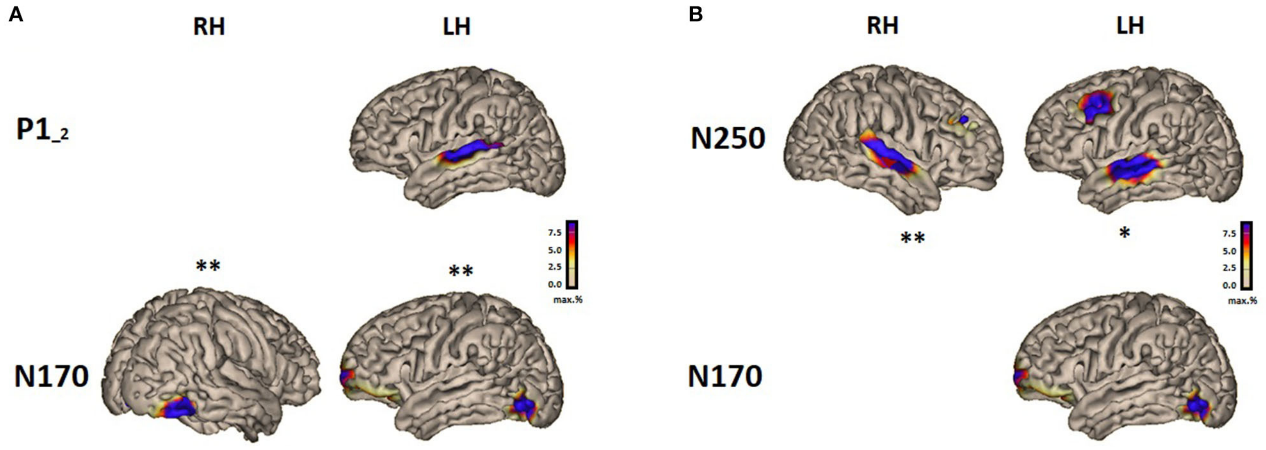


FIGURE 2 | Summary results showing significant correlations between the source activity in speech processing and reading in the brain-to-brain analysis for the P1-$_2$ vs. N170 **(A)** and N250 vs. N170 **(B)**.

TABLE 2 | Brain-to-brain correlation analysis between auditory and visual source activity.

			Auditory components							
			Auditory P1		**Auditory P1_2**		**Auditory N250**		**Auditory N250_2**	
	Sources		**R AC**	**L AC**	**R STA**	**L STA**	**R STA**	**L STA**	**R STA**	**L STA**
Visual N170	L VWFA	Correlation	0.146	0.196	0.121	**0.335**	**0.294**	**0.286**	**0.222**	**0.231**
		Significance	0.197	0.081	0.284	**0.002**[a]	**0.008**[a]	**0.010**[a]	**0.047**	**0.039**
	R VWFA	Correlation	−0.004	0.118	0.180	**0.316**	0.209	0.154	0.122	**0.225**
		Significance	0.972	0.299	0.109	**0.004**[a]	0.063	0.172	0.279	**0.045**

AC, auditory cortex; STA, superior temporal area; VWFA, visual word form area; R, right hemisphere; L, left hemisphere.
The correlations significant before the FDR correlation are shown in bold.
[a]*indicates that the correlations remained significant after the FDR multiple comparison corrections.*

Correlations Between Visual and Auditory Sources

Figure 2 shows the correlations between the scalar value of the visual N170 source and the auditory source activities.

The activity of the auditory P1-$_2$ source (for consonant-vowel transition/vowel onset in "suu") located in the left hemisphere over the temporal area (L STA) correlated significantly with both active sources of the N170 over the left and right hemispheres (L VWFA and R VWFA). The higher the auditory source activity, the higher the activity of the visual sources. The activity of the auditory N250 sources (for the stimulus "suu" onset) located in the left and right hemispheres (L STA and R STA) correlated significantly only with the left source activity of the N170 response (L VWFA). The larger the response to the auditory stimulus, the larger the response to the visual stimulus; see **Table 2** for details.

Partial correlations controlling for reading scores were conducted to investigate whether the brain-to-brain correlation was mainly driven by reading skill level. As shown in

TABLE 3 | Partial correlation (controlling for reading scores) between the auditory and visual source activities conducted for the brain-to-brain correlations (after FDR correction).

	Components correlated							
	Auditory P1_2 * Visual N170				**Auditory N250 * Visual N170**			
Sources correlated	**Auditory L STA**	**Visual L VWFA**	**Auditory L STA**	**Visual R VWFA**	**Auditory R STA**	**Visual L VWFA**	**Auditory L STA**	**Visual L VWFA**
Correlation	0.306		0.288		0.260		0.273	
Significance	0.006		0.010		0.021		0.015	
df	77		77		77		77	

STA, superior temporal area; VWFA, visual word form area; R, right hemisphere; L, left hemisphere; df, degree of freedom.

Table 3, controlling for the reading score did not change the correlations noticeably.

DISCUSSION

This study had two main aims. The first was to investigate the relationship between speech processes and reading fluency, indicated by the PAF score, and visual brain activity in reading, as reflected by the VWFA activation, with the reading score. The second aim was to investigate the brain-to-brain responses for speech and reading processes among a group of children with different reading skills, ranging from good to poor. The study was conducted using brain ERPs for speech stimuli, FRPs for words in sentence stimuli, and source reconstruction for both processes to conduct the correlation analysis. To reveal the link between brain activity and reading skills, we first investigated the correlation between the brain activity of each modality (auditory and visual separately) and reading skills, as indicated by PAF, a reading fluency score derived from three different reading tasks. Our results showed that brain activity correlated with reading scores over the P1 and N250-$_2$ components. The brain activity in reading, as reflected in N170 over the left hemisphere occipital area (L VWFA), correlated significantly with the reading fluency score. However, this correlation did not survive the statistical correction. The brain-to-brain analysis revealed the presence of significant correlations between speech-generated brain responses and reading source activity. The strength of the speech processing sources in the P1-$_2$ and the early N250 showed a correlation with the VWFA source strength for N170. The current results are in line with the trends found in the literature, where the early speech components, P1 and N250, showed correlations with reading. However, our results showed that specific components correlate with behavioral reading skills, whereas other components correlate with brain reading processes. Our findings provide new evidence that there is still reliance on the auditory system and basic speech processes, even after long exposure to print, suggesting that the visual reading system continues to be linked to the auditory system at this developmental age.

In the first part of the study, we investigated the different brain components emerging in speech processing and reading tasks and their cortical sources. In the speech processing task, we examined brain responses to the standard "suu". We chose this stimulus because it was the most repeated speech sound in the oddball paradigm. The literature has shown that stimulus repetition forms a strong memory trace (Jaramillo et al., 2000; Näätänen and Rinne, 2002; Haenschel et al., 2005) and generates a strong neural phonemic/phonetic representation. This phonetic representation was suggested to be linked to the print N170 response (Hsu et al., 2009; Zhang et al., 2021b).

The speech processing ERP results showed two main responses, P1 and N250, both of which have a two-peaked structure reflecting the nature of the syllable-word stimulus "suu". Two similar positive peaks appeared in the early part of the response, one at 80 ms and the second at 170 ms, both of which showed similar scalp topographies with a fronto-central distributed positivity. The first peak seems to be a classic P1 peak emerging in response to the first sound of the syllable /s/, labeled here as P1. The second peak seems to emerge as a response to the second sound of the stimulus, /uu/, labeled as P1-$_2$. This double-peak structure was also found for the second part of the response in the time range of the N250 component. Two similar peaks with similar fronto-central negative topographies appeared at 250 and 370 ms. The first N250 response is likely to reflect the further processing phase of the first sound /s/ (of /suu/), labeled as the early N250, and the second response to reflect the second processing phase of the second sound /uu/ and labeled as N250-$_2$. N250 and N250-$_2$ differed in amplitude, where the second component showed a very high negative amplitude compared to the first. This may be interpreted by a cumulative effect, where the N250-$_2$ compromised the coarticulation processing in addition to the stimulus second sound /uu/ processing. This higher amplitude could also reflect the repetition effect, as both N250 and N250-$_2$ showed higher amplitudes compared to the P1 responses. Another possible interpretation is that this enhanced response is due to the nature of the word stimulus, its strong familiarity, and its well-established neural representation. Early lexical/semantic access in this early phase is also possible. Early semantic access at this time range has been proposed in the literature (Zhao et al., 2016).

Previous studies have identified the early complex P1/N1-P2/N2 as the auditory change complex, reflecting the consonant-vowel transition in naturally produced syllables by children (Boothroyd, 2004). The P1-N250 complex response has been described in the literature as part of the basic auditory processing response (Ceponiene et al., 2005; Gansonre et al., 2018). The

P1 is known to be an obligatory response reflecting sound detection and phoneme identification (Durante et al., 2014; Hämäläinen et al., 2015; Kuuluvainen et al., 2016), whereas the N250 was suggested to reflect phonological processing (Eddy et al., 2016), but also seemed to play a role in memory trace formation (Karhu et al., 1997; Ceponiene et al., 2005; Khan et al., 2011; Hämäläinen et al., 2013). These auditory speech responses have previously been shown to be linked to reading skills and have been studied in the context of typical reading and reading problems (Parviainen et al., 2011; Hämäläinen et al., 2015; Kuuluvainen et al., 2016). Differences between typical and dyslexic readers in these obligatory brain responses were found to emerge between 100 and 250 ms (Bonte and Blomert, 2004b; Hämäläinen et al., 2007, 2015; Khan et al., 2011).

In the reading task, the FRP results showed a typical N170 response. The N170 component has previously been described as reflecting objects and face recognition processes (Rossion et al., 2002; Collin et al., 2012; Hinojosa et al., 2015). It is also known to reflect print and word reading processes. This response was investigated in typical reading and RD and has been shown to have left-lateralized brain activity in reading (Maurer et al., 2005a, 2008; Mahé et al., 2013; Sacchi and Laszlo, 2016; Loberg et al., 2019).

In source reconstructions, the P1 component showed bilateral activation over the primary auditory cortices. In P1-2, the source reconstruction also shows bilateral brain activity in the auditory areas extending to the lateral surface of the STAs in this later response. The sources seem to be similar in both P1 responses, as both reflect similar processes occurring at different time points, where each component reflects the processing of a specific sound of the stimulus. Similar brain areas have been identified for P1 sources when processing auditory stimuli in adults and children (Godey et al., 2001; Shahin et al., 2004; Ruhnau et al., 2011). Our source reconstruction of the N250 component showed more inferior bilateral sources over the auditory areas (superior temporal and middle temporal areas), but an activation of frontal sources was also observed. In the N250-2, bilateral activation was also found in the auditory areas, with slightly more anterior location and with activation of frontal areas. Similar brain areas have previously been defined as the source origins of the N250 component to auditory stimuli (Parviainen et al., 2011; Hämäläinen et al., 2015) and speech processing (Ortiz-Mantilla et al., 2012). The STAs has been said to play a role in phonological (Hickok and Poeppel, 2007) and language processing (Trébuchon et al., 2013). The encoding of speech sounds in the STG was summarized in the review by Yi et al. (2019).

The source reconstruction of the P1-N250 complex showed the basic speech processing temporal and spatial dynamics of the stimulus, suggesting that these responses are more anteriorly located through time. Furthermore, our results suggest that the generators of the P1 and N250 components are different, although very closely located, with our source analysis suggesting more anterior and ventral sources for the N250 responses. The difference in source generators and topographies between the P1 and N250 responses clearly indicates two different processes. We argue that the P1 components seem to reflect the sound detection, phonetic processing, and feature extractions of each stimulus unit, whereas the N250 seems to reflect more complex processes, such as articulation processing and memory trace formation, as introduced above. The differences between the double peaks in P1 (P1 and P1-2) and N250 (N250 and N250-2) probably reflect the transitional state from one processing to the next, notably observed in the second components (P1-2 and N250-2) with slightly different auditory source locations in addition to the emergence of frontal sources. These frontal activations may reflect additional processes. These findings confirm our interpretations of the ERP responses.

The source reconstruction of the N170 shows bilateral activation of the occipital areas over the VWFA and activation of the left frontal area. The activation of VWFA as the source generator of N170 confirms previous findings. The N170 is known as the marker of visual specialization for print processing, and its relationship to the VWFA is well established in the literature (Maurer et al., 2005a; Maurer and McCandliss, 2007; Mahé et al., 2013). The left frontal activation is also in line with previous findings (Maurer et al., 2011). However, previous evidence showed a left lateralization of the N170/VWFA to be characteristic of the visual expertise of reading (Maurer et al., 2008). Interestingly, we observed bilateral activation over the occipital areas. N170 bilateral activation was previously reported in young children, indicating immature development of their reading systems (Uno et al., 2021). Our group sample of children comprise sixth-graders, who were exposed longer to print, but this group comprised both good and poor readers. Given that dyslexic readers have been shown to lack hemispheric lateralization of the N170/VWFA (Maurer et al., 2005a), the atypical activation observed in the right hemisphere in the source analysis most likely comes from the poor reader subsample. This atypical activation may also indicate an immature reading system in the RD subgroup.

The correlation analysis excluded frontal sources found in both speech ERP and reading FRP source reconstructions because they are known to be part of the attention network and the frontal eye field (Ptak, 2012).

In the reading process, N170 correlated with the reading scores, but it did not survive the statistical correction. The relationship between the N170 and reading was expected based on strong evidence in the literature showing the role of this visual component in reading and print processing (Maurer et al., 2005a; Hasko et al., 2013). In line with previous findings, correlation results between the N170 response and reading scores were found over the left occipital area. This left lateralization has also been described in the literature as the neural biomarker of the brain's sensitivity to print and word processing (Simon et al., 2007; Maurer et al., 2008; Zhao et al., 2012). However, it seems that the correlation we found was weak, as it did not survive the statistical corrections. One reason for this result is the methodological approach used in this study. As we have been computing FRPs for a group average containing 80 subjects and for multiple words, the effect may have been weakened through this averaging procedure.

The correlation analysis between cortical brain activity and reading scores in the auditory P1 response showed a significant correlation between left (primary) auditory cortex activity and reading score. Previous studies have shown that time cues and temporal acoustic information are typically processed by the left auditory cortex (Ladeira et al., 2011; Heimrath et al., 2016). Our results also suggest a left lateralization effect of the auditory P1 in response to speech stimuli, which is in line with previous findings. Interestingly, we found a negative correlation with reading skills, showing that the more active this brain area was, the lower the reading skills; this result contradicts previous findings (Shaywitz et al., 2002; Meyler et al., 2007). The smaller response observed in good readers may reflect the maturity of the neural network. Furthermore, correlations were not found in the right hemisphere for this component, which may suggest that brain activity in the right hemisphere may not be linked to reading skills.

N250-2 showed significant correlations between the reading scores and the STAs in both hemispheres. These brain areas were also shown to be part of the N250 component in typical auditory and language processing (Albrecht et al., 2000; Mody et al., 2008; Proverbio et al., 2011). This temporal activation was studied previously, and the role of the temporal areas was discussed in speech sound processing as reflecting low-level speech encoding (Hullett et al., 2016; Berezutskaya et al., 2017; Yi et al., 2019). The literature includes strong evidence of the role of the superior temporal area in reading and demonstrates the function of this brain area in relation to phonological processing in reading (Simos et al., 2000; Mesgarani et al., 2014).

All the correlations found between the auditory/speech brain activity and the reading scores or the visual/reading brain activity and the reading scores were negative. These results show that the more active the brain was, the lower the reading skills were. One possible interpretation is the recruitment of additional neuronal resources to compensate during atypical processing. Recruiting additional resources could be an adaptation to rebalance processing, as previously suggested in the literature (Lohvansuu et al., 2014). Another possible explanation for this result is the developmental phase of this age group. It has been suggested that visual reading skills follow an inverted U-shaped developmental trajectory (González et al., 2016). It is possible that in this age group, reading skills follow the inverted U-shaped curve of expertise in both the visual and auditory domains, which may explain the negative correlation.

We found correlations between brain activity to the visual stimuli and the auditory stimuli. The auditory source activity (in the STA) of the P1-2 response correlated significantly with both N170 sources in the left and right hemispheres (VWFA). The N250 sources (L STA and R STA) correlated only with the left N170 source (L VWFA). The N250-2 sources also showed correlations with the N170 sources over both hemispheres, but these correlations were weak and did not survive the statistical correction. Overall, these brain-to-brain correlation results suggest a strong relationship between the left occipital source in the reading processes and the auditory processes in both hemispheres. This result confirms our hypothesis, assuming that auditory and reading processes are interlinked and is grounded in the literature (Lin et al., 2011). Furthermore, the left lateralization found in the N250 correlation with the N170 is in line with the phonological mapping hypothesis. As this theory proposed that the left lateralization of the VWFA, the source origin of the N170 results from recruiting the left auditory language regions to link the orthography and phonology (Sacchi and Laszlo, 2016). Our correlation analysis suggests that the auditory region recruited for this purpose could be the STA as this area correlated with the VWFA. In addition, the positive correlation results suggest that both modalities behave in the same direction, so when brain activity is higher in one modality, it is also higher in the other modality. This may be interpreted by the presence of a compensatory or a complementary system that seems to act consistently across the two modalities.

Interestingly, the partial correlation analysis did not reveal a significant difference after controlling for reading. This result may indicate that the two modalities may be linked independently of the reading variable, suggesting the presence of possible common mechanism or network between the two modalities. This claim requires further investigation.

In line with our hypothesis, we found correlations between brain activity in speech processing and reading. Correlations between auditory and visual perception and reading have previously been shown on the behavioral level *via* meta-analysis (Kavale and Forness, 2000), and several studies have investigated both processes using simultaneous audiovisual stimuli. No such correlation was investigated *via* neuroimaging, as our findings showed the presence of correlation, even in independent tasks. With this method, we were able to investigate spatio-temporal processing in both processes and reveal, with high temporal accuracy, the different events, which allowed audiovisual sequential partial mapping in relation to reading. Our results confirmed earlier findings of auditory cortex responses to speech stimuli linked to reading skills, suggesting either the activation of the phonological route or the effect of learning to read through phonology still active at sixth grade when reading skills are fluent in most children. Similarly, the fusiform cortex or (STA) activity in response to print and correlation to reading skills confirms earlier findings and suggests this area is sensitive to environmental regularities, which seems to be linked to reading skills. From our results we were able to show the relation between the two routes, suggesting a link between the VWFA and STA.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, upon reasonable request.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethical Committee of the University of Jyväskylä. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

NA, OL, JH, and PL: conceptualization, writing, revising, and editing. OL: programming. NA: data collection, writing the main manuscript, and created the figures. NA and OL: data preprocessing and analysis. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Predictable (Marie Curie Innovative Training Networks, No. 641858), the Academy of Finland TULOS-program project eSeek-Internet and Learning Difficulties: A Multidisciplinary Approach for Understanding Reading in New Media (No. 274022), Niilö Helander foundation and the Department of Psychology at the University of Jyväskylä.

ACKNOWLEDGMENTS

The authors would like to thank the research assistants for their help with data collection.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2022.921977/full#supplementary-material

REFERENCES

Abrams, D. A., Nicol, T., Zecker, S., and Kraus, N. (2009). Abnormal cortical processing of the syllable rate of speech in poor readers. *J. Neurosci.* 29, 7686–7693. doi: 10.1523/JNEUROSCI.5242-08.2009

Alain, C., and Tremblay, K. (2007). The role of event-related brain potentials in assessing central auditory processing. *J. Am. Acad. Audiol.* 18, 573–589. doi: 10.3766/jaaa.18.7.5

Albrecht, R., Suchodoletz, W. V., and Uwer, R. (2000). The development of auditory evoked dipole source activity from childhood to adulthood. *Clin. Neurophysiol.* 111, 2268–2276. doi: 10.1016/S1388-2457(00)00464-8

Aravena, S., Tijms, J., Snellings, P., and van der Molen, M. W. (2018). Predicting individual differences in reading and spelling skill with artificial script-based letter–speech sound training. *J. Learn. Disabil.* 51, 552–564. doi: 10.1177/0022219417715407

Archer, K., Pammer, K., and Vidyasagar, T. R. (2020). A temporal sampling basis for visual processing in developmental dyslexia. *Front. Hum. Neurosci.* 14, 213. doi: 10.3389/fnhum.2020.00213

Baccino, T. (2011). "Chapter 47- Eye movements and concurrent event-related potentials: eye fixation-related potential investigations in reading," in *The Oxford Handbook of Eye Movements*, eds S. P. Liversedge, I. Gilchrist, and S. Everling (Oxford Library of Psychology). doi: 10.1093/oxfordhb/9780199539789.013.0047

Bell, A. J., and Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Comput.* 7, 1129–1159. doi: 10.1162/neco.1995.7.6.1129

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

Berezutskaya, J., Freudenburg, Z. V., Güçlü, U., van Gerven, M. A., and Ramsey, N. F. (2017). Neural tuning to low-level features of speech throughout the perisylvian cortex. *J. Neurosci.* 37, 7906–7920. doi: 10.1523/JNEUROSCI.0238-17.2017

Bishop, D. V. M. (2007). Using mismatch negativity to study central auditory processing in developmental language and literacy impairments: where are we, and where should we be going? *Psychol. Bull.* 133, 651. doi: 10.1037/0033-2909.133.4.651

Blau, V., Reithler, J., van Atteveldt, N., Seitz, J., Gerretsen, P., Goebel, R., et al. (2010). Deviant processing of letters and speech sounds as proximate cause of reading failure: a functional magnetic resonance imaging study of dyslexic children. *Brain.* 133, 868–879.

Blau, V., van Atteveldt, N., Ekkebus, M., Goebel, R., and Blomert, L. (2009). Reduced neural integration of letters and speech sounds links phonological and reading deficits in adult dyslexia. *Curr. Biol.* 19, 503–508. doi: 10.1016/j.cub.2009.01.065

Boersma, P., and Weenink, D. (2010). *Praat: Doing Phonetics by Computer. Version 6.2.04 [Software]*. Downloaded April 4. Available online at: http://www.praat.org/ (accessed December 18, 2021).

Boets, B., Op de Beeck, H. P., Vandermosten, M., Scott, S. K., Gillebert, C. R., Mantini, D., et al. (2013). Intact but less accessible phonetic representations in adults with dyslexia. *Science* 342, 1251–1254. doi: 10.1126/science.1244333

Boets, B., Vandermosten, M., Poelmans, H., Luts, H., Wouters, J., and Ghesquiere, P. (2011). Preschool impairments in auditory processing and speech perception uniquely predict future reading problems. *Res. Dev. Disabil.* 32, 560–570. doi: 10.1016/j.ridd.2010.12.020

Boets, B., Wouters, J., Van Wieringen, A., De Smedt, B., and Ghesquiere, P. (2008). Modelling relations between sensory processing, speech perception, orthographic and phonological ability, and literacy achievement. *Brain Lang.* 106, 29–40. doi: 10.1016/j.bandl.2007.12.004

Bonte, M., and Blomert, L. (2004b). Developmental changes in ERP correlates of spoken word recognition during early school years: a phonological priming study. *Clin. Neurophysiol.* 115, 409–423. doi: 10.1016/S1388-2457(03)00361-4

Bonte, M. L., and Blomert, L. (2004a). Developmental dyslexia: ERP correlates of anomalous phonological processing during spoken word recognition. *Cogn. Brain Res.* 21, 360–376. doi: 10.1016/j.cogbrainres.2004.06.010

Boothroyd, A. (2004). "Measuring auditory speech perception capacity in very young children," in *International Congress Series*, Vol. 1273, 292–295. doi: 10.1016/j.ics.2004.08.024

Breadmore, H. L., and Carroll, J. M. (2018). Sublexical and syntactic processing during reading: evidence from eye movements of typically developing and dyslexic readers. *J. Cogn. Psychol.* 30, 177–197. doi: 10.1080/20445911.2017.1414222

Brem, S., Bach, S., Kucian, K., Kujala, J. V., Guttorm, T. K., Martin, E., et al. (2010). Brain sensitivity to print emerges when children learn letter–speech sound correspondences. *Proc. Nat. Acad. Sci.* 107, 7939–7944. doi: 10.1073/pnas.0904402107

Breznitz, Z. (2006). *Fluency in Reading: Synchronization of Processes*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Breznitz, Z., and Meyler, A. (2003). Speed of lower-level auditory and visual processing as a basic factor in dyslexia: electrophysiological evidence. *Brain Lang.* 85, 166–184. doi: 10.1016/S0093-934X(02)00513-8

Castles, A., and Friedmann, N. (2014). Developmental dyslexia and the phonological deficit hypothesis. *Mind Lang.* 29, 270–285. doi: 10.1111/mila.12050

Catts, H. W., McIlraith, A., Bridges, M. S., and Nielsen, D. C. (2017). Viewing a phonological deficit within a multifactorial model of dyslexia. *Read. Writ.* 30, 613–629. doi: 10.1007/s11145-016-9692-2

Ceponiene, R., Alku, P., Westerfield, M., Torki, M., and Townsend, J. (2005). ERPs differentiate syllable and nonphonetic sound processing in children and adults. *Psychophysiology* 42, 391–406. doi: 10.1111/j.1469-8986.2005.00305.x

Christmann, C. A., Lachmann, T., and Steinbrink, C. (2015). Evidence for a general auditory processing deficit in developmental dyslexia from a discrimination paradigm using speech versus nonspeech sounds matched in complexity. *J. Speech Lang. Hear. Res.* 58, 107–121. doi: 10.1044/2014_JSLHR-L-14-0174

Christoforou, C., Fella, A., Leppänen, P. H., Georgiou, G. K., and Papadopoulos, T. C. (2021). Fixation-related potentials in naming speed: a combined EEG and eye-tracking study on children with dyslexia. *Clin. Neurophysiol.* 132, 2798–2807. doi: 10.1016/j.clinph.2021.08.013

Chyl, K., Kossowski, B., Debska, A., Łuniewska, M., Banaszkiewicz, A., Zelechowska, A., et al. (2018). Prereader to beginning reader: changes induced by reading acquisition in print and speech brain networks. *J. Child Psychol. Psychiatry* 59, 76–87. doi: 10.1111/jcpp.12774

Collin, C. A., Therrien, M. E., Campbell, K. B., and Hamm, J. P. (2012). Effects of band-pass spatial frequency filtering of face and object images on the amplitude of N170. *Perception* 41, 717–732. doi: 10.1068/p7056

Conlon, E. G., Wright, C. M., Norris, K., and Chekaluk, E. (2011). Does a sensory processing deficit explain counting accuracy on rapid visual sequencing tasks in adults with and without dyslexia? *Brain Cogn.* 76, 197–205. doi: 10.1016/j.bandc.2010.10.014

Degno, F., and Liversedge, S. P. (2020). Eye movements and fixation-related potentials in reading: a review. *Vision* 4, 11. doi: 10.3390/vision4010011

Démonet, J. F., Taylor, M. J., and Chaix, Y. (2004). Developmental dyslexia. *Lancet* 363, 1451–1460. doi: 10.1016/S0140-6736(04)16106-0

Desroches, A. S., Newman, R. L., Robertson, E. K., and Joanisse, M. F. (2013). Electrophysiological indices of phonological impairments in dyslexia. *J. Speech Lang. Hear. Res.* 56, 250–264 doi: 10.1044/1092-4388(2012/10-0351)

Dimigen, O., and Ehinger, B. V. (2021). Regression-based analysis of combined EEG and eye-tracking data: theory and applications. *J. Vis.* 213, 1–30. doi: 10.1167/jov.21.1.3

Dimigen, O., Kliegl, R., and Sommer, W. (2012). Trans-saccadic parafoveal preview benefits in fluent reading: a study with fixation-related brain potentials. *Neuroimage* 62, 381–393. doi: 10.1016/j.neuroimage.2012.04.006

Dimigen, O., Sommer, W., Hohlfeld, A., Jacobs, A. M., and Kliegl, R. (2011). Coregistration of eye movements and EEG in natural reading: analyses and review. *J. Exp. Psychol. Gen.* 140, 552. doi: 10.1037/a0023885

Duncan, L. G. (2018). Language and reading: the role of morpheme and phoneme awareness. *Curr. Dev. Disord. Rep.* 5, 226–234. doi: 10.1007/s40474-018-0153-2

Durante, A. S., Wieselberg, M. B., Carvalho, S., Costa, N., Pucci, B., Gudayol, N., et al. (2014). "Cortical auditory evoked potential: evaluation of speech detection in adult hearing aid users," in *Codas, Vol. 26* (Sociedade Brasileira de Fonoaudiologia), 367–373.

Eddy, M. D., Grainger, J., Holcomb, P. J., and Gabrieli, J. D. (2016). Orthographic and phonological processing in developing readers revealed by ERPs. *Psychophysiology* 53, 1776–1783. doi: 10.1111/psyp.12763

Eden, G. F., VanMeter, J. W., Rumsey, J. M., and Zeffiro, T. A. (1996). The visual deficit theory of developmental dyslexia. *Neuroimage* 4, S108–S117. doi: 10.1006/nimg.1996.0061

Ehinger, B. V., and Dimigen, O. (2019). Unfold: an integrated toolbox for overlap correction, non-linear modeling, and regression-based EEG analysis. *PeerJ* 7, e7838. doi: 10.7717/peerj.7838

Ehri, L. C. (2005). Learning to read words: theory, findings, and issues. *Sci. Stud. Read.* 9, 167–188. doi: 10.1207/s1532799xssr0902_4

Eklund, K., Torppa, M., Aro, M., Leppänen, P. H., and Lyytinen, H. (2015). Literacy skill development of children with familial risk for dyslexia through grades 2, 3, and 8. *J. Educ. Psychol.* 107, 126. doi: 10.1037/a0037121

Facoetti, A., Zorzi, M., Cestnick, L., Lorusso, M. L., Molteni, M., Paganoni, P., et al. (2006). The relationship between visuo-spatial attention and nonword reading in developmental dyslexia. *Cogn. Neuropsychol.* 23, 841–855. doi: 10.1080/02643290500483090

Farmer, M. E., and Klein, R. M. (1995). The evidence for a temporal processing deficit linked to dyslexia: a review. *Psychon. Bull. Rev.* 2, 460–493. doi: 10.3758/BF03210983

Ferree, T. C. (2006). Spherical splines and average referencing in scalp electroencephalography. *Brain Topogr.* 19, 43–52. doi: 10.1007/s10548-006-0011-0

Fraga González, G., Žarić, G., Tijms, J., Bonte, M., Blomert, L., and van der Molen, M. (2014). Brain-potential analysis of visual word recognition in dyslexics and typically reading children. *Front. Hum. Neurosci.* 8, 474. doi: 10.3389/fnhum.2014.00474

Fraga-González, G., Smit, D. J., Van der Molen, M. J., Tijms, J., Stam, C. J., de Geus, E. J., et al. (2021). Graph analysis of EEG functional connectivity networks during a letter-speech sound binding task in adult dyslexics. *Front. Psychol.* 12, 5344. doi: 10.3389/fpsyg.2021.767839

Fraga-González, G., Žari,ć, G., Tijms, J., Bonte, M., Blomert, L., and van der Molen, M.W. (2015). A randomized controlled trial on the beneficial effects of training letter-speech sound integration on reading fluency in children with dyslexia. *PLoS ONE* 10, e0143914. doi: 10.1371/journal.pone.0143914

Francisco, A. A., Jesse, A., Groen, M. A., and McQueen, J. M. (2017). A general audiovisual temporal processing deficit in adult readers with dyslexia. *J. Speech Lang. Hear. Res.* 60, 144–158. doi: 10.1044/2016_JSLHR-H-15-0375

Froyen, D., Van Atteveldt, N., and Blomert, L. (2010). Exploring the role of low level visual processing in letter-speech sound integration: a visual MMN study. *Front. Integr. Neurosci.* 4, 9. doi: 10.3389/fnint.2010.00009

Froyen, D. J., Bonte, M. L., van Atteveldt, N., and Blomert, L. (2009). The long road to automation: neurocognitive development of letter–speech sound processing. *J. Cogn. Neurosci.* 21, 567–580. doi: 10.1162/jocn.2009.21061

Gansonre, C., Højlund, A., Leminen, A., Bailey, C., and Shtyrov, Y. (2018). Task-free auditory EEG paradigm for probing multiple levels of speech processing in the brain. *Psychophysiology* 55, e13216. doi: 10.1111/psyp.13216

Georgiou, G. K., Papadopoulos, T. C., Zarouna, E., and Parrila, R. (2012). Are auditory and visual processing deficits related to developmental dyslexia? *Dyslexia* 18, 110–129. doi: 10.1002/dys.1439

Gibson, L. Y., Hogben, J. H., and Fletcher, J. (2006). Visual and auditory processing and component reading skills in developmental dyslexia. *Cogn. Neuropsychol.* 23, 621–642. doi: 10.1080/02643290500412545

Giofrè, D., Toffalini, E., Provazza, S., Calcagn,ì, A., Alto,è, G., and Roberts, D. J. (2019). Are children with developmental dyslexia all the same? A cluster analysis with more than 300 cases. *Dyslexia* 25, 284–295. doi: 10.1002/dys.1629

Godey, B., Schwartz, D., De Graaf, J. B., Chauvel, P., and Liegeois-Chauvel, C. (2001). Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: a comparison of data in the same patients. *Clin. Neurophysiol.* 112, 1850–1859. doi: 10.1016/S1388-2457(01)00636-8

González, G. F., Žari,ć, G., Tijms, J., Bonte, M., Blomert, L., Leppänen, P., et al. (2016). Responsivity to dyslexia training indexed by the N170 amplitude of the brain potential elicited by word reading. *Brain Cogn.* 106, 42–54. doi: 10.1016/j.bandc.2016.05.001

Goswami, U. (2000). Phonological representations, reading development and dyslexia: towards a cross-linguistic theoretical framework. *Dyslexia* 6, 133–151. doi: 10.1002/(SICI)1099-0909(200004/06)6:2<133::AID-DYS160>3.0.CO;2-A

Goswami, U. (2002). In the beginning was the rhyme? A reflection on Hulme, Hatcher, Nation, Brown, Adams, and Stuart (2002). *J. Exp. Child Psychol.* 82, 47–57. doi: 10.1006/jecp.2002.2673

Goswami, U. (2015). Sensory theories of developmental dyslexia: three challenges for research. *Nat. Rev. Neurosci.* 16, 43–54. doi: 10.1038/nrn3836

Goswami, U. (2019). A neural oscillations perspective on phonological development and phonological processing in developmental dyslexia. *Lang. Linguist. Compass* 13, e12328. doi: 10.1111/lnc3.12328

Gu, C., and Bi, H. Y. (2020). Auditory processing deficit in individuals with dyslexia: a meta-analysis of mismatch negativity. *Neurosci. Biobehav. Rev.* 116, 396–405. doi: 10.1016/j.neubiorev.2020.06.032

Haenschel, C., Vernon, D. J., Dwivedi, P., Gruzelier, J. H., and Baldeweg, T. (2005). Event-related brain potential correlates of human auditory sensory memory-trace formation. *J. Neurosci.* 25, 10494–10501. doi: 10.1523/JNEUROSCI.1227-05.2005

Hämäläinen, J., Landi, N., Loberg, O., Lohvansuu, K., Pugh, K., and Leppänen, P. H. (2018). Brain event-related potentials to phoneme contrasts and their correlation to reading skills in school-age children. *Int. J. Behav. Dev.* 42, 357–372. doi: 10.1177/0165025417728582

Hämäläinen, J. A., Leppänen, P. H., Guttorm, T. K., and Lyytinen, H. (2008). Event-related potentials to pitch and rise time change in children with reading disabilities and typically reading children. *Clin. Neurophysiol.* 119, 100–115. doi: 10.1016/j.clinph.2007.09.064

Hämäläinen, J. A., Leppänen, P. H. T., Guttorm, T. K., and Lyytinen, H. (2007). N1 and P2 components of auditory event-related potentials in children with and without reading disabilities. *Clin. Neurophysiol.* 118, 2263–2275. doi: 10.1016/j.clinph.2007.07.007

Hämäläinen, J. A., Lohvansuu, K., Ervast, L., and Leppänen, P. H. (2015). Event-related potentials to tones show differences between children with multiple risk factors for dyslexia and control children before the onset of formal reading instruction. *Int. J. Psychophysiol.* 95, 101–112. doi: 10.1016/j.ijpsycho.2014.04.004

Hämäläinen, J. A., Salminen, H. K., and Leppänen, P. H. (2013). Basic auditory processing deficits in dyslexia: systematic review of the behavioral and event-related potential/field evidence. *J. Learn. Disabil.* 46, 413–427. doi: 10.1177/0022219411436213

Hasko, S., Groth, K., Bruder, J., Bartling, J., and Schulte-Körne, G. (2013). The time course of reading processes in children with and without dyslexia: an ERP study. *Front. Hum. Neurosci.* 7, 570. doi: 10.3389/fnhum.2013.00570

Hayiou-Thomas, M. E., Carroll, J. M., Leavett, R., Hulme, C., and Snowling, M. J. (2017). When does speech sound disorder matter for literacy? The role of disordered speech errors, co-occurring language impairment and family risk of dyslexia. *J. Child Psychol. Psychiatry* 58, 197–205. doi: 10.1111/jcpp.12648

He, Y., Liu, X., Hu, J., Nichols, E. S., Lu, C., and Liu, L. (2021). Difference between children and adults in the print-speech coactivated network. *Sci. Stud. Read.* 1–16. doi: 10.1080/10888438.2021.1965607

Heimrath, K., Fiene, M., Rufener, K. S., and Zaehle, T. (2016). Modulating human auditory processing by transcranial electrical stimulation. *Front. Cell. Neurosci.* 10, 53. doi: 10.3389/fncel.2016.00053

Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402. doi: 10.1038/nrn2113

Hinojosa, J. A., Mercado, F., and Carreti,é, L. (2015). N170 sensitivity to facial expression: a meta-analysis. *Neurosci. Biobehav. Rev.* 55, 498–509. doi: 10.1016/j.neubiorev.2015.06.002

Hoeft, F., McCandliss, B. D., Black, J. M., Gantman, A., Zakerani, N., Hulme, C., et al. (2011). Neural systems predicting long-term outcome in dyslexia. *Proc. Nat. Acad. Sci.* 108, 361–366. doi: 10.1073/pnas.1008950108

Holloway, I. D., van Atteveldt, N., Blomert, L., and Ansari, D. (2015). Orthographic dependency in the neural correlates of reading: evidence from audiovisual integration in English readers. *Cerebral Cortex* 25, 1544–1553. doi: 10.1093/cercor/bht347

Hommet, C., Vidal, J., Roux, S., Blanc, R., Barthez, M. A., De Becque, B., et al. (2009). Topography of syllable change-detection electrophysiological indices in children and adults with reading disabilities. *Neuropsychologia* 47, 761–770. doi: 10.1016/j.neuropsychologia.2008.12.010

Hornickel, J., and Kraus, N. (2013). Unstable representation of sound: a biological marker of dyslexia. *J. Neurosci.* 33, 3500–3504. doi: 10.1523/JNEUROSCI.4205-12.2013

Hsu, C. H., Tsai, J. L., Lee, C. Y., and Tzeng, O. J. L. (2009). Orthographic combinability and phonological consistency effects in reading Chinese phonograms: an event-related potential study. *Brain Lang.* 108, 56–66. doi: 10.1016/j.bandl.2008.09.002

Hullett, P. W., Hamilton, L. S., Mesgarani, N., Schreiner, C. E., and Chang, E. F. (2016). Human superior temporal gyrus organization of spectrotemporal modulation tuning derived from speech stimuli. *J. Neurosci.* 36, 2014–2026. doi: 10.1523/JNEUROSCI.1779-15.2016

Hulme, C., and Snowling, M. J. (2014). The interface between spoken and written language: developmental disorders. *Philos. Trans. R. Soc. B Biol. Sci.* 369, 20120395. doi: 10.1098/rstb.2012.0395

Jainta, S., and Kapoula, Z. (2011). Dyslexic children are confronted with unstable binocular fixation while reading. *PLoS ONE* 6, e18694. doi: 10.1371/journal.pone.0018694

Jaramillo, M., Paavilainen, P., and Näätänen, R. (2000). Mismatch negativity and behavioural discrimination in humans as a function of the magnitude of change in sound duration. *Neurosci. Lett.* 290, 101–104. doi: 10.1016/S0304-3940(00)01344-6

Jarodzka, H., and Brand-Gruwel, S. (2017). Tracking the reading eye: towards a model of real-world reading. *J. Comp. Assist. Learn.* 33, 193–201. doi: 10.1111/jcal.12189

Jednoróg, K., Marchewka, A., Tacikowski, P., and Grabowska, A. (2010). Implicit phonological and semantic processing in children with developmental dyslexia: evidence from event-related potentials. *Neuropsychologia* 48, 2447–2457. doi: 10.1016/j.neuropsychologia.2010.04.017

Kanniainen, L., Kiili, C., Tolvanen, A., Aro, M., and Leppänen, P. H. (2019). Literacy skills and online research and comprehension: struggling readers face difficulties online. *Read. Writ.* 32, 2201–2222. doi: 10.1007/s11145-019-09944-9

Karhu, J., Herrgård, E., Pääkkönen, A., Luoma, L., Airaksinen, E., and Partanen, J. (1997). Dual cerebral processing of elementary auditory input in children. *Neuroreport* 8, 1327–1330. doi: 10.1097/00001756-199704140-00002

Karipidis, I. I., Pleisch, G., Brandeis, D., Roth, A., Röthlisberger, M., Schneebeli, M., et al. (2018). Simulating reading acquisition: the link between reading outcome and multimodal brain signatures of letter–speech sound learning in prereaders. *Sci. Rep.* 8, 1–13. doi: 10.1038/s41598-018-24909-8

Karipidis, I. I., Pleisch, G., Röthlisberger, M., Hofstetter, C., Dornbierer, D., Stämpfli, P., et al. (2017). Neural initialization of audiovisual integration in prereaders at varying risk for developmental dyslexia. *Hum. Brain Mapp.* 38, 1038–1055. doi: 10.1002/hbm.23437

Kavale, K. A., and Forness, S. R. (2000). Auditory and visual perception processes and reading ability: a quantitative reanalysis and historical reinterpretation. *Learn. Disabil. Q.* 23, 253–270. doi: 10.2307/1511348

Khan, A., Hämäläinen, J. A., Leppänen, P. H., and Lyytinen, H. (2011). Auditory event-related potentials show altered hemispheric responses in dyslexia. *Neurosci. Lett.* 498, 127–132. doi: 10.1016/j.neulet.2011.04.074

Kim, S., and Lombardino, L. J. (2016). Simple sentence reading and specific cognitive functions in college students with dyslexia: an eye-tracking study. *Clin. Arch. Commun. Disord.* 1, 48–61. doi: 10.21849/cacd.2016.00073

Klenberg, L., Jäms,ä, S., Häyrinen, T., Lahti-Nuuttila, P. E. K. K.A., and Korkman, M. (2010). The attention and executive function rating inventory (ATTEX): psychometric properties and clinical utility in diagnosing ADHD subtypes. *Scand. J. Psychol.* 51, 439–448. doi: 10.1111/j.1467-9450.2010.00812.x

Kornrumpf, B., Niefind, F., Sommer, W., and Dimigen, O. (2016). Neural correlates of word recognition: a systematic comparison of natural reading and rapid serial visual presentation. *J. Cogn. Neurosci.* 28, 1374–1391. doi: 10.1162/jocn_a_00977

Kronschnabel, J., Brem, S., Maurer, U., and Brandeis, D. (2014). The level of audiovisual print–speech integration deficits in dyslexia. *Neuropsychologia* 62, 245–261. doi: 10.1016/j.neuropsychologia.2014.07.024

Kujala, T., Myllyviita, K., Tervaniemi, M., Alho, K., Kallio, J., and Näätänen, R. (2000). Basic auditory dysfunction in dyslexia as demonstrated by brain activity measurements. *Psychophysiology* 37, 262–266. doi: 10.1111/1469-8986.3720262

Kuuluvainen, S., Leminen, A., and Kujala, T. (2016). Auditory evoked potentials to speech and nonspeech stimuli are associated with verbal skills in preschoolers. *Dev. Cogn. Neurosci.* 19, 223–232. doi: 10.1016/j.dcn.2016.04.001

LaBerge, D., and Samuels, S. J. (2017). *Basic Processes in Reading: Perception and Comprehension*. London: Routledge. doi: 10.4324/9781315467610

Lachmann, T., Berti, S., Kujala, T., and Schröger, E. (2005). Diagnostic subgroups of developmental dyslexia have different deficits in neural processing of tones and phonemes. *Int. J. Psychophysiol.* 56, 105–120. doi: 10.1016/j.ijpsycho.2004.11.005

Ladeira, A., Fregni, F., Campanh,ã, C., Valasek, C. A., De Ridder, D., Brunoni, A. R., et al. (2011). Polarity-dependent transcranial direct current stimulation effects on central auditory processing. *PLoS ONE* 6, e25399. doi: 10.1371/journal.pone.0025399

Leppänen, P. H., Richardson, U., Pihko, E., Eklund, K. M., Guttorm, T. K., Aro, M., et al. (2002). Brain responses to changes in speech sound durations differ between infants with and without familial risk for dyslexia. *Dev. Neuropsychol.* 22, 407–422. doi: 10.1207/S15326942dn2201_4

Leppänen, P. H., Tóth, D., Honbolyg,ó, F., Lohvansuu, K., Hämäläinen, J. A., Demonet, J. F., et al. (2019). Reproducibility of brain responses: high for speech perception, low for reading difficulties. *Sci. Rep.* 9, 1–12. doi: 10.1038/s41598-019-41992-7

Lin, S. E., Chen, H. C., Zhao, J., Li, S., He, S., and Weng, X. C. (2011). Left-lateralized N170 response to unpronounceable pseudo but not false Chinese characters – the key role of orthography. *Neuroscience* 190, 200–206. doi: 10.1016/j.neuroscience.2011.05.071

Lindeman, J. (1998). *ALLU – Ala-asteen lukutesti [ALLU – Reading Test for Primary School]*. University of Turku, Finland: The Center for Learning Research.

Lizarazu, M., Lallier, M., Molinaro, N., Bourguignon, M., Paz-Alonso, P. M., Lerma-Usabiaga, G., et al. (2015). Developmental evaluation of atypical auditory sampling in dyslexia: functional and structural evidence. *Hum. Brain Mapp.* 36, 4986–5002. doi: 10.1002/hbm.22986

Loberg, O., Hautala, J., Hämäläinen, J. A., and Leppänen, P. H. (2019). Influence of reading skill and word length on fixation-related brain activity in school-aged children during natural reading. *Vision Res.* 165, 109–122. doi: 10.1016/j.visres.2019.07.008

Lobier, M., Zoubrinetzky, R., and Valdois, S. (2012). The visual attention span deficit in dyslexia is visual and not verbal. *Cortex* 48, 768–773. doi: 10.1016/j.cortex.2011.09.003

Lobier, M. A., Peyrin, C., Pichat, C., Le Bas, J. F., and Valdois, S. (2014). Visual processing of multiple elements in the dyslexic brain: evidence for a superior parietal dysfunction. *Front. Hum. Neurosci.* 8, 479. doi: 10.3389/fnhum.2014.00479

Lochy, A., Van Reybroeck, M., and Rossion, B. (2016). Left cortical specialization for visual letter strings predicts rudimentary knowledge of letter-sound association in preschoolers. *Proc. Nat. Acad. Sci.* 113, 8544–8549. doi: 10.1073/pnas.1520366113

Lohvansuu, K., Hämäläinen, J. A., Ervast, L., Lyytinen, H., and Leppänen, P. H. (2018). Longitudinal interactions between brain and cognitive measures on reading development from 6 months to 14 years. *Neuropsychologia* 108, 6–12. doi: 10.1016/j.neuropsychologia.2017.11.018

Lohvansuu, K., Hämäläinen, J. A., Tanskanen, A., Ervast, L., Heikkinen, E., Lyytinen, H., et al. (2014). Enhancement of brain event-related potentials to speech sounds is associated with compensated reading skills in dyslexic children with familial risk for dyslexia. *Int. J. Psychophysiol.* 94, 298–310. doi: 10.1016/j.ijpsycho.2014.10.002

Mahé, G., Bonnefond, A., and Doignon-Camus, N. (2013). Is the impaired N170 print tuning specific to developmental dyslexia? A matched reading-level study with poor readers and dyslexics. *Brain Lang.* 127, 539–544. doi: 10.1016/j.bandl.2013.09.012

Malins, J. G., Pugh, K. R., Buis, B., Frost, S. J., Hoeft, F., Landi, N., et al. (2018). Individual differences in reading skill are related to trial-by-trial neural activation variability in the reading network. *J. Neurosci.* 38, 2981–2989. doi: 10.1523/JNEUROSCI.0907-17.2018

Maurer, J., Collet, L., Pelster, H., Truy, E., and Gallégo, S. (2002). Auditory late cortical response and speech recognition in digisonic cochlear implant users. *Laryngoscope* 112, 2220–2224. doi: 10.1097/00005537-200212000-00017

Maurer, U., Blau, V. C., Yoncheva, Y. N., and McCandliss, B. D. (2010). Development of visual expertise for reading: rapid emergence of visual familiarity for an artificial script. *Dev. Neuropsychol.* 35, 404–422. doi: 10.1080/87565641.2010.480916

Maurer, U., Brandeis, D., and McCandliss, B. D. (2005a). Fast, visual specialization for reading in English revealed by the topography of the N170 ERP response. *Behav. Brain Funct.* 1, 1–12. doi: 10.1186/1744-9081-1-13

Maurer, U., Brem, S., Bucher, K., and Brandeis, D. (2005b). Emerging neurophysiological specialization for letter strings. *J. Cogn. Neurosci.* 17, 1532–1552. doi: 10.1162/089892905774597218

Maurer, U., and McCandliss, B. D. (2007). "The development of visual expertise for words: the contribution of electrophysiology," in *Single-Word Reading* (Psychology Press), 57–77.

Maurer, U., Schulz, E., Brem, S., van der Mark, S., Bucher, K., Martin, E., et al. (2011). The development of print tuning in children with dyslexia: evidence from longitudinal ERP data supported by fMRI. *Neuroimage* 57, 714–722. doi: 10.1016/j.neuroimage.2010.10.055

Maurer, U., Zevin, J. D., and McCandliss, B. D. (2008). Left-lateralized N170 effects of visual expertise in reading: evidence from Japanese syllabic and logographic scripts. *J. Cogn. Neurosci.* 20, 1878–1891. doi: 10.1162/jocn.2008.20125

Meng, X., Sai, X., Wang, C., Wang, J., Sha, S., and Zhou, X. (2005). Auditory and speech processing and reading development in Chinese school children: behavioural and ERP evidence. *Dyslexia* 11, 292–310. doi: 10.1002/dys.309

Mesgarani, N., Cheung, C., Johnson, K., and Chang, E. F. (2014). Phonetic feature encoding in human superior temporal gyrus. *Science* 343, 1006–1010. doi: 10.1126/science.1245994

Meyler, A., Keller, T. A., Cherkassky, V. L., Lee, D., Hoeft, F., Whitfield-Gabrieli, S., et al. (2007). Brain activation during sentence comprehension among good and poor readers. *Cerebral Cortex* 17, 2780–2787. doi: 10.1093/cercor/bhm006

Mody, M., Wehner, D. T., and Ahlfors, S. P. (2008). Auditory word perception in sentence context in reading-disabled children. *Neuroreport* 19, 1567. doi: 10.1097/WNR.0b013e328311ca04

Moll, K., Hasko, S., Groth, K., Bartling, J., and Schulte-Körne, G. (2016). Sound processing deficits in children with developmental dyslexia: an ERP study. *Clin. Neurophysiol.* 127, 1989–2000. doi: 10.1016/j.clinph.2016.01.005

Monzalvo, K., Fluss, J., Billard, C., Dehaene, S., and Dehaene-Lambertz, G. (2012). Cortical networks for vision and language in dyslexic and normal children of variable socio-economic status. *Neuroimage* 61, 258–274. doi: 10.1016/j.neuroimage.2012.02.035

Näätänen, R., and Rinne, T. (2002). Electric brain response to sound repetition in humans: an index of long-term-memory-trace formation? *Neurosci. Lett.* 318, 49–51. doi: 10.1016/S0304-3940(01)02438-7

Navas, A. L. G. P., Ferraz, É. D. C., and Borges, J. P. A. (2014). "Phonological processing deficits as a universal model for dyslexia: evidence from different orthographies," in *CoDAS, Vol. 26* (Sociedade Brasileira de Fonoaudiologia), 509–519.

Nevala, J., and Lyytinen, H. (2000). *Sanaketjutesti [Word chain test]*. Jyväskylä, Finland: Niilo Mäki Instituutti ja Jyväskylän yliopiston lapsitutkimuskeskus [Niilo Mäki Institute and Child Research Center of University of Jyväskylä].

Nilsson Benfatto, M., Öqvist Seimyr, G., Ygge, J., Pansell, T., Rydberg, A., and Jacobson, C. (2016). Screening for dyslexia using eye tracking during reading. *PLoS ONE* 11, e0165508. doi: 10.1371/journal.pone.0165508

Noordenbos, M. W., Segers, E., Serniclaes, W., Mitterer, H., and Verhoeven, L. (2012). Neural evidence of allophonic perception in children at risk for dyslexia. *Neuropsychologia* 50, 2010–2017. doi: 10.1016/j.neuropsychologia.2012.04.026

Norton, E. S., Beach, S. D., and Gabrieli, J. D. (2015). Neurobiology of dyslexia. *Curr. Opin. Neurobiol.* 30, 73–78. doi: 10.1016/j.conb.2014.09.007

Ortiz-Mantilla, S., Hämäläinen, J. A., and Benasich, A. A. (2012). Time course of ERP generators to syllables in infants: a source localization study using age-appropriate brain templates. *Neuroimage* 59, 3275–3287. doi: 10.1016/j.neuroimage.2011.11.048

Ozeri-Rotstain, A., Shachaf, I., Farah, R., and Horowitz-Kraus, T. (2020). Relationship between eye-movement patterns, cognitive load, and reading ability in children with reading difficulties. *J. Psycholinguist. Res.* 49, 491–507. doi: 10.1007/s10936-020-09705-8

Parviainen, T., Helenius, P., Poskiparta, E., Niemi, P., and Salmelin, R. (2011). Speech perception in the child brain: cortical timing and its relevance to literacy acquisition. *Hum. Brain Mapp.* 32, 2193–2206. doi: 10.1002/hbm.21181

Pennington, B. F., and Bishop, D. V. (2009). Relations among speech, language, and reading disorders. *Annu. Rev. Psychol.* 60, 283–306. doi: 10.1146/annurev.psych.60.110707.163548

Peterson, R. L., and Pennington, B. F. (2015). Developmental dyslexia. *Annu. Rev. Clin. Psychol.* (New York, NY: Oxford university Press) 11, 283–307.

Pleisch, G., Karipidis, I. I., Brem, A., Röthlisberger, M., Roth, A., Brandeis, D., et al. (2019). Simultaneous EEG and fMRI reveals stronger sensitivity to orthographic strings in the left occipito-temporal cortex of typical versus poor beginning readers. *Dev. Cogn. Neurosci.* 40, 100717. doi: 10.1016/j.dcn.2019.100717

Prestes, M. R. D., and Feitosa, M. A. G. (2017). Theories of dyslexia: support by changes in Auditory perception1. *Psicologia Teoria e Pesquisa* 32, 1–9. doi: 10.1590/0102-3772e32ne24

Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage* 62, 816–847. doi: 10.1016/j.neuroimage.2012.04.062

Proverbio, A. M., D'Aniello, G. E., Adorni, R., and Zani, A. (2011). When a photograph can be heard: vision activates the auditory cortex within 110 ms. *Sci. Rep.* 1, 1–11. doi: 10.1038/srep00054

Ptak, R. (2012). The frontoparietal attention network of the human brain: action, saliency, and a priority map of the environment. *Neuroscientist* 18, 502–515. doi: 10.1177/1073858411409051

Ramus, F. (2014). Neuroimaging sheds new light on the phonological deficit in dyslexia. *Trends Cogn. Sci.* 18, 274–275. doi: 10.1016/j.tics.2014.01.009

Ramus, F., and Szenkovits, G. (2008). What phonological deficit? *Q. J. Exp. Psychol.* 61, 129–141. doi: 10.1080/17470210701508822

Raven, J. C., and Court, J. H. (1998). *Raven's Progressive Matrices and Vocabulary Scales*, Vol. 759. Oxford: Oxford Psychologists Press.

Richardson, U., Thomson, J. M., Scott, S. K., and Goswami, U. (2004). Auditory processing skills and phonological representation in dyslexic children. *Dyslexia.* 10, 215–233.

Robertson, E. K., and Gallant, J. E. (2019). Eye tracking reveals subtle spoken sentence comprehension problems in children with dyslexia. *Lingua* 228, 102708. doi: 10.1016/j.lingua.2019.06.009

Rossion, B., Curran, T., and Gauthier, I. (2002). A defense of the subordinate-level expertise account for the N170 component. *Cognition* 85, 189–196. doi: 10.1016/S0010-0277(02)00101-4

Rossion, B., Joyce, C. A., Cottrell, G. W., and Tarr, M. J. (2003). Early lateralization and orientation tuning for face, word, and object processing in the visual cortex. *Neuroimage* 20, 1609–1624. doi: 10.1016/j.neuroimage.2003.07.010

Ruhnau, P., Herrmann, B., Maess, B., and Schröger, E. (2011). Maturation of obligatory auditory responses and their neural sources: evidence from EEG and MEG. *Neuroimage* 58, 630–639. doi: 10.1016/j.neuroimage.2011.06.050

Rüsseler, J., Ye, Z., Gerth, I., Szycik, G. R., and Münte, T. F. (2018). Audio-visual speech perception in adult readers with dyslexia: an fMRI study. *Brain Imaging Behav.* 12, 357–368. doi: 10.1007/s11682-017-9694-y

Sacchi, E., and Laszlo, S. (2016). An event-related potential study of the relationship between N170 lateralization and phonological awareness in developing readers. *Neuropsychologia* 91, 415–425. doi: 10.1016/j.neuropsychologia.2016.09.001

Salmelin, R., Kiesil,ä, P., Uutela, K., Service, E., and Salonen, O. (1996). Impaired visual word processing in dyslexia revealed with magnetoencephalography. *Ann. Neurol.* 40, 157–162. doi: 10.1002/ana.410400206

Sánchez-Vincitore, L. V., Avery, T., and Froud, K. (2018). Word-related N170 responses to implicit and explicit reading tasks in neoliterate adults. *Int. J. Behav. Dev.* 42, 321–332. doi: 10.1177/0165025417714063

Schulte-Körne, G., Deimel, W., Bartling, J., and Remschmidt, H. (1998). Auditory processing and dyslexia: evidence for a specific speech processing deficit. *Neuroreport* 9, 337–340. doi: 10.1097/00001756-199801260-00029

Shahin, A., Roberts, L. E., and Trainor, L. J. (2004). Enhancement of auditory cortical development by musical experience in children. *Neuroreport* 15, 1917–1921. doi: 10.1097/00001756-200408260-00017

Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Mencl, W. E., Fulbright, R. K., Skudlarski, P., et al. (2002). Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biol. Psychiatry* 52, 101–110. doi: 10.1016/S0006-3223(02)01365-3

Siegel, L. S. (2006). Perspectives on dyslexia. *Paediatr. Child Health* 11, 581–587. doi: 10.1093/pch/11.9.581

Simon, G., Petit, L., Bernard, C., and Reba,ï, M. (2007). N170 ERPs could represent a logographic processing strategy in visual word recognition. *Behav. Brain Funct.* 3, 1–11. doi: 10.1186/1744-9081-3-21

Simos, P. G., Breier, J. I., Wheless, J. W., Maggio, W. W., Fletcher, J. M., Castillo, E. M., et al. (2000). Brain mechanisms for reading: the role of the superior temporal gyrus in word and pseudoword naming. *Neuroreport* 11, 2443–2446. doi: 10.1097/00001756-200008030-00021

Smith-Spark, J. H., Henry, L. A., Messer, D. J., and Ziecik, A. P. (2017). Verbal and non-verbal fluency in adults with developmental dyslexia: phonological processing or executive control problems? *Dyslexia* 23, 234–250. doi: 10.1002/dys.1558

Snowling, M. (1998). Dyslexia as a phonological deficit: evidence and implications. *Child Psychol. Psychiatry Rev.* 3, 4–11. doi: 10.1017/S1360641797001366

Snowling, M. J. (2001). From language to reading and dyslexia 1. *Dyslexia* 7, 37–46. doi: 10.1002/dys.185

Snowling, M. J., Hulme, C., and Nation, K. (2020). Defining and understanding dyslexia: past, present and future. *Oxford Rev. Educ.* 46, 501–513. doi: 10.1080/03054985.2020.1765756

Stein, J. (2018). What is developmental dyslexia? *Brain Sci.* 8, 26. doi: 10.3390/brainsci8020026

Tiffin-Richards, S. P., and Schroeder, S. (2015). Children's and adults' parafoveal processes in German: phonological and orthographic effects. *J. Cogn. Psychol.* 27, 531–548. doi: 10.1080/20445911.2014.999076

Trébuchon, A., Démonet, J. F., Chauvel, P., and Liégeois-Chauvel, C. (2013). Ventral and dorsal pathways of speech perception: an intracerebral ERP study. *Brain Lang.* 127, 273–283. doi: 10.1016/j.bandl.2013.04.007

Uno, T., Kasai, T., and Seki, A. (2021). The developmental change of print-tuned N170 in highly transparent writing systems 1. *Jpn. Psychol. Res.* doi: 10.1111/jpr.12397

Van Atteveldt, N., Formisano, E., Goebel, R., and Blomert, L. (2004). Integration of letters and speech sounds in the human brain. *Neuron* 43, 271–282. doi: 10.1016/j.neuron.2004.06.025

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., and Scanlon, D. M. (2004). Specific reading disability (dyslexia): what have we learned in the past four decades? *J. Child Psychol. Psychiatry* 45, 2–40. doi: 10.1046/j.0021-9630.2003.00305.x

Vernon, M. D. (2016). *Backwardness in Reading*. Cambridge: Cambridge University Press.

Vidal, J., Bonnet-Brilhault, F., Roux, S., and Bruneau, N. (2005). Auditory evoked potentials to tones and syllables in adults: evidence of specific influence on N250 wave. *Neurosci. Lett.* 378, 145–149. doi: 10.1016/j.neulet.2004.12.022

Vogel, A. C., Church, J. A., Power, J. D., Miezin, F. M., Petersen, S. E., and Schlaggar, B. L. (2013). Functional network architecture of reading-related regions across development. *Brain Lang.* 125, 231–243. doi: 10.1016/j.bandl.2012.12.016

Wang, Y., Huang, H., Yang, H., Xu, J., Mo, S., Lai, H., et al. (2019). Influence of EEG references on N170 component in human facial recognition. *Front. Neurosci.* 13, 705. doi: 10.3389/fnins.2019.00705

Wass, S. V., Daubney, K., Golan, J., Logan, F., and Kushnerenko, E. (2019). Elevated physiological arousal is associated with larger but more variable neural responses to small acoustic change in children during a passive auditory attention task. *Dev. Cogn. Neurosci.* 37, 100612. doi: 10.1016/j.dcn.2018.12.010

Wenzel, M. A., Golenia, J. E., and Blankertz, B. (2016). Classification of eye fixation related potentials for variable stimulus saliency. *Front. Neurosci.* 10, 23. doi: 10.3389/fnins.2016.00023

Wright, C. M., and Conlon, E. G. (2009). Auditory and visual processing in children with dyslexia. *Dev. Neuropsychol.* 34, 330–355. doi: 10.1080/87565640902801882

Yang, Y., Yang, Y. H., Li, J., Xu, M., and Bi, H. Y. (2020). An audiovisual integration deficit underlies reading failure in nontransparent writing systems: an fMRI study of Chinese children with dyslexia. *J. Neurolinguistics* 54, 100884. doi: 10.1016/j.jneuroling.2019.100884

Ye, Z., Rüsseler, J., Gerth, I., and Münte, T. F. (2017). Audiovisual speech integration in the superior temporal region is dysfunctional in dyslexia. *Neuroscience* 356, 1–10. doi: 10.1016/j.neuroscience.2017.05.017

Yi, H. G., Leonard, M. K., and Chang, E. F. (2019). The encoding of speech sounds in the superior temporal gyrus. *Neuron* 102, 1096–1110. doi: 10.1016/j.neuron.2019.04.023

Ylinen, S., Junttila, K., Laasonen, M., Iverson, P., Ahonen, L., and Kujala, T. (2019). Diminished brain responses to second-language words are linked with native-language literacy skills in dyslexia. *Neuropsychologia* 122, 105–115. doi: 10.1016/j.neuropsychologia.2018.11.005

Žarić, G., Fraga González, G., Tijms, J., van der Molen, M. W., Blomert, L., and Bonte, M. (2014). Reduced neural integration of letters and speech sounds in dyslexic children scales with individual differences in reading fluency. *PLoS ONE* 9, e110337. doi: 10.1371/journal.pone.0110337

Zhang, G., Yuan, B., Hua, H., Lou, Y., Lin, N., and Li, X. (2021a). Individual differences in first-pass fixation duration in reading are related to resting-state functional connectivity. *Brain Lang.* 213, 104893. doi: 10.1016/j.bandl.2020.104893

Zhang, J., and McBride-Chang, C. (2010). Auditory sensitivity, speech perception, and reading development and impairment. *Educ. Psychol. Rev.* 22, 323–338. doi: 10.1007/s10648-010-9137-4

Zhang, R., Wang, Z., Wang, X., and Yang, J. (2021b). N170 adaptation effect of the sub-lexical phonological and semantic processing in Chinese character reading. *Acta Psychol. Sinica* 53, 807. doi: 10.3724/SP.J.1041.2021.00807

Zhao, B., Dang, J., and Zhang, G. (2016). "EEG evidence for a three-phase recurrent process during spoken word processing," in *2016 10th International Symposium on Chinese Spoken Language Processing (ISCSLP)* (IEEE), 1–5.

Zhao, J., Kipp, K., Gaspar, C., Maurer, U., Weng, X., Mecklinger, A., et al. (2014). Fine neural tuning for orthographic properties of words emerges early in children reading alphabetic script. *J. Cogn. Neurosci.* 26, 2431–2442. doi: 10.1162/jocn_a_00660

Zhao, J., Li, S., Lin, S. E., Cao, X. H., He, S., and Weng, X. C. (2012). Selectivity of N170 in the left hemisphere as an electrophysiological marker for expertise in reading Chinese. *Neurosci. Bull.* 28, 577–584. doi: 10.1007/s12264-012-1274-y

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Neuroscience

TYPE Original Research
PUBLISHED 24 August 2022
DOI 10.3389/fnins.2022.983084

OPEN ACCESS

EDITED BY
Li-Hai Tan,
Shenzhen Institute of Neuroscience,
China

REVIEWED BY
Wai Ting Siok,
The University of Hong Kong,
Hong Kong SAR, China
Xinyang Liu,
East China Normal University, China
Hong-Yan Bi,
Institute of Psychology (CAS), China

*CORRESPONDENCE
Sha Tao
taosha@bnu.edu.cn

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Neurodevelopment, a section of the journal Frontiers in Neuroscience

RECEIVED 30 June 2022
ACCEPTED 08 August 2022
PUBLISHED 24 August 2022

CITATION
Wang Y, Luo J, Ma L, Chen R, Wang J, Chu C, Men W, Tan S, Gao J-H, Qin S, He Y, Dong Q and Tao S (2022) Learning to read Chinese promotes two cortico-subcortical pathways: The development of thalamo-occipital and fronto-striatal circuits. *Front. Neurosci.* 16:983084. doi: 10.3389/fnins.2022.983084

Learning to read Chinese promotes two cortico-subcortical pathways: The development of thalamo-occipital and fronto-striatal circuits

Yanpei Wang[1,2†], Jie Luo[3†], Leilei Ma[1,2†], Rui Chen[1,2], Jiali Wang[1,2], Congying Chu[1,2], Weiwei Men[4], Shuping Tan[5], Jia-Hong Gao[4], Shaozheng Qin[1,2], Yong He[1,2], Qi Dong[1,2] and Sha Tao[1,2]*

[1]State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China, [2]IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, China, [3]Department of Psychological Sciences, University of Connecticut, Storrs, CT, United States, [4]Center for MRI Research, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China, [5]Psychiatry Research Center, Beijing HuiLongGuan Hospital, Peking University, Beijing, China

Learning to read may result in network reorganization in the developing brain. The thalamus and striatum are two important subcortical structures involved in learning to read. It remains unclear whether the thalamus and striatum may form two independent cortico-subcortical reading pathways during reading acquisition. In this prospective longitudinal study, we aimed to identify whether there may be two independent cortico-subcortical reading pathways involving the thalamus and striatum and to examine the longitudinal predictions between these two cortico-subcortical pathways and reading development in school-age children using cross-lagged panel modeling. A total of 334 children aged 6–12 years completed two reading assessments and resting functional imaging scans at approximately 12-month intervals. The results showed that there were two independent cortico-subcortical pathways, the thalamo-occipital and fronto-striatal circuits. The former may be part of a visual pathway and was predicted longitudinally by reading ability, and the prediction was stronger in children in lower grades and weaker in children in higher grades. The latter may be part of a cognitive pathway related to attention, memory, and reasoning, which was bidirectionally predicted with reading ability, and the predictive effect gradually increasing with reading development. These results extend previous findings on the relationship between functional connectivity and reading competence in children, highlighting the dynamic relationships between the thalamo-occipital and fronto-striatal circuits and reading acquisition.

KEYWORDS

learning to read, thalamo-occipital circuits, fronto-striatal circuits, school-age children, longitudinal development

Introduction

Learning to read may result in brain network reorganization (Houde et al., 2010; Martin et al., 2015). In one resting-state functional magnetic resonance imaging (fMRI) study of illiterate adults, Skeide et al. (2017) observed that a 6-month literacy intervention altered the cortico-subcortical crosstalk in the visual system of illiterate individuals. Some studies, including studies based on comparisons between children and adults (Koyama et al., 2020), literacy training studies (Alcauter et al., 2017; Hancock et al., 2017; Koyama et al., 2020; Mohammadi et al., 2020), and child development studies, (Alcauter et al., 2017) also found that cortico-subcortical alterations play an important role at the early stage of learning to read. Among these alterations, the most frequently affected subcortical structures were the striatum and thalamus. However, no study has explored the respective roles of cortico-thalamic and the cortico-striatal connection in the early stages of reading development.

The thalamus is a large mass of gray matter located in the dorsal part of the diencephalon. Nerve fibers project out of the thalamus to the cerebral cortex in various directions, allowing hub-like exchanges of information. The thalamus is critical for the detection of visual changes (Rima and Schmid, 2020). Effective temporal and spatial interpretation of text by the top-down attention network of the visual system is a critical early stage of reading, and any lesions that impair this process can lead to dyslexia (Vidyasagar, 2019). Actually, there is evidence that the important role of the thalamus is mainly reflected early in the stage of reading development. Koyama et al. (2011) found that the thalamus is specific to reading processing brain regions in children based on the results of two meta-analyses (Bolger et al., 2005; Houde et al., 2010). Siok et al. (2020) depicted a lifespan developmental trajectory of the activation intensity of related brain regions during reading task execution, which indicated that thalamic activation is gradually reduced. Correlation analysis of the thalamo-cortical visual pathway (functional connectivity) and reading competence showed a significant positive relationship in children but a non-significant negative association in adults (Koyama et al., 2011). This suggests that children's reading ability relies on the thalamo-cortical visual pathway, which does not appear to be necessarily beneficial for reading in adults. However, how the role of the thalamo-cortical visual pathway in learning to read is gradually changing and the bidirectional relationship between the two regions remains unclear, which requires further confirmation from a longitudinal study.

The striatum, as a part of the basal ganglia, receives information from the cortex and forms the corticostriatal loop projecting to the frontal lobe. The striatum is associated with semantic, phonological, and articulatory processing while reading (Xu et al., 2005; Binder et al., 2006; Bitan et al., 2007; Brem et al., 2009). The dysfunction of fronto-striatal circuits was confirmed by a meta-analysis to cause fundamental impairments in reading-related processing (Hancock et al., 2017). Fronto-striatal functional connectivity was significantly weaker in illiterate individuals than in literate controls (Mohammadi et al., 2020). Some previous studies have suggested that the striatum may be more involved in reading in adults than in children. For example, a lifespan fMRI study found that the striatum was activated in adults when reading but not in children (Siok et al., 2020). Several meta-analyses of reading task-based fMRI studies failed to identify striatum activation in children (Houde et al., 2010; Li and Bi, 2022) and found striatum activation specific to adults (Richlan et al., 2011). However, a study found that fronto-striatal functional connectivity significantly predicted reading performance in children aged 6-9 years (Alcauter et al., 2017). Thus, the striatum may be mainly activated for adult reading, and the fronto-striatal circuits may also be involved in reading when children are learning to read. However, the above studies are based on cross-sectional data, and how the roles of the fronto-striatal circuit develop and change during when learning to read remains unclear and requires longitudinal studies to offer clear evidence.

The thalamo-occipital circuit is an important visuospatial pathway and involves visual processing and visual pathway reorganization in early reading (Muller-Axt et al., 2017; Skeide et al., 2017; Tschentscher et al., 2019), and its damage can cause blindsight and developmental dyslexia (Rima and Schmid, 2020). Fronto-striatal connectivity is a critical cortico-subcortical pathway involved in language and cognitive processing (Gordon et al., 2021). The fronto-striatal pathway has been shown to be closely associated with a variety of cognitive abilities, including inhibitory control (Ojha et al., 2022), working memory(Rodrigue et al., 2020; Hidalgo-Lopez and Pletzer, 2021), executive function (Galandra et al., 2019), and cognitive flexibility (Banaie Boroujeni, 2021); its impairment can lead to problems, such as attention-deficit/hyperactivity (Cupertino et al., 2020; Mamiya et al., 2021) and reading disorder (Hancock et al., 2017). From recent evidence, the thalamo-occipital and fronto-striatal connectivities seem to be two functional independent pathways. However, it is not clear whether the two pathways maintain functional independence in the processes of learning to read and further developing reading skills.

Based on the previous research mentioned above, there may be two important cortico-subcortical pathways relevant to learning to read that may involve the thalamus and striatum. More research is needed to examine how thalamo-cortical and cortico-striatal pathways play roles in children's process of learning to read. In this study, we conducted a longitudinal brain-behavioral study among several hundred school-aged children, varying in ability from beginning readers to intermediate readers, and used cross-lagged panel analyses to explore how the two cortico-subcortical pathways and reading development may predict each other over one year. Further association analyses were conducted to examine the cognitive basis of both pathways. We hypothesized that the thalamo-occipital and fronto-striatal circuits may be two important pathways involved in learning to read. The former may be part of visual-spatial processing that plays important roles in the early stage of learning to read and gradually weakens with the development of reading ability. In contrast, the latter may be a complex cognitive pathway that always plays an important role in learning to read and gradually strengthens with the development of reading.

Materials and methods

Participants

Neuroimaging and behavioral data were obtained from the Children School Functions and Brain Development Project (CBD, Beijing Cohort: Tao, 2019). Comprehensive assessments have been conducted yearly, including MRI brain scans, reading achievement, cognition and others. Children were recruited from dozens of primary schools in Beijing. Informed consent was obtained from the parents or guardians (written) and children (oral). The exclusion criteria included a history of neurological or psychiatric disorders, the use of psychoactive drugs, significant head injury, and physical illness that prevented MRI scanning. All study procedures were reviewed and approved by the Institutional Review Boards at Beijing Normal University in accordance with the Declaration of Helsinki.

This study included 334 children with complete MRI scans and reading and cognition scores at both baseline and one-year follow-up assessments. More detailed information about the participants is presented in **Table 1**. Referring to a previous study (Siok et al., 2020), we categorized the participants as beginning readers (grade 3 and below, $n = 167$) and intermediate readers (grade 4 and above, $n = 167$).

Reading achievement test

Based on the national curriculum, the reading achievement test was developed by the project team of the National Children's Study of China (NCSC) (Dong and Lin, 2011). It assessed character and word recognition as well as sentence and short passage comprehension. Item response theory (IRT) scores, with an average of 500 and a standard deviation of 100, were computed based on the comprehensive national representative sample of 140,000 children and adolescents in more than 600 primary and junior high schools from 100 counties and 31 provinces around mainland China. According to the technical report of the NCSC (Dong and Lin, 2011) as well as previous research (Wang et al., 2022), this test showed good psychometric properties. The Cronbach's alpha coefficients were 0.72–0.94

TABLE 1 Characteristics of the participants at baseline and follow-up

	Baseline ($n = 334$)	**Follow-up ($n = 334$)**	***t* value**
Age (mean ± SD)	9.03 ± 1.33	10.20 ± 1.41	
Sex:Females, *n* (%)	157(47.0%)		
Parental Education (mean ± SD)	8.82 ± 2.55		
Family Income (mean ± SD)	8.80 ± 2.72		
Reading achievements (mean ± SD)	524.34 ± 98.27	548.45 ± 105.43	4.37***
Attention (mean ± SD)	93.59 ± 10.64	100.90 ± 11.38	13.87***
Memory (mean ± SD)	91.42 ± 11.07	97.71 ± 13.59	9.18***
Visuospatial Perceptive (mean ± SD)	97.97 ± 12.83	103.90 ± 13.06	9.52***
Reasoning (mean ± SD)	95.99 ± 11.99	100.09 ± 12.44	7.03***

Follow-up = the assessment after one year. The parent's education level refers to the highest level of education between children's parents. Parental Education: 1 = Uneducated; 2 = Primary education; 3 = Junior school; 4 = High school; 5 = Secondary vocational school; 6 = Polytechnic school; 7 = Higher vocational education; 8 = Junior college(part-time); 9 = Junior college(full-time); 10 = Bachelor degree (part-time); 11 = Bachelor degree (full-time); 12 = Graduate education or above. Family Income (RMB/year): 1 = Less than 3,000; 2 = 3,001–6,000; 3 = 6,001–10,000; 4 = 10,001–30,000; 5 = 30,001–50,000; 6 = 50,001–100,000; 7 = 100,001–150,000; 8 = 150,001–200,000; 9 = 200,001–400,000; 10 = 400,001–600,000; 11 = Over 600,000. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

at various grades, and the average difficulty coefficient was 0.69. Children completed the test in small groups on a computer within 45 min.

Cognitive abilities

The cognitive assessment battery developed by the NCSC project team (Dong and Lin, 2011) was used. There are four subtests, including attention, memory, visuospatial perception, and reasoning. This battery has been used in previous studies (Ren et al., 2013, 2015; Tao et al., 2015; Wang et al., 2016).

Attention

This subtest consists of four sets of number cancellations. In each set, participants were asked to cross out a number with specific marks from 200 items randomly arranged and presented within a 20 × 10 matrix that included 44 targets among the non-targets within 1 min. The correlation between this test and the Cancellation subtest of the Wechsler Intelligence Scale for Children (WISC-IV, Chinese version) (Zhang, 2009) was 0.72 ($p < 0.01$) among 114 children (Dong and Lin, 2011). The raw score was computed by subtracting the number of false hits from the total number of hits and transferred into a norm-based standardized score based on the national representative datasets. The internal consistency (Cronbach's α) was 0.94.

Memory

This subtest consists of 27 items, among which 12 are number recognition, and 15 are object pair recognition. Participants were asked to select numbers or matched objects presented previously from alternatives immediately and with a delay of 30 min, respectively. Among 110 children, the correlations between this memory test and the memory subtest of the WISC-IV (WISC-IV, Chinese version) (Zhang, 2009) were 0.53 ($p < 0.01$) and 0.46 ($p < 0.01$) for the number and object tests, respectively (Dong and Lin, 2011). The total number of correct responses was transferred into a norm-based standardized score based on the national representative datasets. The internal consistency (Cronbach's α) was 0.81 for the number recognition test and 0.74 for the object pair test.

Visuospatial perception

There were 27 items, among which 11 are hidden figures and 16 are mental rotation. In the hidden figure subtest, participants were asked to assess 4 options and identify the figure that was not in the complex figure presented previously. In the mental rotation subtest, participants were asked to identify the rotated figure among the 4 options that had been presented previously. Among 116 children, the correlation between the outcomes of the hidden figure subtest and the Test of Visual Perceptual Skills (TVPS-3) was 0.51 ($p < 0.01$), and the correlation between the outcomes of the mental rotation subtest and the Motor-free Visual Perceptual Test (MVPT-3) was 0.57 ($p < 0.01$) (Dong and Lin, 2011). The number of correct responses was transferred into a norm-based standardized score based on the national representative datasets. The internal consistency (Cronbach's α) was 0.74 for the hidden figure subtest and 0.77 for the mental rotation subtest.

Reasoning

This subtest consists of 40 items of figures and numbers. Participants were asked to choose one of four alternatives to complete figure or number sequences according to the rules embedded in the presented figure or number sequences. Among 111 children, the correlations among the figural reasoning subtest, the numerical reasoning subtest and the matrix reasoning subtest of the WISC-IV (WISC-IV, Chinese version) (Zhang, 2009) were 0.66 ($p < 0.01$) and 0.64 ($p < 0.01$), respectively, (Dong and Lin, 2011). The number of correct responses was transferred into norm-based standardized scores based on the national representative datasets. The internal consistency (Cronbach's α) was 0.77 for the figural reasoning test and 0.86 for the numerical reasoning test.

Image acquisition

All MRI scans were acquired on two 3T Siemens Prisma scanners with a 64-channel head coil at Peking University and Beijing HuiLongGuan Hospital using the same imaging sequences. Blood oxygen level-dependent (BOLD) fMRI data were acquired using a whole-brain, single-shot, multislice, echo-planar imaging (EPI) sequence of 240 volumes with the following parameters: repetition time/echo time (TR/TE) = 2000/30 ms, flip angle = 90°, field of view (FOV) = 224 × 224 mm, matrix = 64 × 64, slice thickness = 3.5 mm, and slices = 33. The resulting nominal voxel size was 3.5 mm × 3.5 mm × 3.5 mm. A fixation cross was displayed as images were acquired. Subjects were instructed to remain awake, keep their eyes open, fixate on the displayed blank screen, and remain still. Prior to time-series acquisition, a 6-min magnetization-prepared, rapid acquisition gradient-echo T1-weighted (MPRAGE) image (TR = 2530 ms, TE 2.98 ms, FOV 256 mm × 224 mm, matrix, effective voxel resolution of 1 mm × 1 mm × 1 mm, slice thickness = 1 mm, and slices = 192) was acquired to aid spatial normalization to standard atlas space. Prior to scanning, to acclimate subjects (children) to the MRI environment, a mock scanning session was conducted for each individual using a decommissioned MRI scanner and head coil. Mock scanning was accompanied by acoustic recordings of the noise produced by gradient coils for each scanning pulse sequence. To further minimize motion, subjects' heads were stabilized in the head coil using one foam pad over each ear.

MRI quality control

All MRI scan quality control procedures are described below. (i) Individual images were subjected to a careful visual examination performed by an experienced radiologist to exclude incidental abnormalities, such as arachnoid cysts, neuroepithelial cysts and other intracranial space-occupying lesions. (ii) Careful visual inspections with a scan rating procedure were separately conducted by five experienced raters using a protocol similar to that used in the Human Connectome Project (Marcus et al., 2013). (iii) Images considered to have a better than fair quality by both raters were retained. We quantified the head motion during resting-state fMRI acquisition as framewise displacement (FD) (Power et al., 2012). The participants were also excluded if the mean FD exceeded 0.5 mm during resting-state scans (Xia et al., 2018). In this study, a total of 12 children have been excluded from the data analysis because of substandard quality control.

Image data analysis

Resting fMRI preprocessing was performed using DPARSF software[1] (Yan and Zang, 2010). Preprocessing included the following steps: (1) slice-timing correction; (2) head-motion correction; (3) spatial normalization (MNI); (4) whole-brain and white matter signals and 24 motion parameters being regressed out; (5) spatial smoothing with a 6-mm 3D full-width half-maximum kernel; and (6) temporal bandpass filtering (0.01–0.1 Hz).

Literature-based spherical seed regions with a radius of 4 mm were created using DPARSF (Yan and Zang, 2010) in the bilateral thalamus [Montreal Neurological Institute (MNI) coordinates, left: $x = 6$, $y = -18$, $z = -3$; right: $x = -6$, $y = -21$, $z = -3$] (Skeide et al., 2017) and the bilateral striatum (MNI coordinates, left: $x = -18$, $y = 18$, $z = 0$; right: $x = -10$, $y = 14$, $z = 8$) (Alcauter et al., 2017). Mean time series were extracted by averaging the time series of all voxels in the seed region, and the correlation coefficients between this time course and all other brain voxels were computed. The correlation maps were then z-normalized using Fisher's r-to-z transformation to approximate a normal distribution. In addition, the Automated Anatomical Labeling (AAL) atlas was used for anatomical labeling of the MRI peaks/clusters in this study.

Statistical analysis

Cross-lagged panel analyses were performed using AMOS 21.0 (IBM). All statistical analyses of MRI data were performed in DPABI software (see footnote 1) (Yan et al., 2016). Pearson's correlation was used to evaluate seed-based connectivity, with a significance threshold set at a voxel-size value of $p < 0.001$ and a familywise error-corrected cluster probability of $P < 0.05$. Brain-behavior correlations were performed using SPSS 21 (IBM) with a significance threshold set 0.0125 (Bonferroni correction: alpha/number of tests = 0.05/4 = 0.0125). The directional association between functional connectivity and reading ability was determined by the cross-lagged panel model (CLPM) (Hamaker et al., 2015). Age, sex, handedness, site, household income, parental education, and head motion were controlled as covariates.

Results

Sample characteristics

All sample characteristics are presented in **Table 1**. Both reading ($t = 4.37$, $p < 0.001$) and cognitive abilities (attention, memory visuospatial perception and reasoning, all $ts \geq 7.03$, all $ps < 0.001$) performance increased at the follow-up assessment compared with that at the baseline assessment (**Table 1**).

The two cortico-subcortical reading pathways were identified: Thalamo-occipital and fronto-striatal circuits

To explore the thalamus-occipital and fronto-striatal circuits in reading development, we performed seed-based functional connectivity analysis. First, the mean time series were extracted by averaging the time series of all voxels in the seed region, including the bilateral thalamus and striatum, and the correlation coefficients between this time course and all other brain voxels were computed. Then, brain-behavior correlations were performed between the z *value* maps and reading ability. Striatum-based analysis showed that reading ability was related to the functional connectivity between the left striatum and left middle frontal gyrus (MNI coordinates: $x = -30$, $y = 51$, $z = 18$; **Figure 1A** and **Table 2**); thalamus-based analysis revealed that reading was associated with the functional connectivity between the right thalamus and left superior occipital gyrus (MNI coordinates: $x = 21$, $y = -66$, $z = 42$; **Figure 1B** and **Table 2**). The right striatum and left thalamus-based

---

[1] http://rfmri.org/

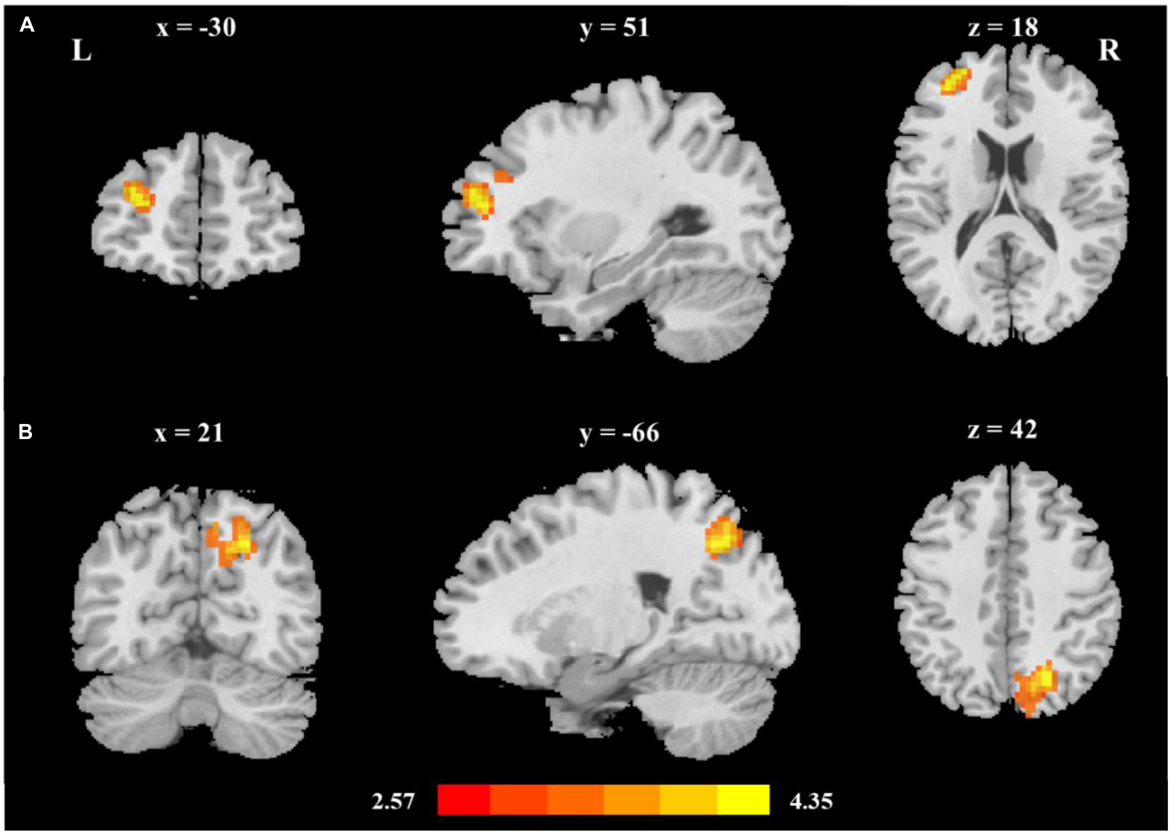


FIGURE 1
The two cortico-subcortical reading pathways **(A)** Reading ability was related to fronto-striatal functional connectivity (left middle frontal gyrus, MNI coordinates: x = −30, y = 51, z = 18); **(B)** Reading ability was related to thalamo-occipital functional connectivity (right superior occipital gyrus, MNI coordinates: x = 21, y = −66, z = 42). Age, sex, handedness, site, household income, parental education, and head motion were used as covariates of no interest.

TABLE 2 Seed-based functional connectivity related to reading.

Seed	Regions	HS	MNI Coordinate			Voxel	*z*-value
			x	y	Z		
Left striatum	MFG/SFG	L	−30	51	18	78	4.35
Right thalamus	PCUN/SOG	R	21	−66	42	244	4.18

MFG, middle frontal gyrus; SFG, superior frontal gyrus; PCUN, precuneus; SOG, superior occipital gyrus; L, left; R, right; HS, hemisphere.

analyses did not reveal significant results after correction for multiple comparisons.

Longitudinal prediction between reading and the two cortico-subcortical pathways: CLPM analysis

Children's reading performance and two cortico-subcortical pathways (striatum-MFG and thalamus-SOG) at baseline and follow-up were significantly correlated (Table 3). CLPM analysis showed that reading and functional connectivity between the left striatum and left middle frontal gyrus could bidirectionally predict each other's development one year later (Figure 2A), while reading could only unidirectionally predict the development of functional connectivity between the right thalamus and right superior occipital gyrus one year later (Figure 2B).

To examine the development of and change in the relationship, we categorized readers into primary reading and intermediate reading groups. To further examine the development and change in the relationship between

TABLE 3 The correlation matrix between reading and cortico-subcortical functional connectivity at baseline and follow-up assessments

		Striatum-MFG		Thalamus-SOG	
		BL	**FU**	**BL**	**FU**
Reading	BL	0.231***	0.210***	0.212**	0.191*
	FU	0.188***	0.198***	0.117*	0.161**

BL, baseline; FU, follow-up; MFG, middle frontal gyrus; SOG,superior occipital gyrus. Age, sex, handedness, site, household income, parental education, and head motion were used as covariates of no interest. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

cortical-subcortical functional connectivity and reading, referring to a previous study (Siok et al., 2020), we categorized the participants as beginning readers (grade 3 and below, $n = 167$) and intermediate readers (grade 4 and above, $n = 167$) (**Figures 3**, **4**). First, we found that reading could predict the connectivity between the left striatum and left middle frontal gyrus both in beginning readers and intermediate readers, and the effect did not show significant differences between the two groups ($\chi^2 = 0.375$, $p = 0.540$), while the reverse prediction was only significant in intermediate readers (**Figure 3**). Second, we found that reading could predict the connectivity between the right thalamus and right superior occipital gyrus both in beginning readers and intermediate readers, and the effect was weaker in intermediate readers ($\chi^2 = 5.447$, $p = 0.020$), while the reverse prediction was not significant in either group (**Figure 4**).

The cognitive basis of the thalamo-occipital and fronto-striatal circuits

To explore the cognitive basis of cortical-subcortical functional connectivity, we performed a partial correlation analysis controlling for age, sex, handedness, site, household income, parental education, and head motion. This analysis showed that the striatal frontal pathway was associated with attention, reasoning, and memory but not visuospatial perceptive ability (**Figure 5** upper layer graph) and that the thalamic occipital pathway was related to visuospatial perceptive and reasoning but not attention and memory ability (**Figure 5** lower layer graph).

Because attention, reasoning and memory tests contain visuospatial perceptive processing, to exclude its influence on the other three cognitive abilities, we controlled for visuospatial perceptive processing with other covariates and found that the correlations between the striatal frontal pathway and attention ($r = 0.162$, $p = 0.003$), between the striatal frontal pathway and reasoning ($r = 0.163$, $p = 0.003$), and between the striatal frontal pathway and memory ($r = 0.149$, $p = 0.007$) were still significant. In addition, the correlation between the thalamic occipital pathway and reasoning was no longer significant ($r = 0.095$, $p = 0.086$). This finding suggests that the relationship between the thalamic occipital pathway and reasoning observed in the reasoning test may be caused by visuospatial processing rather than reasoning processing.

Discussion

This longitudinal study of reading development in school-age children identified two important cortico-subcortical pathways—thalamo-occipital and fronto-striatal circuits—and provided novel evidence for understanding the developmental connections between reading and cortico-subcortical crosstalk. We demonstrated that the thalamo-occipital and fronto-striatal circuits are two important pathways of learning to read. The former is a visual pathway that can be predicted by reading ability, and the prediction gradually weakens with the improvement in reading ability, while the latter is a

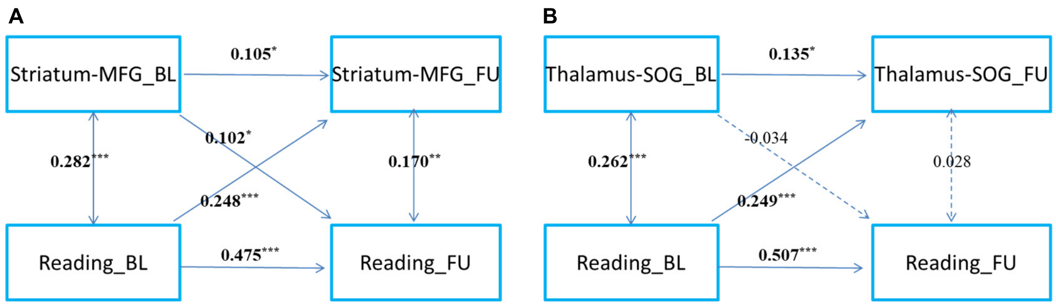


FIGURE 2
The cross-time predictions between cortical-subcortical functional connectivity and reading development in school-age children The cross-lagged panel models of **(A)** left striatum and left middle frontal gyrus connectivity, **(B)** right thalamus and superior occipital gyrus and reading development. MFG, middle frontal gyrus; SOG, superior occipital gyrus. Standardized estimates are presented. Age, sex, handedness, site, household income, parental education, and head motion were used as covariates of no interest. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

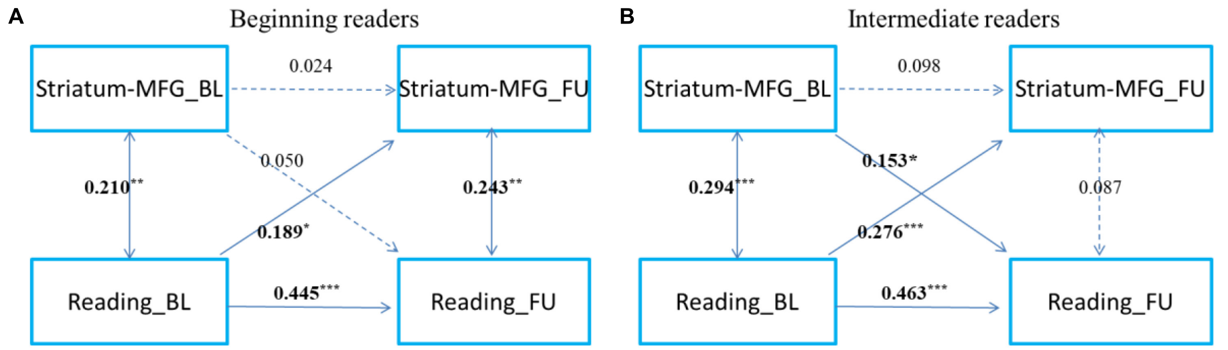


FIGURE 3
The cross-time predictions between left striatum and left middle frontal gyrus functional connectivity and reading development in school-age children The cross-lagged panel models in **(A)** beginning readers (grade 3 and below), and **(B)** intermediate readers (grade 4 and above). Standardized estimates are presented. MFG, middle frontal gyrus. Age, sex, handedness, site, household income, parental education, and head motion were used as covariates of no interest. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

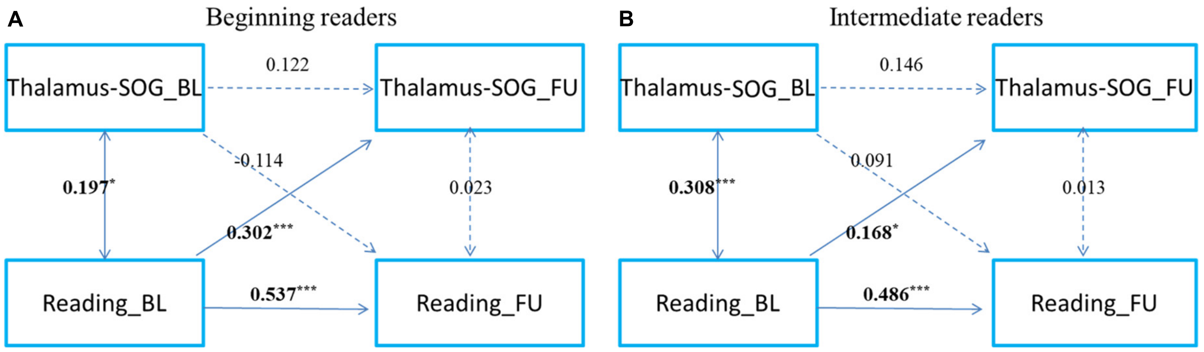


FIGURE 4
The cross-time predictions between right thalamus and right superior occipital gyrus functional connectivity and reading development in school-age children The cross-lagged panel models in **(A)** beginning readers (grade 3 and below), and **(B)** intermediate readers (grade 4 and above). Standardized estimates are presented. SOG, superior occipital gyrus. Age, sex, handedness, site, household income, parental education, and head motion were used as covariates of no interest. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

complex cognitive pathway related to attention, memory, and reasoning. And it can predict each other with reading, and its predictive effect on reading increases with the improvement of reading ability.

The fronto-striatal circuit: A cognitive pathway

The first pathway we found was the fronto-striatal cognitive pathway, which is formed by the functional connectivity between the left striatum and left middle frontal gyrus. This pathway has been thought to play an important role in reading, and a meta-analysis found that fronto-striatal abnormalities in reading disorders could arise from fundamental impairments in reading-related processes, such as phonological processing and implicit sequence learning, relevant to early language acquisition (Hancock et al., 2017). Consistent with our study, one cross-sectional, small sample study ($n = 60$) found that resting-state functional connectivity of the striatum and prefrontal cortex predicts reading performance in children aged 6–9 years old (Alcauter et al., 2017). However, in Alcauter et al.'s (2017) study, the prediction was not directional, simply providing a correlation based on cross-sectional data. In this study, using longitudinal cross-lagged panel analyses, we found that fronto-striatal circuitry and reading predicted each other in children aged 6–12 years. In addition, we found that fronto-striatal functional connectivity was predicted by learning to read at an early age and gradually became one of the important predictors of reading ability as one's reading ability developed. We confirmed and advanced the conclusions of Alcauter et al. (2017). We found that the predictive relationship demonstrated in Alcauter et al.'s (2017) study was for the prediction of reading based on the fronto-striatal pathway, and the predictive effect of the fronto-striatal pathway on reading was not significant until after grade 3.

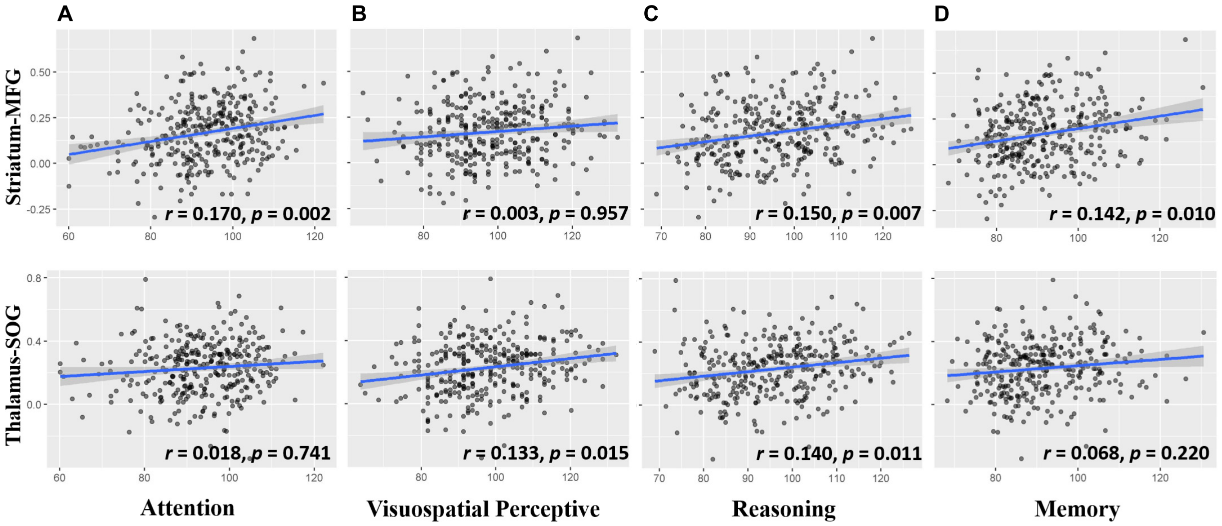


FIGURE 5
Scatter plots for the correlation between cortico-subcortical functional connectivity and cognitive ability **(A)** Attention ability related to striatum-MFG, but not thalamus-SOG; **(B)** visuospatial perceptive ability related to thalamus-SOG, but not striatum-MFG; **(C)** reasoning ability was associated with striatum-MFG and thalamus-SOG; **(D)** memory ability was associated with striatum-MFG, but not thalamus-SOG. The upper row is the striatum-MFG and the lower row is the thalamus-SOG. Striatum-MFG, functional connectivity between left striatum and left middle frontal gyrus; Thalamus-SOG, functional connectivity between right thalamus and right superior occipital gyrus. Age, sex, handedness, site, household income, parental education, and head motion were used as covariates of no interest.

The left middle frontal gyrus was recognized as critical for Chinese reading and reading acquisition. Learning to read could increase the activation in the left middle frontal gyrus (Li et al., 2006; Siok et al., 2020), and activation and gray matter volume in this region decreased in dyslexic Chinese children (Siok et al., 2004, 2008, 2009, 2020; Tan et al., 2005; Xu et al., 2015). In this study, the fronto-striatal circuit was composed of the connection between the left middle frontal gyrus and the striatum. Is the fronto-striatal functional connectivity found in this study a pathway specific to Chinese reading? To answer this question, we need to address three related issues. First, the middle frontal gyrus found in this study is not the middle frontal gyrus specific for Chinese reading; the Chinese reading-specific brain region found by Tan et al. (2005) is located in BA9, while this study found that the region is located in BA46. Second, does this pathway only exist in Chinese reading and not in alphabetic languages? In fact, the answer is no. Several studies based on alphabetic language reading have found that this pathway plays an important role (Alcauter et al., 2017; Hancock et al., 2017; Mohammadi et al., 2020). Third, is fronto-striatal connectivity stronger in Chinese reading than in alphabetic language reading? This study was unable to answer this question due to the lack of data on reading alphabetic languages. This question may need to be answered in future bilingual studies. In conclusion, recent evidence does not support fronto-striatal connectivity as a pathway specific to Chinese reading, and further research is needed to explore whether this pathway is more important in Chinese reading than in alphabetic languages.

The thalamo-occipital circuit: A visual pathway

The second pathway we found was the thalamo-occipital visual pathway formed by the functional connectivity between the right thalamus and right superior occipital gyrus. This was consistent with a previous study that trained illiterate individuals to be literate and found that training improved the degree of the thalamic activity and the strength of its connection to the occipital lobe (Skeide et al., 2017). In addition, we further investigated the predictive relationship between reading ability and the thalamo-occipital pathway and found that the predictive effect of reading on the thalamo-occipital pathway was unidirectional in both beginning and intermediate readers, and the predictive effect decreased gradually with reading ability. This finding suggests that the thalamo-occipital pathway may be closely related to visual processing in early reading. Effective temporal and spatial interpretation of text by the visual system is a critical early stage of reading, and any lesions that impair this process can lead to dyslexia, including downstream effects on the phonological domain (Vidyasagar, 2019). In this study, we only found that reading promoted the thalamo-occipital pathway but did not find the thalamo-occipital pathway to support or restrict the development of reading ability. This may be mainly because our samples were all developing children whose visual system and reading ability were well developed; therefore, there was no significant predictive effect on reading.

In this study, we found two cortico-subcortical pathways, thalamo-occipital and fronto-striatal circuits; however, in the study of literacy training, only the thalamus and thalamo-occipital pathways were observed (Skeide et al., 2017). Why did the authors not find the fronto-striatal pathway? We think there might be two reasons. First, the thalamus and striatum play different roles in the development of reading ability. Previous meta-analyses have found that the thalamus is more involved in childhood reading processing, while the striatum is more involved in adult reading processing (Koyama et al., 2011; Hancock et al., 2017). The study of illiteracy training belongs to the early stage of reading development, so the thalamus is the most prominent and easy to discover. For example, another study of training based on functional illiteracy found that although the connectivity strength of the fronto-striatal and thalamo-visual pathways was significantly lower in the illiterate group than in the control group, only the thalamic network, not the striatum network, was enhanced after a short training session (Mohammadi et al., 2020). This finding suggests that the thalamic network is more likely to be increased early in training, while the striatum network requires more training time to become more involved in reading ability. This suggestion has also been confirmed in this study. Second, the selection of measuring characteristics may also be an important reason. The thalamus is an important relay station for sensory and perceptual processing, and nerve fibers project out of the thalamus to the cerebral cortex in all directions, allowing hub-like exchanges of information. Skeide et al. (2017) used a measure called degree to examine brain changes in illiterate individuals before and after training. Degree refers to the strength of connection between a brain region and other brain regions of the whole brain and is an important feature to measure hub attributes. As a result, variations in the thalamus are easier to identify. In this study, we used the striatum as the seed point for functional connectivity analysis and found variations in the fronto-striatal circuit associated with reading development. Similar findings were made by Alcauter et al. (2017) using a combination of independent component analysis and functional connectivity.

Limitations

Several limitations of this study should be noted and need further research. First, in this study, we used a CLPM to examine the longitudinal relationship between attention and reading development after one year, which offered important empirical evidence for our understanding of the connection between attention and reading development. Future studies may further address this important question over a longer period. Second, the thalamus and striatum are relatively large subcortical structures and contain more subareas than other regions. In this study, we selected highly representative coordinates to characterize the thalamus and striatum based on previous studies and found thalamo-occipital and fronto-striatal circuits entailed two pathways for reading development. Future studies can fully explore all subareas of the thalamus and striatum to discover other possible pathways. Third, we found that reading could longitudinally predict the thalamus-occipital functional connectivity, but the thalamus-occipital pathways could not support or restrict reading development, which may be related to our sample being composed of normal developing children. Future research can adopt children with visual or reading difficulty to further verify whether the thalamo-occipital pathway could predict the longitudinal development of reading. Finally, the thalamo-occipital and fronto-striatal circuits are shared by reading and domain-general cognitive skills, but are not specific to reading development. Future studies should use caution when citing this conclusion.

Conclusion

This study clarifies the vague description of the cortico-subcortical crosstalk that learning to read promotes and clearly describes two important pathways: one is the thalamo-occipital visual pathway centered in the thalamus, and the other is the frontal lobe-striatum cognitive pathway centered in the striatum. The former plays an important role in the early stage of learning to read and gradually decreases as reading ability improves, while the latter plays an important role in learning to read and gradually increases as reading ability improves.

Data availability statement

The original contributions presented in this study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by Institutional Review Boards at Beijing Normal University in accordance with the Declaration of Helsinki.

Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

ST, SPT, SQ, YH, and QD conceived and designed the study. YW, JL, LM, RC, and JW collected the data under the supervision of ST, SPT, WM, and J-HG. YW, JL, and LM performed data analysis under the supervision of CC and ST. YW, JL, LM, and ST wrote the manuscript. ST, YW, JL, LM, RC, and JW amend and proofread the draft of the manuscript. All authors reviewed and commented on the study and manuscript.

Funding

This study was supported by the Beijing Brain Initiative of Beijing Municipal Science & Technology Commission (Z181100001518003), the 111 Project (BP0719032), and the National Natural Science Foundation of China (31521063).

Acknowledgments

We thank the National Center for Protein Sciences at Peking University in Beijing, China, for assistance with MRI data acquisition.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Alcauter, S., Garcia-Mondragon, L., Gracia-Tabuenca, Z., Moreno, M. B., Ortiz, J. J., and Barrios, F. A. (2017). Resting state functional connectivity of the anterior striatum and prefrontal cortex predicts reading performance in school-age children. *Brain Lang.* 174, 94–102. doi: 10.1016/j.bandl.2017.07.007

Banaie Boroujeni, K. (2021). *Mechanisms of cognitive flexibility in primate fronto-striatal circuits*. Nashville, TN: Vanderbilt University ProQuest Dissertations Publishing.

Binder, J. R., Medler, D. A., Westbury, C. F., Liebenthal, E., and Buchanan, L. (2006). Tuning of the human left fusiform gyrus to sublexical orthographic structure. *Neuroimage* 33, 739–748. doi: 10.1016/j.neuroimage.2006.06.053

Bitan, T., Cheon, J., Lu, D., Burman, D. D., Gitelman, D. R., Mesulam, M. M., et al. (2007). Developmental changes in activation and effective connectivity in phonological processing. *Neuroimage* 38, 564–575. doi: 10.1016/j.neuroimage.2007.07.048

Bolger, D. J., Perfetti, C. A., and Schneider, W. (2005). Cross-cultural effect on the brain revisited: Universal structures plus writing system variation. *Hum. Brain Mapp.* 25, 92–104. doi: 10.1002/hbm.20124

Brem, S., Halder, P., Bucher, K., Summers, P., Martin, E., and Brandeis, D. (2009). Tuning of the visual word processing system: Distinct developmental ERP and fMRI effects. *Hum. Brain Mapp.* 30, 1833–1844. doi: 10.1002/hbm.20751

Cupertino, R. B., Soheili-Nezhad, S., Grevet, E. H., Bandeira, C. E., Picon, F. A., Tavares, M. E. A., et al. (2020). Reduced fronto-striatal volume in attention-deficit/hyperactivity disorder in two cohorts across the lifespan. *Neuroimage Clin.* 28:102403. doi: 10.1016/j.nicl.2020.102403

Dong, Q., and Lin, C. D. (2011). *Standardized tests in children and adolescent mental development in China*. Beijing: Science Press.

Galandra, C., Basso, G., Manera, M., Crespi, C., Giorgi, I., Vittadini, G., et al. (2019). Abnormal fronto-striatal intrinsic connectivity reflects executive dysfunction in alcohol use disorders. *Cortex* 115, 27–42. doi: 10.1016/j.cortex.2019.01.004

Gordon, E. M., Laumann, T. O., Marek, S., Newbold, D. J., Hampton, J. M., Seider, N. A., et al. (2021). Human fronto-striatal connectivity is organized into discrete functional subnetworks. *bioRxiv* [Preprint]. doi: 10.1101/2021.04.12.439415

Hamaker, E. L., Kuiper, R. M., and Grasman, R. P. (2015). A critique of the cross-lagged panel model. *Psychol. Methods* 20, 102–116. doi: 10.1037/a0038889

Hancock, R., Richlan, F., and Hoeft, F. (2017). Possible roles for fronto-striatal circuits in reading disorder. *Neurosci. Biobehav. Rev.* 72, 243–260. doi: 10.1016/j.neubiorev.2016.10.025

Hidalgo-Lopez, E., and Pletzer, B. (2021). Fronto-striatal changes along the menstrual cycle during working memory: Effect of sex hormones on activation and connectivity patterns. *Psychoneuroendocrinology* 125:105108. doi: 10.1016/j.psyneuen.2020.105108

Houde, O., Rossi, S., Lubin, A., and Joliot, M. (2010). Mapping numerical processing, reading, and executive functions in the developing brain: An fMRI meta-analysis of 52 studies including 842 children. *Dev. Sci.* 13, 876–885. doi: 10.1111/j.1467-7687.2009.00938.x

Koyama, M. S., Di Martino, A., Zuo, X. N., Kelly, C., Mennes, M., Jutagir, D. R., et al. (2011). Resting-state functional connectivity indexes reading competence in children and adults. *J. Neurosci.* 31, 8617–8624. doi: 10.1523/JNEUROSCI.4865-10.2011

Koyama, M. S., Molfese, P. J., Milham, M. P., Mencl, W. E., and Pugh, K. R. (2020). Thalamus is a common locus of reading, arithmetic, and IQ: Analysis of local intrinsic functional properties. *Brain Lang.* 209:104835. doi: 10.1016/J.Bl.2020.104835

Li, G., Cheung, R. T., Gao, J. H., Lee, T. M., Tan, L. H., Fox, P. T., et al. (2006). Cognitive processing in Chinese literate and illiterate subjects: An fMRI study. *Hum. Brain Mapp.* 27, 144–152. doi: 10.1002/hbm.20173

Li, Y., and Bi, H. Y. (2022). Comparative research on neural dysfunction in children with dyslexia under different writing systems: A meta-analysis study. *Neurosci. Biobehav. Rev.* 137:104650. doi: 10.1016/j.neubiorev.2022.104650

Mamiya, P. C., Richards, T. L., Edden, R. A. E., Lee, A. K. C., Stein, M. A., and Kuhl, P. K. (2021). Task-related excitatory/inhibitory ratios in the fronto-striatal circuitry predict attention control deficits in attention-deficit/hyperactivity disorder. *medRxiv* [Preprint]. doi: 10.1101/2021.03.25.21254355

Marcus, D. S., Harms, M. P., Snyder, A. Z., Jenkinson, M., Wilson, J. A., Glasser, M. F., et al. (2013). Human connectome project informatics: Quality control, database services, and data visualization. *Neuroimage* 80, 202–219. doi: 10.1016/j.neuroimage.2013.05.077

Martin, A., Schurz, M., Kronbichler, M., and Richlan, F. (2015). Reading in the Brain of children and adults: A meta-analysis of 40 functional magnetic resonance imaging studies. *Hum. Brain Mapp.* 36, 1963–1981. doi: 10.1002/hbm.22749

Mohammadi, B., Munte, T. F., Cole, D. M., Sami, A., Boltzmann, M., and Russeler, J. (2020). Changed functional connectivity at rest in functional illiterates after extensive literacy training. *Neurol. Res. Pract.* 2:12. doi: 10.1186/s42466-020-00058-0

Muller-Axt, C., Anwander, A., and von Kriegstein, K. (2017). Altered structural connectivity of the left visual thalamus in developmental dyslexia. *Curr. Biol.* 27, 3692–3698.e4. doi: 10.1016/j.cub.2017.10.034

Ojha, A., Parr, A. C., Foran, W., Calabro, F. J., and Luna, B. (2022). Puberty contributes to adolescent development of fronto-striatal functional connectivity supporting inhibitory control. *bioRxiv* [Preprint]. doi: 10.1101/2022.05.02.490303

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., and Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59, 2142–2154. doi: 10.1016/j.neuroimage.2011.10.018

Ren, X. Z., Schweizer, K., and Xu, F. (2013). The sources of the relationship between sustained attention and reasoning. *Intelligence* 41, 51–58. doi: 10.1016/j.intell.2012.10.006

Ren, X. Z., Schweizer, K., Wang, T. F., and Xu, F. (2015). The prediction of students' academic performance with fluid intelligence in giving special consideration to the contribution of learning. *Adv. Cogn. Psychol.* 11, 97–105. doi: 10.5709/acp-0175-z

Richlan, F., Kronbichler, M., and Wimmer, H. (2011). Meta-analyzing brain dysfunctions in dyslexic children and adults. *Neuroimage* 56, 1735–1742. doi: 10.1016/j.neuroimage.2011.02.040

Rima, S., and Schmid, M. C. (2020). V1-bypassing thalamo-cortical visual circuits in blindsight and developmental dyslexia. *Curr. Opin. Physiol.* 16, 14–20. doi: 10.1016/j.cophys.2020.05.001

Rodrigue, K. M., Daugherty, A. M., Foster, C. M., and Kennedy, K. M. (2020). Striatal iron content is linked to reduced fronto-striatal brain function under working memory load. *Neuroimage* 210:116544. doi: 10.1016/j.neuroimage.2020.116544

Siok, W. T., Jia, F., Liu, C. Y., Perfetti, C. A., and Tan, L. H. (2020). A lifespan fMRI study of neurodevelopment associated with reading Chinese. *Cereb. Cortex* 30, 4140–4157. doi: 10.1093/cercor/bhaa038

Siok, W. T., Niu, Z., Jin, Z., Perfetti, C. A., and Tan, L. H. (2008). A structural-functional basis for dyslexia in the cortex of Chinese readers. *Proc. Natl. Acad. Sci. U.S.A.* 105, 5561–5566. doi: 10.1073/pnas.0801750105

Siok, W. T., Perfetti, C. A., Jin, Z., and Tan, L. H. (2004). Biological abnormality of impaired reading is constrained by culture. *Nature* 431, 71–76. doi: 10.1038/nature02865

Siok, W. T., Spinks, J. A., Jin, Z., and Tan, L. H. (2009). Developmental dyslexia is characterized by the co-existence of visuospatial and phonological disorders in Chinese children. *Curr. Biol.* 19, R890–R892. doi: 10.1016/j.cub.2009.08.014

Skeide, M. A., Kumar, U., Mishra, R. K., Tripathi, V. N., Guleria, A., Singh, J. P., et al. (2017). Learning to read alters cortico-subcortical cross-talk in the visual system of illiterates. *Sci. Adv.* 3:e1602612. doi: 10.1126/sciadv.1602612

Tan, L. H., Laird, A. R., Li, K., and Fox, P. T. (2005). Neuroanatomical correlates of phonological processing of Chinese characters and alphabetic words: A meta-analysis. *Hum. Brain Mapp.* 25, 83–91. doi: 10.1002/hbm.20134

Tao, S. (2019). Intelligence development and school adjustment of school-age children and adolescents: A follow-up cohort study. *Psychol. Commun.* 2, 88–90.

Tao, S., Liu, H., Zhou, C., Wang, C., Sun, C., Xu, F., et al. (2015). The roles of school psychological environment in grades 4∼ 6 students' cognitive development: A multilevel analysis of the national representative data. *J. Psychol. Sci.* 1, 2–10.

Tschentscher, N., Ruisinger, A., Blank, H., Diaz, B., and von Kriegstein, K. (2019). Reduced structural connectivity between left auditory thalamus and the motion-sensitive planum temporale in developmental dyslexia. *J. Neurosci.* 39, 1720–1732. doi: 10.1523/Jneurosci.1435-18.2018

Vidyasagar, T. R. (2019). Visual attention and neural oscillations in reading and dyslexia: Are they possible targets for remediation? *Neuropsychologia* 130, 59–65. doi: 10.1016/j.neuropsychologia.2019.02.009

Wang, T. F., Ren, X. Z., Schweizer, K., and Xu, F. (2016). Schooling effects on intelligence development: Evidence based on national samples from urban and rural China. *Educ. Psychol.* 36, 831–844. doi: 10.1080/01443410.2015.1099618

Wang, Y., Guan, H., Ma, L., Luo, J., Chu, C., Hu, M., et al. (2022). Learning to read may help promote attention by increasing the volume of the left middle frontal gyrus and enhancing its connectivity to the ventral attention network. *Cereb. Cortex* bhac206. doi: 10.1093/cercor/bhac206 [Epub ahead of print].

Xia, C. H., Ma, Z., Ciric, R., Gu, S., Betzel, R. F., Kaczkurkin, A. N., et al. (2018). Linked dimensions of psychopathology and connectivity in functional brain networks. *Nat. Commun.* 9:3003. doi: 10.1038/s41467-018-05317-y

Xu, J., Kemeny, S., Park, G., Frattali, C., and Braun, A. (2005). Language in context: Emergent features of word, sentence, and narrative comprehension. *Neuroimage* 25, 1002–1015. doi: 10.1016/j.neuroimage.2004.12.013

Xu, M., Wang, T., Chen, S., Fox, P. T., and Tan, L. H. (2015). Effective connectivity of brain regions related to visual word recognition: An fMRI study of Chinese reading. *Hum. Brain Mapp.* 36, 2580–2591. doi: 10.1002/hbm.22792

Yan, C. G., Wang, X. D., Zuo, X. N., and Zang, Y. F. (2016). DPABI: Data processing & analysis for (Resting-State) brain imaging. *Neuroinformatics* 14, 339–351. doi: 10.1007/s12021-016-9299-4

Yan, C., and Zang, Y. (2010). DPARSF: A MATLAB toolbox for "Pipeline". Data Analysis of Resting-State fMRI. *Front. Syst. Neurosci.* 4:13. doi: 10.3389/fnsys.2010.00013

Zhang, H. (2009). The revision of WISC-IV Chinese version. *J. Psychol. Sci.* 32, 1177–1179. doi: 10.16719/j.cnki.1671-6981.2009.05.026

frontiers | Frontiers in Neuroscience

TYPE Original Research
PUBLISHED 08 September 2022
DOI 10.3389/fnins.2022.934590

Modeling developmental changes in print tuning in a transparent alphabetic orthography

Ludo Verhoeven[1,2]*, Marinus Voeten[1] and Jos Keuning[3]

[1]Behavioural Science Institute, Radboud University, Nijmegen, Netherlands, [2]Faculty of Arts, University of Curaçao Moises Da Costa Gomez, Willemstad, Curaçao, [3]Cito, Arnhem, Netherlands

Check for updates

OPEN ACCESS

EDITED BY
Li-Hai Tan,
Shenzhen Institute of Neuroscience,
China

REVIEWED BY
Léon Franzen,
University of Lübeck, Germany
Daniela Traficante,
Catholic University of the Sacred
Heart, Italy

*CORRESPONDENCE
Ludo Verhoeven
Ludo.verhoeven@ru.nl

SPECIALTY SECTION
This article was submitted to
Neurodevelopment,
a section of the journal
Frontiers in Neuroscience

RECEIVED 02 May 2022
ACCEPTED 22 August 2022
PUBLISHED 08 September 2022

CITATION
Verhoeven L, Voeten M and Keuning J
(2022) Modeling developmental
changes in print tuning in a transparent
alphabetic orthography.
Front. Neurosci. 16:934590.
doi: 10.3389/fnins.2022.934590

The development of print tuning involves the increased specificity and redundancy for orthographic representations. However, it is by no means clear how decoding accuracy and efficiency are related over the years and how it affects reading disability. In the present study, we monitored the development of accuracy and efficiency of decoding in Dutch as a relatively transparent orthography as a function of orthographic complexity and lexical status throughout the primary grades. There was clear evidence that development of decoding accuracy preceded development of decoding efficiency and that a certain threshold of accuracy is needed for decoding efficiency to evolve. Furthermore, it was shown that pseudoword decoding efficiency predicted growth in word decoding efficiency, especially for the higher levels of orthographic complexity. There was also evidence that accuracy precedes efficiency across different profiles of readers and that decoding strength can be defined as a function of orthographic complexity and lexicality.

KEYWORDS

print tuning, decoding accuracy, decoding efficiency, development, orthographic complexity, lexicality effects

Introduction

Reading involves tuning to orthographic information to access phonological word patterns and corresponding meanings in the brain (e.g., Dehaene et al., 2005; Dehaene and Cohen, 2007). In learning to read, children must learn the inventory of graphic forms for a given writing system, the orthographic units (graphemes) that connect to spoken language, and how specific orthographic units map onto specific units of the spoken language. Beyond discovering the mapping principle of their writing system, children develop print tuning, that is they acquire precise connections of their orthographic inventory with language units, allowing them to compute orthographic representations and to gain orthographic fluency through reading (Maurer et al., 2011).

Across languages, word identification shifts from computation to memory-based retrieval for words when they become familiar. The importance of orthographic knowledge, sometimes neglected for alphabetic writing, spans from initial learning to later automatized word identification. Across languages, written words can become familiar perceptual objects that are then recognized more quickly. Learning to read fluently builds on this increasing familiarity. Reading fluency becomes the distinguishing marker of skill once children have reached high accuracy for word decoding. Highly fluent word reading is an effortless perceptual response that can include the automatization of word decoding, familiarity-based memory retrieval, and the attainment of fluent skilled reading. Accordingly, it is suggested that the left dorsal parietal-temporal is mainly involved in the establishment of accurate word reading and that it later supports the rapid word identification subserved by the occipitotemporal system which is associated with orthographic processing and coding (McCandliss and Noble, 2003). How this transition takes place over the grades is far from clear. It is unclear when children make the transition from accurate to fast word decoding, how word and pseudoword decoding development are related, to what extent decoding accuracy and efficiency problems in poor readers can be differentiated and what the role is of orthographic complexity and lexical status. Therefore, in the present study the development of tuning for print was examined in children learning to read Dutch as a transparent orthography throughout the primary grades.

Print tuning development

Several attempts have been made to model the processing of visual word forms. Dual-route theories of reading propose that both lexical and non-lexical routes can be followed to arrive at a phonological representation of a word. In the lexical route, it is assumed that access of the representation of the word is derived from the orthographic input lexicon with its spoken form being retrieved from the phonological output lexicon. In the non-lexical route, it is assumed that a set of grapheme-to-phoneme correspondence (GPC) rules is applied to the string of letters which are subsequently assembled to the phonological representation of the word (see Coltheart et al., 2001). In order to shed more light on the underlying perceptual and cognitive processes of decoding, a Bayesian approach of modeling has been followed (see Norris, 2013). Eye-movement studies (see Sereno and Rayner, 2003; Reichle et al., 2013) and event-related potential studies (see Balass et al., 2000; Lemons et al., 2010) have also provided insight into the temporal and spatial progression of oculomotor control during decoding. And functional magnetic resonance imaging data revealed distributed neural systems for mapping orthography directly to phonology, involving left supramarginal, posterior middle temporal, and fusiform gyri. Distinct from these were areas reflecting semantic processing, including left middle temporal gyrus/inferior-temporal sulcus, bilateral angular gyrus, and precuneus/posterior cingulate (Graves et al., 2010). It has also been shown that words can be read via two neural pathways working in close collaboration with each other: a dorsal pathway for phonological recognition and a ventral pathway for the retrieval of already existing orthographic representations from memory (Cohen and Dehaene, 2009; Das et al., 2011).

Importantly, decoding problems may arise under the condition of developmental dyslexia (see Verhoeven et al., 2019). Research has consistently evidenced that compared to typical readers, readers with dyslexia showed a phonological processing deficit. It is assumed that deficits in phonological representations negatively impact orthography-to-phonology mapping and ultimately limit adequate development of orthographic-to-phonology and orthography-to-semantic pathways (Richlan et al., 2013). Interestingly, time-sensitive event-related potentials studies revealed reduced inferior occipito-temporal N1 tuning for print in children with dyslexia being associated with a developmental delay in the fast retrieval of written words (Maurer et al., 2011). A functional imaging study on the neural correlates of reading fluency problems in dyslexia evidenced that readers with dyslexia exhibited lower gains in activation in the left prefrontal and left superior temporal regions associated with semantic retrieval and semantic and phonological representations (Christodoulou et al., 2014).

A developmental mechanism serving reading of both deeper and shallower orthographies is the self-teaching procedure identified by Share (2004). It is assumed that in learning to read, children start out learning the decoding principle along with the graphs of the orthography and that with growing reading experience, word meanings are identified more holistically and no longer via the application of grapheme-level conversion. Each encounter with a word is supposed to result into phonological recoding, which is then fed back to the orthographic representation of the word, triggering a word-specific identification process leading to a (partial) storage of the word in memory. Thanks to children's coarse and fine tuning for print, lexical representations become available for frequently occurring words to enable holistic processing (Eberhard-Moscicka et al., 2015). This mechanism makes it clear that a limited number of exposures to the same word can be sufficient for storage of its orthographic representation (see Ziegler et al., 2014). Besides the computation of accurate word representations, reading can also provide gains in word reading fluency. With repeated exposure, the status of words gradually changes from unfamiliar to familiar. Highly fluent word reading is an effortless perceptual response that can include the automatization of word decoding, familiarity-based memory retrieval, and the attainment of fluent skilled reading (Verhoeven and Perfetti, 2017, 2022).

Role of orthographic complexity

Developmental changes of print tuning are assumed to be highly related to orthographic complexity (Borleffs et al., 2017). A direct influence of word length and word frequency on children's word decoding has also been evidenced in a variety of orthographies, such as English (Rau et al., 2015), Italian (De Luca et al., 1999) and German (Tiffin-Richards and Schroeder, 2015). Research has shown that the word length effect applies particularly to beginning readers. The longer the length of a word, the more time it takes to read the word (Spinelli et al., 2005). However, this word length effect diminishes with decoding proficiency (Zoccolotti et al., 2005), which may be due to a shift from a serial letter-by-letter approach to a more holistic word processing approach (van den Boer and de Jong, 2015).

Becoming fluent in word reading is also dependent on orthographic depth (Ziegler and Goswami, 2005). Becoming fluent in word reading is relatively simple in shallow alphabetic writing. The first encounter with a new written word leads to decoding of the written form into its phonological form and initial familiarity; greater exposure may be needed for familiarity with deeper orthographies to grow. When a systematic comparison of the development of word decoding in different alphabetic languages was undertaken by Seymour et al. (2003), the speed and accuracy of the reading of familiar words by normal readers was found to be affected by syllabic complexity and orthographic depth. Syllabic complexity involved the distinction between open consonant-vowel (CV) syllables with few initial or final consonant clusters and closed consonant-vowel-consonant (CVC) syllables with complex consonant clusters in both the onset and coda positions. Orthographic depth involved the degree to which relevant orthographic patterns did not reflect and parallel phonemic representations. Decoding performance was relatively high in transparent orthographies, such as French, Portuguese, and Danish, and low in opaque English orthography. In a functional magnetic resonance imaging study, it was evidenced that the convergence of brain activity for print and speech was higher in learning to read transparent Polish as compared to opaque English in the right temporal region, associated with phonological processing, whereas it was lower in the left fusiform region, associated with visual word recognition (Chyl et al., 2021).

The lexicality effect

A developmental increase in specialization of the brain mechanisms engaged for word and pseudoword processing has been evidenced. Consistent with adult studies, children demonstrated a greater activation for words as compared to pseudowords in the anterior left ventral occipito-temporal cortex (Weiss and Booth, 2017). Abundant behavioral studies have shown that knowledge of word meanings is inextricably involved in word reading (see Taylor et al., 2015). In learning to read, children must develop orthographic representations of words from their oral vocabulary which is supported by an underlying self-teaching mechanism of pseudoword decoding. Given word exposure effects, it can be expected that the growth for word reading will be faster than for pseudoword reading. Pseudoword tasks are commonly used to assess children's decoding ability and to diagnose the non-word reading deficit, i.e., phonological processing deficit in dyslexia (De Luca et al., 2002). Lexical status may provide two indicators of accurate and efficient word reading ability in alphabetic orthographies. A lexicality effect shows the advantage of development of word decoding accuracy and efficiency beyond the decoding of pseudowords across varying orthographies (see Verhoeven et al., 2019).

A word reading advantage provides evidence of the fact that orthographic representations have become lexicalized (Perfetti, 2007). Its strength is modulated by several factors, such as age, reading ability and orthographic complexity (length and syllabic complexity). An orthographic complexity effect within orthographies was evidenced by Rahbari and Sénéchal (2010). They compared decoding efficiency of transparent and opaque words and found greater lexicality effects for words with transparent mappings and smaller lexicality effect for more opaque words of which orthographic representations are more slowly acquired. Interactions between lexicality effects and orthographic depth effects parallel to those for the reading of familiar words were also found for the reading of simple non-words in the before mentioned study by Seymour et al. (2003). Faster and more accurate decoding were apparent for the simple syllable languages of French, Portuguese, and Danish; for the more complex syllable languages of Swedish and Dutch, this occurred to a lesser extent. The most striking outcome was the evidence of profound delays in the development of simple non-word decoding skills in English (Seymour et al., 2003, p. 160). Furthermore, Caravolas (2018) tested growth models of word and pseudoword decoding efficiency in early readers of opaque English and transparent Czech and Slovak orthographies. Growth was faster for word than pseudoword decoding efficiency, and strong lexicality effects that increased over time were obtained across languages. In line with predictions about the costs of learning lower-consistency orthographies, readers of English experienced relatively slower growth on both reading skills.

Modeling longitudinal changes in print tuning

Neurocognitive studies have evidenced that tuning to both words and pseudowords can be considered fundamental to reading development in alphabetic orthographies. In adult skilled readers, the N1 component in the visual event-related potential appeared as index of visual expertise for print (Maurer et al., 2005a). MEG-data also showed that print-specific activation typically occur in the inferior parts of the occipito-temporal cortex (Maurer et al., 2005b) which is in agreement with fMRI evidence of the identification of the visual word form area as marker for sensitivity to orthographic word forms (Cohen et al., 2000) and the more extended visual word form system showing a posterior-to-anterior gradient of word form specificity particularly tuned for print (Dehaene et al., 2015).

Pseudoword decoding abilities address the degree to which children have acquired the basic self-teaching device of phonological recoding whereas word decoding abilities display the accuracy and efficiency of retrieving orthographic representations. Pseudoword decoding ability involves the accuracy and efficiency of phonological recoding. It arises in the early grades as children have learned all the letters along with the alphabetic principle that letters in a word can be associated with their corresponding sounds, which can successively be blended into word pronunciations. Word decoding efficiency is defined as the accurate and fluent reading of words and is usually assessed under time pressure, for instance as the number of words correctly read in 1 min. As children successfully apply the phonological recoding procedure to newly encountered words, they become capable to build word-specific orthographic representations in their mental lexicon. Repeated exposure to words leads to incrementally refined and redundant orthographic representations that facilitate word identification. Thus, word decoding drives robust orthographic representations and enables efficient word identification. Indeed, Verhoeven and Perfetti (2017) evidenced that the growth of word decoding across languages and writing systems is largely a matter of increased speed.

It is important to note, however, that little is known about concurrent growth trajectories of accuracy and efficiency of pseudoword decoding and word decoding skills in the early grades among children learning different alphabetic orthographies (Caravolas, 2018). Juul et al. (2014) examined the relation between accuracy and speed of word reading in first and second grade in Danish children. They found that speed of word recognition mainly developed after a student had reached an accuracy level of 70% correct. Word recognition speed was found to be dependent on the amount of time a student has been able to read with basic accuracy. Karageorgos et al. (2019) followed the accuracy and speed of word decoding in a representative sample of German primary school children from grades 1–4. They found the growth curves of word-recognition speed to be steeper for children who achieved a basic word-recognition accuracy of 71% compared with children who failed to reach this threshold by the end of Grade 1. Children who reached the basic word-recognition accuracy in later grades showed flatter trajectories of word-recognition speed over the primary school years. These findings suggest that good word-recognition accuracy lays the foundation for the development of word-recognition speed of primary school children. In a follow-up study, Karageorgos et al. (2020) investigated whether word-reading speed starts increasing only after German fourth graders have reached a basic level of word-decoding accuracy. The results based on the full sample suggest that a specific level of word-decoding accuracy seems to be required before word-reading speed starts improving. They also examined for children with lower reading abilities whether a word-decoding intervention has differential effects depending on the level of accuracy a child has reached before the intervention. The trained readers showed positive treatment effects on word-decoding accuracy for readers below the accuracy level and on word-decoding speed regardless of their accuracy. The results suggest that a sufficient level of word-decoding accuracy is an important precondition for the development of fluent reading. Longitudinal changes in reading network connectivity in children have also been studied by Wise Younger et al. (2017) across two moments in time. They evidenced longitudinal increases in word decoding to be related with higher initial connectivity in the dorsal stream between fusiform and inferior parietal cortex, implicating phonological recoding. Increases in word reading were also associated with maintenance of connectivity in the ventral stream between inferior occipital and fusiform cortex, implicating automatic orthographic decoding. It was also shown that readers with little efficiency improvement over time showed low levels of connectivity in the dorsal stream and a decrease in ventral activity over time.

Zhang and Peng (2021) conducted a meta-analysis of neuroimaging studies on the development of decoding in children with developmental dyslexia as compared to their typically reading peers. They found that readers with dyslexia showed hypoactivity in the left-lateralized reading network. It included the occipitotemporal regions, temporoparietal regions, and inferior frontal gyri in real word and pseudoword decoding. In pseudoword compared with real word decoding, hypoactivity was more severely reduced in the inferior frontal gyrus. Meta-regression showed no hypoactivity to be related with grade in real word decoding, whereas in pseudoword decoding, hypoactivity in the left superior temporal gyrus was found to be negatively associated with grade. These findings show that reading problems may be associated with abnormalities in both the direct and indirect pathways in word and pseudoword decoding. Compared with word decoding, pseudoword decoding in poor readers was found to be more associated with abnormalities in the indirect pathway. With development, abnormalities in both pathways appeared stable in word decoding, whereas in pseudoword decoding abnormalities in the indirect pathway were initially more severe but improved later, and abnormalities in the direct pathway tended to become more severe with age.

The research so far shows that tuning for print involves learning to decode pseudowords and words across an extended period for differential graphic forms. However, due to the lack of extended longitudinal data and a poor operationalization of decoding skills it is by no means clear how decoding accuracy and efficiency of words and pseudowords at different length develop, to what extent they are related over the years and how it affects reading disability. At least, three issues related to the development of print tuning remain unresolved. First of all, it is by no means clear how accuracy and efficiency in word reading development are associated throughout the primary grades. Previous studies focused on grades 1–2 (Juul et al., 2014) or grades 1–4 (Karageorgos et al., 2020) without considering orthographic complexity and lexicality. Second, it is still far from clear how the developmental trajectories of pseudoword and word reading across the grades are related. Previous research focused on the first two grades without taking into account lexical complexity (see Caravolas, 2018). Finally, it is unclear to what extent decoding accuracy and efficiency problems in poor readers can be differentiated. In the research so far, the focus has been on print tuning development without separating accuracy from efficiency in decoding words and pseudowords with varying orthographic complexity (see Maurer et al., 2011).

The present study

The aim of the present study was to examine the early stages of print tuning in Dutch as a relatively transparent orthography throughout the primary grades. To uncover the consequences of neural adaptation while familiarizing with the Dutch script, the longitudinal changes of accuracy and efficiency of decoding in Dutch were investigated in typical and poor readers as a function of lexical status (word vs. pseudoword) and orthographic complexity. The Dutch language offers an interesting case because Dutch orthography is largely phonemic, although the basic letter to phoneme correspondences in Dutch are not strictly one-to-one or invariant (see Verhoeven and van Leeuwe, 2009). In short Dutch words, a rather straightforward mapping of graphemes to phonemes applies, but Dutch syllable structure can be complex because multiple consonants can occur in both the onset and coda positions. In longer words, several deviations from a one-to-one correspondence between letters and sounds can occur. The basic task for children learning to read Dutch is thus to progress from the sequential grapheme-to-phoneme decoding of (pseudo)words to the fast, parallel, and largely phonology-based processing of different (pseudo)word classes. In the present study, we monitored the Dutch decoding development for words and pseudowords in four types of orthographic patterns that varied in a principled manner regarding orthographic transparency (Nunn, 1998; Verhoeven, 2017): (i) regular consonant-vowel-consonant patters, (ii) monosyllabic patterns with consonant clusters in prevocalic and postvocalic positions, (iii) bisyllabic patterns and (iv) polysyllabic patterns. In the present study, word and pseudoword decoding development was studied considering efficiency and accuracy with an accelerated longitudinal design covering Grade 1 to Grade 6 of elementary education. Therefore, students were instructed to read aloud words and pseudowords for each of the four orthographic patterns by the end of each grade.

An attempt was made to find an answer to the following research questions:

(1) How are accuracy and efficiency of decoding related over the grades? We approached this question by first estimating a growth model for decoding efficiency. Next, we added decoding accuracy as a dichotomous moderator to this model. This dichotomous moderator indicated whether a student had reached a threshold percentage of accuracy. These analyses were done separately for all pseudoword and word reading tests, to see how development differed as a function of orthographic complexity and lexical status.

(2) How are pseudoword decoding and word decoding efficiency related throughout the elementary school grades? And how do these developmental relations differ at different levels of orthographic complexity?

(3) To what extent can decoding accuracy and efficiency problems in poor readers be differentiated?

Growth modeling was applied to answer the first two research questions. Given the fact that in consistent orthographies like Dutch there is a high emphasis on phonological recoding, a relatively fast development of pseudoword reading and a small lexicality effect was expected. Furthermore, we hypothesized that decoding efficiency would be preceded by decoding accuracy as a function of orthographic complexity and lexical status. Stability of individual differences was also expected. For the third research question, we used latent class analysis to search for categories of poor readers differing in accuracy and efficiency of decoding and possibly differentiated by orthographic complexity and lexical status. We expected to find at least three subclasses of readers: inaccurate and inefficient, accurate and inefficient, and both accurate and efficient, and an interaction between orthographic complexity and lexical status and subclass of readers.

Materials and methods

Design and participants

An accelerated longitudinal design was used to study the development of decoding accuracy and efficiency across elementary education from Grade 1 to Grade 6. The data were obtained from a national test norming study in the Netherlands (Verhoeven et al., 2013). A stratified random sample of schools resulted in 70 schools participating, stratified by socioeconomic status of the school population (see Verhoeven and Keuning, 2018). Four cohorts of students were included, each one starting at a different grade level, respectively Grade 1, Grade 2, Grade 3, and Grade 4. Only the cohorts of students with measurements in three consecutive school years were included, see the scheme in **Table 1**. Each school contributed one or two cohorts of students. For the longitudinal analyses in the present study, we used the data from three measurement occasions in each of four cohorts of students ($N = 946$, 457 boys and 489 girls), see the numbers of students per cohort and per grade level in **Table 1**. The mean ages of these students were at the first measurement occasion 6.9 years for Grade 1 (Cohort 1), 7.8 years for Grade 2, 8.9 years for Grade 3, and 10.0 years for Grade 4.

For each cohort the data for three grade levels were missing by design. In addition, there were missing values because of longitudinal drop-out, varying per cohort from 4 to 29%. Occasionally, a few students missed a test occasion or one of the tests. Missingness seemed not related to word types, nor to lexical status. But Cohort 5 showed more missing values for pseudowords than for real words. Missingness seemed not related to level of accuracy or efficiency of word decoding. All available data for the four cohorts were kept in the analyses using full information maximum likelihood.

Measures

Word decoding accuracy and efficiency was assessed with four cards of the Dutch Decoding Test (Verhoeven and Keuning, 2018). Students were instructed to read aloud unrelated words from a card as quickly and accurately as possible for 1 min. The words for each test were printed in columns of 30 words. The efficiency score was determined as the number of words read correctly in 1 min. The accuracy score was determined as the percentage of words read correctly; this percentage was taken from the total number of words read by the student for 1 min. The four cards differed by the orthographic structure of the words on it. The first card was composed of 150 CVC (consonant – vowel – consonant) words. The words were regular Dutch words, thought to be familiar for most 6-year-old Dutch children. The second card also included 150 monosyllabic words but with added complexity because of consonant clusters in the onset and/or coda position of the word, CCVC, CVCC, CCVCC, CCCVC, or CVCCC. We denoted this card as CC. The third card included 120 bisyllabic words (Bisyl), and the fourth card was composed of 120 polysyllabic words (Poly, three or four syllables).

TABLE 1 Design of data collection.

Cohort	Grade						N
	1	2	3	4	5	6	
1	X	X	X				203
2		X	X	X			219
3			X	X	X		208
4				X	X	X	316
N	203	422	630	743	524	316	946

X, decoding tests administered at the end of the school year.

Pseudoword decoding accuracy and efficiency was in a similar way assessed with four cards containing lists of pseudowords. The pseudowords were words that do not exist in the Dutch language but that were constructed in a way that they obeyed the orthographic rules of Dutch and that they were therefore still pronounceable. The task again was to read aloud the words as quickly and accurately as possible, but now 2 min were given for each of the four cards. The four cards of pseudowords were composed of the same structures and the same numbers of items as the four cards of existing words. The scores were determined in the same way as for reading existing words. Pseudoword decoding efficiency was determined as the number of words read correctly in 2 min. For better comparability with the scores for real words, we divided by two to obtain an average efficiency score per minute. This transformation did not influence the analyses, only the presentation of results. The pseudoword decoding accuracy score was determined as the percentage of words read correctly; this percentage was taken from the total number of pseudowords read by the student for 2 min.

Procedures

All tests were administered individually, in a quiet place outside the classroom. Students were tested during school hours. Test administration was performed by well-trained graduate students. The tests were administered as part of a larger collection of reading and language tests (Verhoeven and Keuning, 2018). The four Dutch word reading tests were presented successively in a randomized order for each student. The four pseudoword reading tests were presented in the same way. For each cohort, testing was done toward the end of the school year, in June, during three consecutive school years, starting in 2003–2004.

To get information about the reliability of the decoding tests, test–retest correlations were computed for each of the test cards at each of the six grade levels. Between the two test occasions an interval of 3–4 months was maintained. The test–retest correlations, which give a lower bound for reliability, varied by test card and grade level between 0.83 and 0.93 for the decoding efficiency scores with existing words, for grades 2 to 6. For decoding efficiency of pseudowords the test-retest correlations varied between 0.76 and 0.92, with again most of them above 0.82. In Grade 1 the correlations were somewhat lower, between 0.76 and 0.84 for Dutch words, and between 0.72

and 0.77 for pseudowords. This makes sense, since in first grade stronger development takes place than in later grades.

Analysis

Growth modeling was used to answer the first research question, for each of the eight decoding tests separately. In each of the four cohorts, data were available at the end of three consecutive grade levels. The decoding efficiency data were analyzed in long file format with multilevel analysis, using MLwiN 2.36 (Rasbash et al., 2016).

The second research question about developmental relations between word and pseudoword decoding efficiency was studied using bivariate change score analysis (McArdle, 2009). Does pseudoword decoding efficiency predict word decoding efficiency? Or is it the other way around? And at what grade levels in elementary school do predictive relations between these two variables exist? Multiple Cohort Multiple Group analysis with Mplus 7.2 (Muthén and Muthén, 1998–2012), was used to fit a bivariate change score model in four cohorts, separately at each of the four levels of orthographical complexity (cf. Gniewosz and Gniewosz, 2018).

To answer the third research question, we searched for poor readers in Grade 1 with different score profiles on the available decoding tasks, accuracy as well as efficiency. We then followed up the identified groups of readers during grades 2 and 3 to assess how they performed on word decoding accuracy and efficiency. The same process was carried out in the second cohort of students, identifying groups of readers with similar profiles in Grade 2 and following up their word reading performance in grades 3 and 4. The second cohort was included in these analyses because this was the first cohort where the test with polysyllabic words and pseudowords was administered. Latent Profile Analysis (LPA; Masyn, 2013; Ferguson et al., 2020) was used with maximum likelihood estimation by Mplus 7.2 to identify groups of (poor) readers in Grade 1 (Cohort 1) and in Grade 2 (Cohort 2).

Results

Descriptive statistics

Means and standard deviations of the test scores are shown in **Tables 2**, **3**. **Table 2** presents the statistics for word

TABLE 2 Means and standard deviations (within parentheses) of word decoding efficiency for different types of words and pseudowords in four cohorts of students at the end of three school years in each cohort.

		Words[1]				Pseudowords[2]				
Coh[3]	**Grade**	**CVC**	**CC**	**Bisyl**	**Poly**	**CVC**	**CC**	**Bisyl**	**Poly**	***N***
1	1	39.2 (20.34)	25.6 (16.94)	12.2 (10.15)		28.9 (17.22)	18.2 (11.21)	8.7 (7.64)		198
1	2	65.8 (21.12)	51.0 (21.76)	30.7 (16.02)		47.0 (17.17)	32.4 (15.86)	18.4 (10.38)		183
1	3	81.6 (18.22)	69.7 (19.77)	51.0 (18.59)		60.2 (14.41)	47.5 (16.66)	30.9 (12.38)		166
2	2	61.4 (19.06)	48.2 (21.00)	28.2 (15.70)	19.8 (11.14)	44.9 (15.70)	30.7 (14.39)	17.3 (9.55)	11.7 (7.08)	216
2	3	79.9 (17.05)	66.0 (19.62)	46.4 (18.40)	34.1 (14.52)	56.1 (14.12)	42.3 (16.17)	25.5 (11.92)	17.1 (8.78)	173
2	4	91.9 (18.51)	79.3 (20.01)	62.1 (19.38)	47.8 (16.97)	62.9 (11.99)	52.3 (16.91)	33.5 (13.01)	23.6 (10.00)	153
3	3	76.6 (17.46)	64.1 (20.27)	45.5 (19.29)	32.8 (14.36)	54.6 (14.86)	40.0 (15.93)	23.9 (11.53)	16.6 (8.96)	205
3	4	89.5 (16.97)	77.1 (18.83)	60.9 (19.07)	45.6 (15.88)	62.0 (13.19)	49.1 (16.99)	31.1 (12.96)	21.9 (9.84)	198
3	5	97.6 (17.05)	85.5 (17.95)	71.05 (17.87)	56.3 (16.26)	68.13 (10.20)	58.9 (14.44)	40.3 (13.46)	30.9 (12.16)	181
4	4	90.1 (15.87)	77.1 (18.25)	63.7 (18.48)	48.4 (14.63)	63.6 (11.77)	52.0 (15.00)	34.0 (11.60)	24.7 (8.91)	199
4	5	101.2 (17.64)	89.4 (18.11)	75.6 (18.54)	59.8 (16.45)	68.3 (9.89)	59.8 (14.33)	40.1 (12.11)	29.5 (9.81)	189
4	6	109.1 (17.90)	97.6 (18.66)	84.2 (17.56)	68.4 (16.15)	70.3 (7.49)	63.7 (12.68)	45.2 (11.66)	34.3 (10.66)	296

[1] Number of words correctly read in 1 min. [2] Number of words correctly read in 2 min divided by 2. [3] Number of cohort of students is the grade level at which the measurements started.

TABLE 3 Means and standard deviations (within parentheses) of word decoding accuracy for different types of words and pseudowords in four cohorts of students at the end of three school years in each cohort.

		Words[1]				Pseudowords[2]				
Coh[3]	**Grade**	**CVC**	**CC**	**Bisyl**	**Poly**	**CVC**	**CC**	**Bisyl**	**Poly**	***N***
1	1	90.7 (12.25)	81.4 (20.26)	59.2 (26.22)		83.0 (16.77)	75.2 (18.98)	47.8 (24.78)		198
1	2	94.8 (10.52)	92.45 (13.93)	85.0 (17.86)		89.9 (12.74)	84.7 (16.44)	71.1 (20.23)		183
1	3	98.3 (4.10)	97.7 (6.93)	95.0 (11.20)		96.1 (6.79)	94.9 (9.60)	88.1 (14.90)		166
2	2	95.6 (5.66)	93.2 (7.83)	83.8 (16.17)	79.5 (17.65)	91.2 (8.98)	85.5 (13.10)	71.2 (19.03)	61.8 (21.62)	216
2	3	97.7 (3.09)	96.4 (4.46)	93.2 (7.56)	90.0 (9.69)	93.5 (7.61)	88.9 (12.17)	77.7 (18.26)	68.2 (20.87)	173
2	4	98.7 (2.09)	98.1 (2.77)	97.2 (3.79)	94.8 (5.72)	95.6 (5.24)	93.6 (6.44)	86.3 (12.53)	78.6 (14.97)	153
3	3	97.5 (3.70)	95.5 (6.00)	91.8 (9.45)	89.1 (9.69)	92.8 (8.72)	87.6 (11.40)	75.5 (17.97)	66.6 (20.13)	205
3	4	98.7 (2.06)	97.8 (3.53)	96.5 (4.95)	94.3 (6.16)	95.2 (5.20)	92.7 (8.75)	84.5 (13.07)	76.3 (15.38)	198
3	5	99.3 (1.40)	99.1 (2.30)	98.2 (2.92)	97.5 (5.72)	97.4 (4.28)	96.5 (6.09)	91.3 (11.23)	85.0 (15.56)	181
4	4	98.9 (1.64)	98.0 (3.15)	97.0 (4.42)	95.6 (5.31)	96.1 (4.68)	93.5 (7.17)	87.3 (11.52)	81.5 (13.34)	199
4	5	99.1 (1.36)	99.0 (2.52)	98.2 (3.06)	97.2 (3.47)	96.8 (4.07)	95.4 (6.53)	90.7 (9.23)	83.6 (12.79)	189
4	6	99.5 (1.12)	99.4 (1.51)	99.1 (1.84)	98.5 (2.45)	97.7 (3.08)	97.1 (4.08)	93.4 (7.28)	88.2 (10.95)	296

[1] Percentage of words correctly read in 1 min. [2] Percentage of words correctly read in 2 min. [3] Number of cohort of students is the grade level at which the measurements started.

and pseudoword decoding efficiency, number of words read correctly in 1 min. Table 3 presents the statistics for accuracy on the same tasks, the percentage of words and pseudowords read correctly. For each student three scores were available in three consecutive school years. The data in both tables were from four cohorts of students with test scores at overlapping school years, three grade levels in each cohort, from grades 1 to 3 in the first cohort up to grades 4 to 6 in the last cohort. The means in Table 2 show some clear patterns. As to be expected, there were differences between word types and between grade levels, not so much between cohorts.

In general, at all grade levels, the average number of words read correctly was higher for CVC words than for CC words, and the difference was largely the same at all grade levels (about 12 to 14 words). The same appeared true for CVC and CC pseudowords. Also the means for CC words were clearly larger than for bisyllabic words to about the same extent at all grade levels (13–20 words), and approximately the same was true for pseudowords. Longer words having more than two syllables were not used in Grade 1 and were therefore not available for students of the first cohort. For the other three cohorts it appeared that the means were consistently lower for the longer words and pseudowords (10–15 words) compared with bisyllabic (and one-syllabic) words and pseudowords. Thus, the four levels of orthographic complexity were in a clear order of difficulty. And, as could be expected, reading pseudowords appeared to be more difficult than reading real words. Moreover, Table 2 shows that the average lexical difference was larger at higher compared with lower grade levels. (For instance, for CVC in Cohort 1 the average lexical difference was 10.3 words in Grade 1, 18.8 words in Grade 2, and 21.4 in Grade 3, which was an increase from about a half to more than one standard deviation. And for polysyllabic words in Cohort 2 the average lexical difference equaled 8.1 (0.73 s.d.) in Grade 2, 17.0 (1.17 s.d.) in Grade 3 and 24.2 (1.43 s.d.) in Grade 4.

As expected, means were growing larger by grade level, and the mean differences between consecutive grade levels became smaller in higher compared with lower grades, suggesting a non-linear relation between test score and grade. For longer words, however, the relationship between test score and grade was close to linear within cohorts. No large differences between cohorts were visible, except for the first cohort and especially with longer words.

Table 3 shows very high percentages correct for CVC words. Even in first grade the percentage of correctly read words was on average above 90%, but with a relatively high standard deviation.

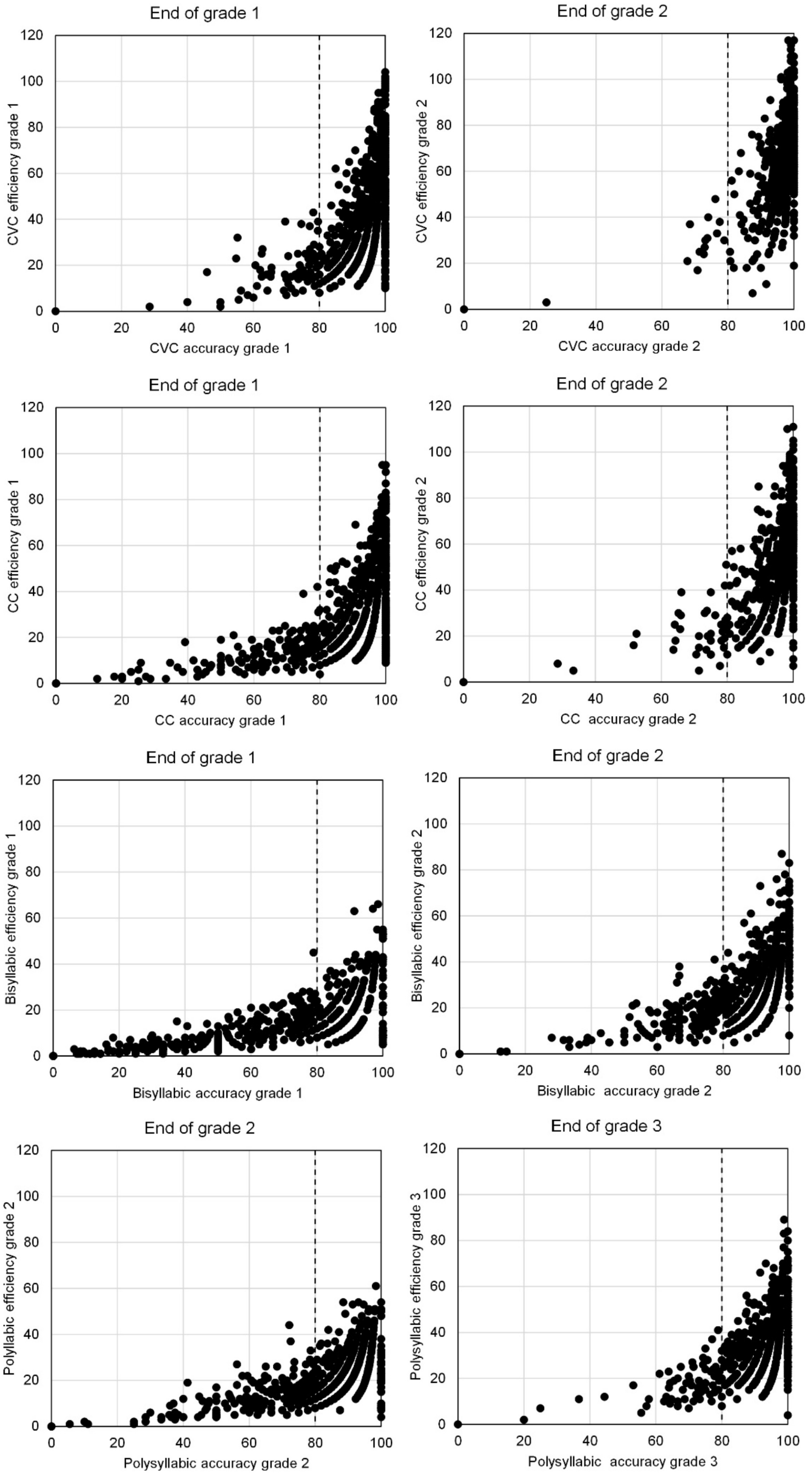


FIGURE 1
Scatterplots of accuracy (percentage correct) and efficiency (number of words correct in 1 min) of word decoding at the end of Grade 1 and Grade 2, for CVC, CC, and bisyllabic words. Also scatterplot of accuracy and efficiency of reading polysyllabic words at the end of Grades 2 and 3.

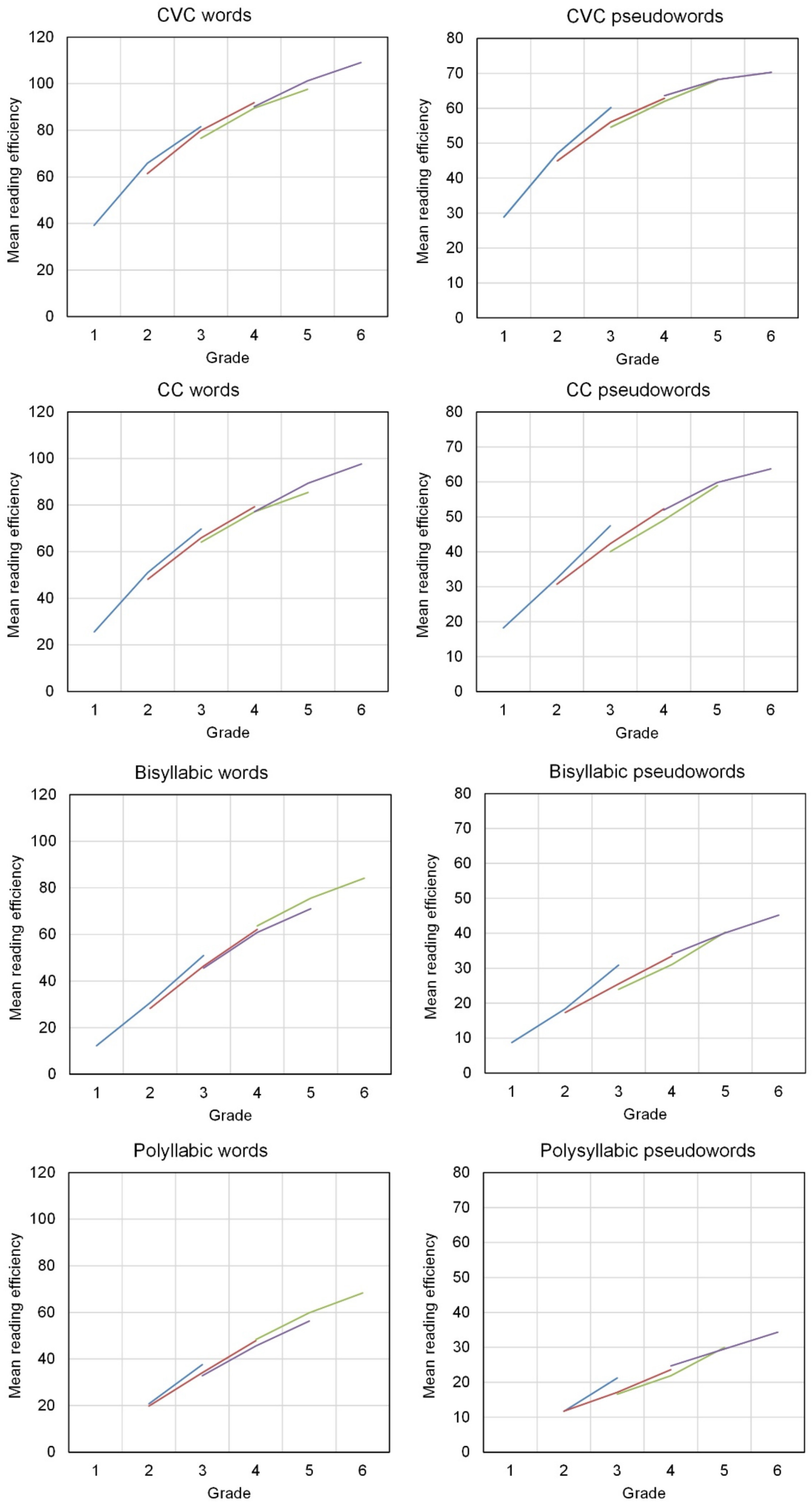


FIGURE 2
Average development of word decoding efficiency (number of words correctly read in 1 min) across Grade 1 thru 6 (blue line: Cohort 1, red line: Cohort 2, green line: Cohort 3, purple line: Cohort 4).

A very small number of students had an accuracy of 65% or less. In higher grade levels the average accuracy soon increased to 99% or higher and the standard deviation sharply decreased from Grade 3 on. For CC words the average accuracy started at 81% in Grade 1 with a very high standard deviation. One grade later the mean and standard deviation were at about the same level as for CVC in Grade 1. From Grade 3 about the same high average accuracy level was reached as for CVC words. A low average accuracy of 59% was achieved for bisyllabic words in Grade 1, but from Grade 3 on high levels of accuracy were achieved, with still a relatively high standard deviation though. The same pattern was observed for polysyllabic words but starting from Grade 2.

As expected, accuracy was clearly lower for decoding pseudowords than for decoding real Dutch words. And Table 3 also shows that accuracy of decoding pseudowords increased quickly with grade level. For monosyllabic pseudowords (CVC and CC) in Grade 6 almost the same average accuracy was achieved as for decoding real CVC and CC words. For longer pseudowords the average level of accuracy still stayed behind that for real words, even in Grade 6.

Developmental relations between accuracy and efficiency

The first research question was about the development of decoding efficiency across grades and the developmental relations between accuracy and efficiency of word and pseudoword reading. Dummy variables were used to represent cohort differences in intercepts. A quadratic growth model was fit. The parameters of change were the slopes of the linear and quadratic component, an intercept and intercept differences between cohorts. Age of the students at the first measurement occasion was used as a covariate, predicting word or pseudoword decoding efficiency at the first measurement occasion in each cohort. A threshold for decoding accuracy was used as a time-varying covariate in the growth models to study the hypothesized difference in growth of decoding efficiency for students who did or did not meet the accuracy threshold. We compared three growth models for each word or pseudoword decoding test: (1) a model including only the development of word decoding efficiency disregarding accuracy, (2) the same model including dichotomized decoding accuracy as a predictor of decoding efficiency, and (3) a model of the development of word decoding efficiency differentiated for two levels of decoding accuracy. As a fit index to compare these models we used Akaike's Information Criterion (Burnham and Anderson, 2004).

We first took a graphical look at the relation between accuracy and efficiency. Comparing Tables 2, 3 suggests that accuracy developmentally precedes efficiency of word decoding. This becomes clearer when plotting efficiency versus accuracy for the same reading task at the same time point (see Figure 1). The figure shows the two lowest grade levels for each type of words. For monosyllabic and bisyllabic words these are Grades 1 and 2. The plots suggest that a certain level of accuracy is needed before decoding efficiency can start to rise. The point of 80% accuracy is marked in the figures. Beyond that point decoding efficiency was clearly growing. Not only the mean, also the variance in scores of word decoding efficiency increased strongly when a minimum level of accuracy was reached. Accuracy is a necessary condition for decoding efficiency but is by no means a sufficient condition. A sizable number of students lags behind in decoding efficiency despite a satisfying level of accuracy. In case of monosyllabic words almost all students surpassed 80% accuracy in Grade 2. For polysyllabic words, Figure 1 shows accuracy-efficiency plots for Grades 2 and 3. In these plots a gradual increase in word decoding efficiency is already seen for students with less than 80% accuracy in Grade 2.

For word decoding efficiency to grow a minimum level of accuracy is needed. Therefore, we developed a model for growth of word decoding efficiency with a threshold for decoding accuracy. We developed this model in two steps. First, we studied a growth model for decoding efficiency for words and pseudowords at various levels of orthographic complexity. Next, we added an accuracy threshold to the resulting models to study how decoding accuracy moderates the growth of decoding efficiency.

Development of word and pseudoword decoding efficiency

Figure 2 shows the average development of word decoding efficiency throughout grades 1–6, for the four types of words and the four types of pseudowords. Each plot shows developmental curves for the four cohorts of students. In most cases the lines for the cohorts coincide nicely; a curve can be drawn that well represents all cohorts, with a few deviations. The average developmental curve is clearly curvilinear. The average efficiency score is increasing with grade level, but the average growth is diminishing in higher grades. Having only three time points per cohort, the graphs suggest a quadratic model of decoding efficiency. We specified word decoding efficiency as a quadratic function of grade level. Dummy variables for cohort effects were in the model as well as the age in months of the students at the first measurement occasion for each cohort. We estimated this quadratic model for each word decoding test separately. The average developmental curve is determined by two parameters: the constant increase with grade level, which we called linear change, and the quadratic component, which we called acceleration. The acceleration of the growth was in all cases negative; the growth diminished with increasing grade level.

Table 4 shows the fixed effect estimates for all eight models: the intercepts (for Cohort 1), the intercept differences of cohorts compared with Cohort 1, the linear and quadratic slopes. Grade

TABLE 4 Growth curve analyses of word decoding efficiency (number of words correctly read in 1 Min): fixed effect estimates[c].

	Intercept	Linear change	Acceleration	Intercept cohort 2[a]	Intercept cohort 3[a]	Intercept cohort 4[a]	Age[b]
Words							
CVC	70.62	15.81	−1.85	3.00	7.65	16.65	−0.60
CC	54.47	16.41	−1.88	5.24	10.34	22.41	−0.70
Bisyllabic	34.66	16.40	−1.49	4.43	13.85	27.77	−0.82
Polysyllabic	23.23	14.12	−0.98	5.15	11.81	23.02	−0.69
Pseudowords							
CVC	50.33	10.54	−1.78	0.31	5.09	12.59	−0.49
CC	32.38	11.50	−1.19	4.24	9.38	19.29	−0.68
Bisyllabic	18.76	9.15	−0.68	2.79	5.99	14.28	−0.54
Polysyllabic	10.01	6.76	−0.31	3.66	6.92	15.53	−0.50

[a] Difference in intercept compared with the intercept of the youngest cohort. [b] Age in months centered at the grand mean. [c] All estimated coefficients are statistically significant ($p < 0.05$), except for the intercept difference between Cohort 1 and 2 with CVC pseudowords.

level was centered at Grade 3. Thus, the intercepts refer to the expected number of words or pseudowords read correctly at the end of Grade 3. The variances (not in the table) were allowed to differ by grade level. The variance was decreasing by grade level for CVC words and pseudowords. This pattern of development, growing mean and decreasing variance, signifies that most students in higher grades reach a high level of decoding efficiency with simple, short words and pseudowords.

Table 4 shows large differences between intercepts, lower intercepts for longer and more complex words and lower intercepts for pseudowords than for real words. The linear change per school year did hardly differ for the various orthographic complexities, was only a bit lower for polysyllabic words. For pseudowords the linear change was clearly lower than for real words, especially for the longer pseudowords. So, for pseudowords, both the level of performance, as shown by the intercepts, and the growth of decoding efficiency, as shown by the linear change, were lower than for real words. The quadratic coefficients, representing the acceleration of the growth, were less negative for the longer words, especially the longer pseudowords. The intercepts showed clear cohort differences. In a cohort starting at a higher grade level the intercept difference was higher. For polysyllabic words and pseudowords, the between-students variance was linearly increasing with grade level. For CC words and pseudowords, the between-student variance increased up to Grade 4 and then decreased with further growth of decoding efficiency. The effect of age was negative. Younger students within a cohort appeared to score higher than older students.

TABLE 5 Fit comparison of growth models: Akaike's information criterion (AIC).

Variables	Growth model (model of Table 4)	Accuracy threshold (80%) added	Differential growth below and above threshold (Figure 4)
Words			
CVC	19508.71	19410.46	19390.08
CC	19786.60	19646.55	19625.54
Bisyllabic	19286.23	19130.95	19078.41
Polysyllabic	17077.94	16909.18	16899.18
Pseudowords			
CVC	17839.13	17523.00	17444.83
CC	18156.72	17927.13	17905.88
Bisyllabic	16988.60	16445.52	16423.50
Polysyllabic	14760.31	14167.20	14134.81

Development of word decoding efficiency moderated by accuracy

As the last step to answer the first research question we added decoding accuracy as a predictor to the growth curve models. Following our exploration of the accuracy-efficiency relationship in our data (see **Figure 1**) and the studies by Juul et al. (2014) for Danish and Karageorgos et al. (2019) for German, we hypothesized that a certain level of reading accuracy is needed for decoding efficiency to develop. Both studies suggested a threshold of 70%. We used a threshold of 80% because this is common in mastery learning (see for instance Reynolds et al., 2021), and because of the small number of students scoring below 70% in the easiest reading tasks. Unlike the other studies, we used several reading tasks with differing degrees of lexical complexity. For most of these tasks, a threshold of 80% seemed better than a threshold of 70%, see **Figure 1**. We expected some development of decoding efficiency below 80% accuracy, and a much stronger development for students with more than 80% accuracy. Therefore, we introduced decoding accuracy in our models as a dichotomized variable indicating whether a student had reached 80% accuracy by the end of a grade level. This accuracy variable was supposed to potentially influence all parameters of the growth curve: the intercept, the linear change, the acceleration, and the intercept

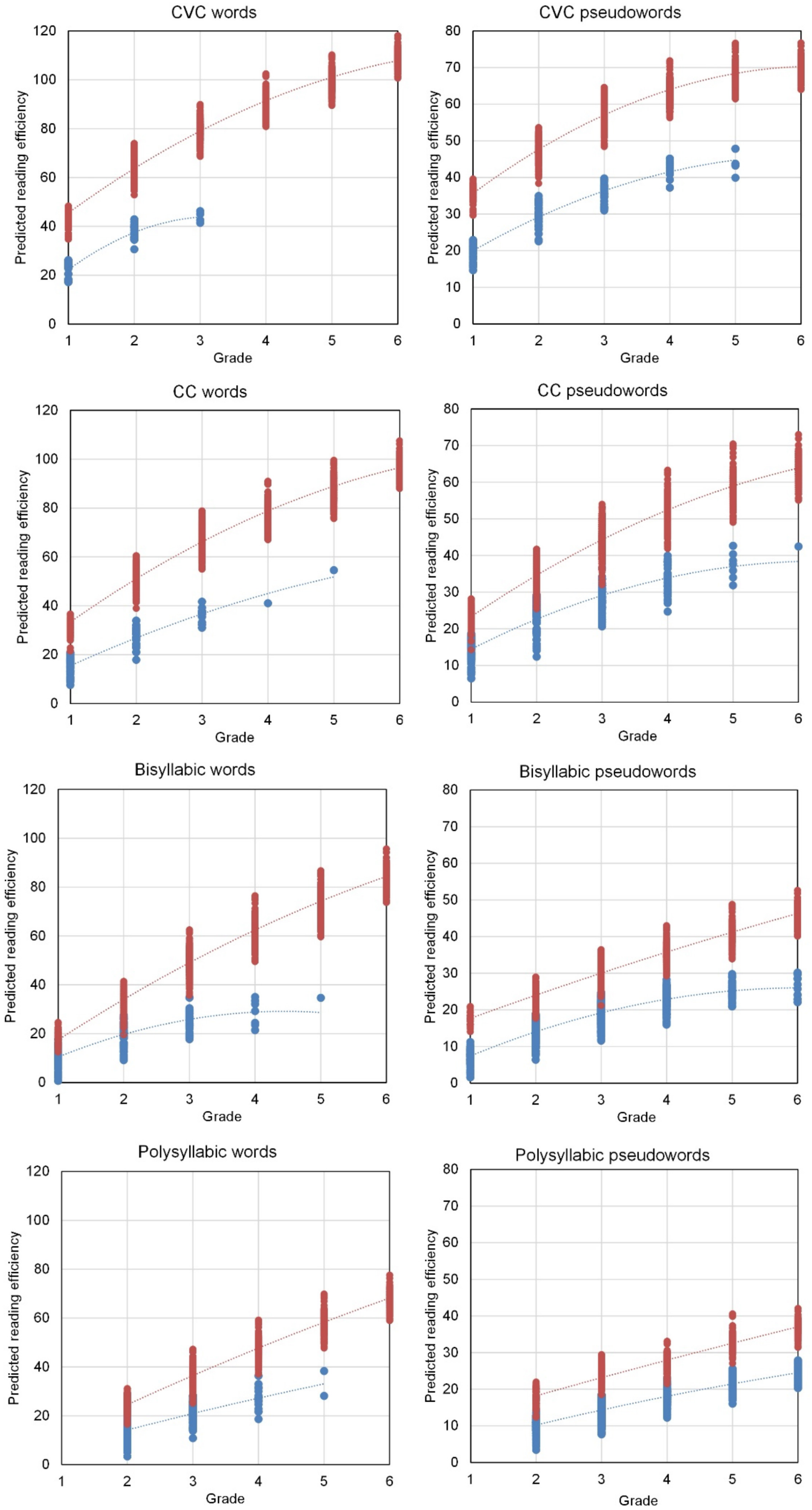


FIGURE 3

Predicted reading efficiency (words per minute) per grade level for students with 80% or higher reading accuracy (in red above) and for students with less than 80% accuracy (in blue below).

differences between cohorts. All these effects were in the model as interaction effects.

To select the most appropriate model, we compared the fit of three growth models for each word decoding test. We used AIC to compare these models (see **Table 5**). The best model is the model with the lowest AIC. The AIC's in **Table 5** show that the model with the accuracy threshold included is in all cases clearly preferable to the growth model of **Table 4**. In addition, the AIC's confirm that growth of word decoding efficiency differs for students with word decoding accuracy below the threshold from the growth for students that met the accuracy threshold.

The differential growth model resulted in two different growth curves for decoding efficiency, one for students that reached the 80% accuracy criterion and one for students that did not (yet) reach this accuracy criterion. The results are in **Figure 3**. The eight plots in this figure show the predictions of decoding efficiency derived from the growth model for individual students. The prediction plots for the grade levels were connected by interpolation lines. The predicted decoding efficiency appeared lower and appeared to have less steep development for students with accuracy below the threshold. Consequently, the two growth curves diverged across the grades. For CVC words, there were no students with less than 80% accuracy beyond Grade 3. For CC words, only few students scored below the accuracy threshold in grades 4 and 5. For bisyllabic words, only few students scored below 80% accuracy in Grade 5 and none in Grade 6. For polysyllabic words, decoding efficiency was linearly increasing with grade level and the same for each cohort. The linear change was lower for students with decoding accuracy below 80%.

The plot for CVC pseudowords shows a curvilinear development but the effect of accuracy on efficiency appeared linear, and again development of decoding efficiency was slower for students with below 80% accuracy. Accuracy of reading CC pseudowords affected both the level and the linear change in decoding efficiency across grade levels. For bisyllabic pseudowords, the accuracy dichotomy additionally affected the acceleration of decoding efficiency across grade levels. Finally, the model for decoding efficiency of polysyllabic pseudowords involved all interaction effects including cohort effects. In all cases, the development of decoding efficiency showed a clearly slower pace when accuracy was not yet above 80%.

Developmental relations between word and pseudoword decoding efficiency

To answer the second research question, we performed multiple-group multiple-cohort analyses for each of the four levels of orthographic complexity separately. The analyses involved word decoding and pseudoword decoding efficiency simultaneously to determine to what extent the amount of change in word decoding efficiency between two adjacent grade levels was related to the level of pseudoword decoding efficiency. Also, we wanted to determine to what extent the amount of change in pseudoword decoding efficiency was related to the level of word decoding efficiency (see **Figure 4**). Decoding accuracy was not included, because it approached its ceiling already in an early school grade. Like in the growth curve

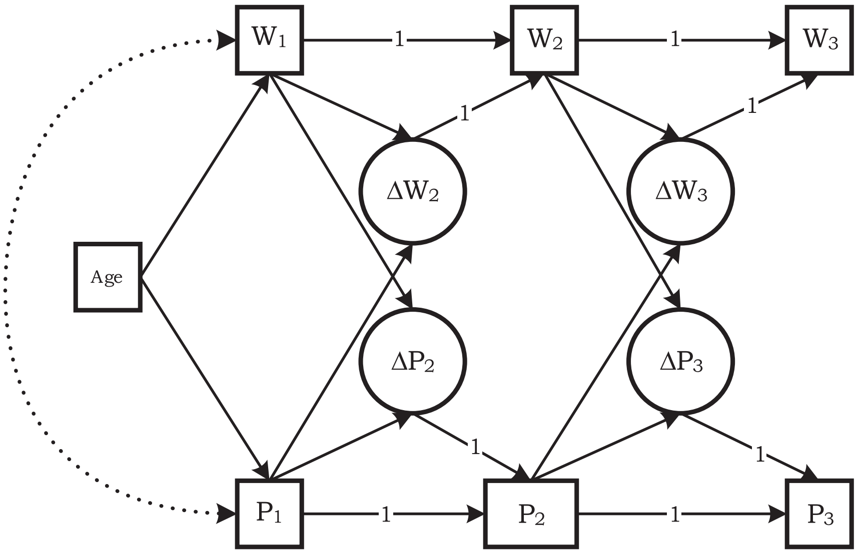


FIGURE 4
Bivariate change score model of Word and Pseudoword decoding Efficiency for the first cohort of students, grades 1–3. W1 is Word decoding Efficiency and P1 is Pseudoword decoding Efficiency observed at the end of Grade 1; ΔW2 is the change score in W2. Paths with a coefficient of 1 were fixed at a value of 1; all other paths had a coefficient to be estimated. Intercepts and residual variances and covariances omitted.

analyses reported in **Table 4**, we employed children's age in months (centered at the grand mean) as a covariate to account for cohort differences.

In each of the four cohorts two change scores were identified both for word decoding efficiency and for pseudoword decoding efficiency: the change between the first and second measurement occasion and the change between the second and third measurement occasion in the cohort (see **Figure 4**). For instance, the word decoding efficiency score at Time 2 (W2) can be written as the word decoding efficiency score at Time 1 (W1) plus the score change from Time 1 to Time 2 (ΔW2). The latent change score ΔW2 is defined by W2 = 1^{*}W1 + $1^{*}\Delta$W2. Thus, this change score is the part of the score of W2 that is not identical to W1 (McArdle, 2009, p. 583). A similar model was specified for word decoding efficiency at the third measurement point (W3). Also, the same model with two latent change scores was specified for pseudoword decoding efficiency at the three time points (P1–P3). The model also accounted for dependencies of change scores on the score level one year earlier. For instance, we regressed the change score ΔW2 on W1. Our second research question pertained to the paths from P1 to ΔW2, from P2 to ΔW3, from W1 to ΔP2, and from W2 to ΔP3. The model for Cohort 1 (**Figure 4**) applied to Grade 1 to Grade 3. This was repeated for each cohort. Due to the overlapping cohorts some change scores were present in two cohorts. For instance, the change from Grade 2 to Grade 3 was measured in the first and the second cohort. Therefore, equality restrictions were placed on the slopes relating to such overlapping change scores. The models were fitted, for each level of orthographic complexity separately, using maximum likelihood estimation. The goodness of fit of models was evaluated by the chi-square statistic and the Root Mean Square Error of Approximation (RMSEA).

The models showed a good fit at three of the four levels of orthographic complexity [for CVC $\chi^2(28) = 45.89$, $p = 0.018$, RMSEA = 0.053, for CC $\chi^2(28) = 38.10$,

TABLE 6 Change of word and pseudoword decoding efficiency in adjacent grades regressed on the preceding levels of word and pseudoword decoding efficiency (estimated unstandardized regression coefficients).

Predictors	**Change in word decoding efficiency**					**Change in pseudoword decoding efficiency**				
	Δ W2	**Δ W3**	**Δ W4**	**Δ W5**	**Δ W6**	**Δ P2**	**Δ P3**	**Δ P4**	**Δ P5**	**Δ P6**
CVC	**−0.245**	**−0.381**	**−0.303**	**−0.303**	0.040	**0.393**	**0.175**	0.022	0.056	−0.011
P-CVC	0.157	**0.210**	**0.299**	**0.256**	−0.140	**−0.598**	**−0.483**	**−0.226**	**−0.388**	**−0.338**
CC	−0.300	**−0.591**	**−0.435**	**−0.317**	0.108	**0.290**	**0.215**	**0.166**	**0.128**	0.019
P-CC	**0.533**	**0.577**	**0.406**	**0.207**	−0.241	−0.235	**−0.335**	**−0.227**	**−0.315**	−0.201
Bisyl	0.026	**−0.255**	**−0.165**	**−0.153**	−0.157	**0.278**	0.095	**0.113**	**0.152**	0.010
P-Bisyl	**0.481**	**0.504**	**0.209**	0.041	0.102	**−0.229**	−0.088	**−0.157**	**−0.287**	−0.112
Poly		−0.086	**−0.248**	−0.057	−0.064		**0.202**	**0.115**	**0.143**	−0.043
P-Poly		**0.459**	**0.512**	0.047	−0.032		**−0.213**	**−0.134**	**−0.216**	0.076

ΔW2 is the change score end of Grade 2 in word decoding efficiency, ΔP2 idem in pseudoword decoding efficiency. Predictors assessed at the first time point for each change. Analyses separately for the four levels of orthographic complexity; words and pseudowords analyzed together. Equality restrictions on slopes across cohorts. Statistically significant coefficients ($p < 0.05$) in boldface type.

TABLE 7 Goodness-of-fit statistics for LPA models for grade 1 (cohort 1, *n* = 198) and for grade 2 (cohort 2, *n* = 216).

Model	**AIC**	**SABIC**	**Entropy**	**Smallest class %**	**LMR *p***
Grade 1					
1	19900.24	19903.13			
2	18545.41	18549.86	0.973	35.8	0.014
3	17919.27	17925.28	0.973	9.7	0.271
4	17355.09	17362.67	0.977	7.0	0.323
5	17064.11	17073.25	0.965	6.6	0.026
Grade 2					
1	27577.47	27584.08			
2	25884.60	25894.72	0.955	36.8	0.053
3	24848.97	24862.59	0.970	24.3	0.054
4	24376.38	24393.51	0.962	10.1	0.127
5	24046.38	24067.03	0.965	6.9	0.408

Model denotes the number of profiles distinguished. AIC, Akaike's Information Criterion; SABIC, Sample-Adjusted Bayesian Information Criterion; LMR *p*, *p*-value for the LMR test.

$p = 0.096$, RMSEA = 0.049, for Bisyllabic $\chi^2(28) = 41.28$, $p = 0.051$, RMSEA = 0.045]. The exception was the model for polysyllabic words and pseudowords [$\chi^2(20) = 71.04$, $p < 0.001$, RMSEA = 0.102]. According to the information criteria AIC and BIC, however, the model with the appropriate equality restrictions was a better model for the polysyllabic data than the model without these equality restrictions. Intercept differences between cohorts were freely estimated and not tested in these analyses. We were only interested in the slopes, especially for the relations between word and pseudoword decoding efficiency in two consecutive grade levels.

Table 6 shows the relevant parameter estimates. For each level of orthographic complexity, word and pseudoword decoding efficiency were analyzed simultaneously. The criterion variables were the yearly changes in word and pseudoword decoding efficiency between Grade 1 and Grade 6. The table shows the effects of the levels of word and pseudoword decoding efficiency on the next changes. In general, change in decoding efficiency appeared smaller for students with higher scores on decoding efficiency. The regression coefficients of change on momentary level of decoding efficiency were mostly negative in all grade levels, except Grade 6. The change from Grade 5 to Grade 6 appeared unrelated to the level of word decoding efficiency achieved in Grade 5. But for the change in pseudoword decoding efficiency from Grade 5 to Grade 6 a statistically significant negative coefficient was observed.

For the short words (CVC and CC) the immediately preceding level of pseudoword decoding efficiency was positively associated with the increase in word decoding efficiency in all grades up to Grade 5 (see Table 6). But for the longer words (bisyllabic and polysyllabic) there was no significant positive effect anymore after Grade 4. The positive effects of pseudoword decoding efficiency on growth in word decoding efficiency were strongest in the early grades, especially for the higher levels of orthographic complexity.

The change in pseudoword decoding efficiency appeared positively related to the preceding level of word decoding efficiency, again except in Grade 6. For CVC words, the positive effects of word on pseudoword decoding efficiency appeared only in the lower grades (1–3), while the effects of pseudoword on CVC word decoding continued to exist until Grade 5 (Table 6). For the three more complex word types, the growth of pseudoword decoding efficiency was positively associated with the level of word decoding efficiency throughout the whole range of elementary school grades until Grade 5.

Differentiating accuracy and efficiency problems in decoding development

To answer the third research question Latent Profile Analyses were performed. We concentrated on the early grades. Profiles of decoding performance were searched for at the end of Grade 1. All available test scores at the end of Grade 1 were used, word decoding accuracy as well as efficiency at both levels of lexicality and at three levels of orthographic complexity, 12 variables in total. To evaluate the number of profiles we used the AIC, the Sample-Adjusted Bayesian Information Criterion (SABIC), and the Entropy of the classification (Ferguson et al., 2020). The entropy is a measure of the quality of the assignment of the students to the latent profiles. The classification is considered good when the entropy is above 0.80

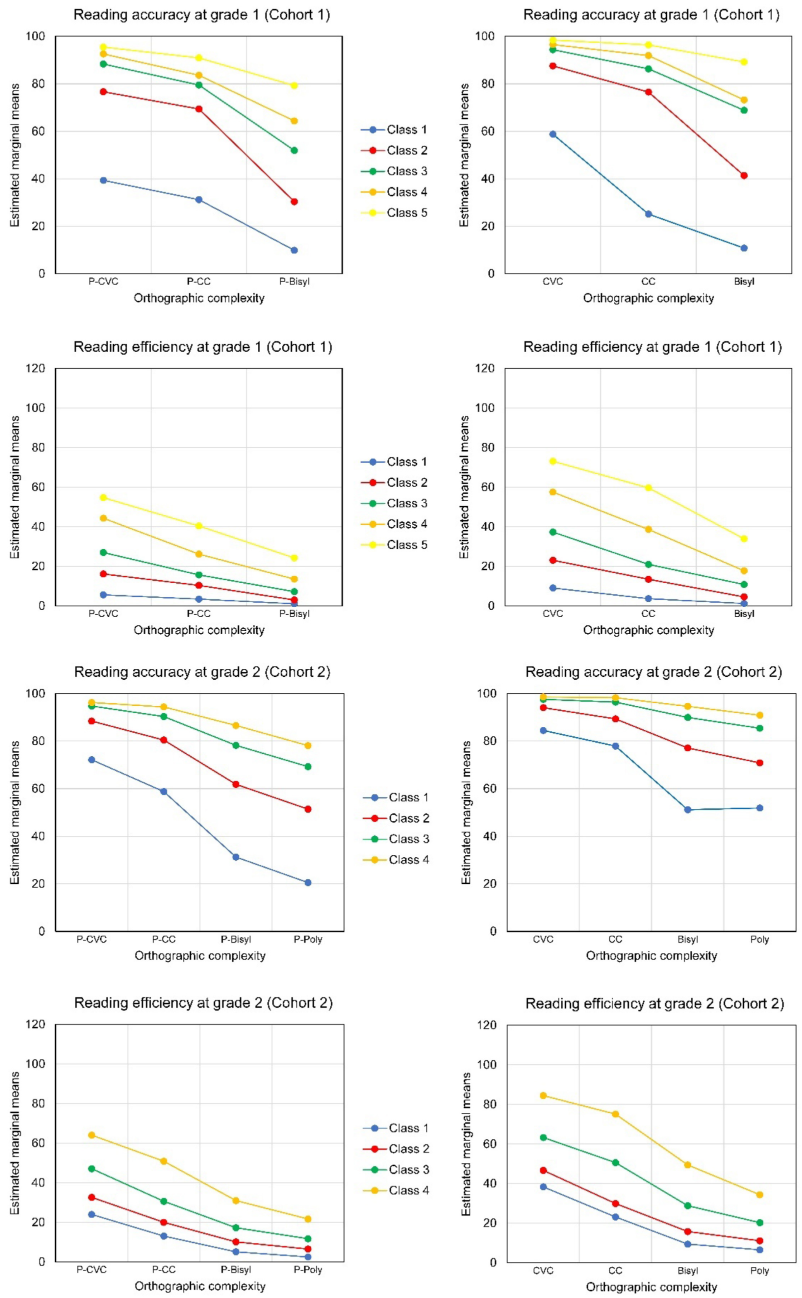
TABLE 8 Estimated means of five-profile model for grade 1 (cohort 1)**.

Variable	Profile 1 ($n = 13$)	Profile 2 ($n = 66$)	Profile 3 ($n = 54$)	Profile 4 ($n = 42$)	Profile 5 ($n = 23$)	Overall mean
	Word decoding accuracy					
CVC (7.83)*	58.78	87.57	**94.38**	**96.52**	**98.42**	90.71
CC (11.69)	25.12	76.56	**86.32**	**91.94**	**96.40**	81.43
Bisyl (15.92)	10.78	41.37	**68.84**	**73.23**	**89.18**	59.22
	Word decoding efficiency					
CVC (7.75)	9.07	23.06	37.32	**57.56**	**73.09**	39.19
CC (5.32)	3.68	13.44	21.02	**38.66**	**59.64**	25.60
Bisyl (3.63)	1.23	4.52	10.79	**17.73**	**33.91**	12.24
	Pseudoword decoding accuracy					
CVC (9.81)	39.33	76.69	**88.32**	**92.57**	**95.47**	82.99
CC (13.11)	31.22	69.41	**79.46**	**83.58**	**90.95**	75.17
Bisyl (15.43)	9.88	30.36	**51.92**	**64.38**	**79.27**	47.84
	Pseudoword decoding efficiency					
CVC (5.99)	5.57	16.14	26.99	**44.29**	**54.72**	28.88
CC (4.21)	3.42	10.36	15.69	**26.18**	**40.42**	18.22
Bisyl (3.14)	1.04	3.03	7.18	**13.56**	**24.26**	8.74

*Within-class standard deviation between parentheses. The standard deviation was restricted to be equal for all latent classes. **Class means above the overall mean are in boldface type.

(Ferguson et al., 2020). To further test the number of latent profiles, the Lo–Mendell–Rubin (LMR) test (Lo et al., 2001) was used; this test compares the present model to the model with one profile less. The follow-up data were analyzed with SPSS GLM Repeated Measures.

To determine the best number of latent profiles, we estimated five models with 1–5 latent profiles. The fit statistics of these models are summarized in **Table 7**. On both information criteria, AIC and SABIC, the model with five latent profiles appeared to be the best model. The drop in value for both AIC and SABIC was relatively small, though, between 4 and 5 classes. The index of entropy was equally good for all classifications; there appeared to be only very low classification uncertainty. The LMR statistic points to two or five classes as the best option. The smallest latent class had a rather small number of students when more than two latent classes were distinguished.


Reading accuracy at grade 1 (Cohort 1)

Reading accuracy at grade 1 (Cohort 1)

Reading efficiency at grade 1 (Cohort 1)

Reading efficiency at grade 1 (Cohort 1)

Reading accuracy at grade 2 (Cohort 2)

Reading accuracy at grade 2 (Cohort 2)

Reading efficiency at grade 2 (Cohort 2)

Reading efficiency at grade 2 (Cohort 2)

FIGURE 5
Average profiles of Reading Accuracy (% correct) and Efficiency (number of words per minute) for Latent Classes in Cohort 1 and Cohort 2; pseudowords to the left, real words to the right.

The model with only two latent classes did not serve our purposes well. It only distinguished students with high scores on all tests (the smallest class) from the others. The five-profile solution seemed well interpretable, see the estimated means in Table 8 and the average profiles in the upper part of Figure 5. The profiles varied by accuracy on shorter versus longer words and by lexicality differences in accuracy followed by similar differences in decoding efficiency. The profiles are in the order of performance on accuracy and efficiency. Multivariate analysis of variance showed the three-way interaction to be statistically significant [for accuracy Wilks' Lambda = 0.76, $F(8,384) = 7.05$, $p < 0.001$; for efficiency Wilks' Lambda = 0.66, $F(8,384) = 10.97$, $p < 0.001$]. Thus, the lexicality*orthographic complexity interaction differed significantly for the five latent classes. The first part of Figure 5 for Cohort 1 depicts this three-way interaction, separately for pseudowords and real Dutch words. The latent classes differed more in accuracy for the bisyllabic than for the single-syllabic words. On the other hand, for decoding efficiency the differences between profiles were largest for CVC words and pseudowords.

Class 1 stands out by very low accuracy, especially with the bisyllabic words. This class concerned poor readers with low scores on all variables. Class 5 concerned students with relatively high scores on all 12 variables. The other classes were in between. In the latent class with Profile 2, students had reached an adequate level of accuracy for reading shorter words (CVC and CC) less so, though, for the shorter pseudowords (P-CVC and P-CC). Decoding efficiency was still clearly lagging behind. Profile 3 showed all accuracies on average above the overall mean, but the accuracy for decoding bisyllabic words and especially pseudowords was still at a low level. Decoding efficiency was close to average. In Profile 4, accuracy was at a very high level for short words and pseudowords. But for the bisyllabic pseudowords much room for improvement was left. Students with Profile 5 scored relatively high on all variables including bisyllabic words and pseudowords. Decoding efficiency differed by orthographic complexity, more in latent classes 4 and 5 than in the other three classes.

For Grade 2, using four levels of orthographic complexity and therefore 16 variables in total, the model with four latent profiles seemed preferable, according to the information criteria and the LMR, see Table 7. When five latent profiles were distinguished, two classes with less students than variables appeared. Multivariate analysis of variance for decoding accuracy showed that the lexicality*orthographic complexity interaction did not significantly differ between the four latent classes [for the three-way interaction Wilks' Lambda = 0.94, $F(9,511.24) = 1.47$, $p = 0.156$]. But the accuracy scores of the four latent classes differed significantly in lexicality and in orthographic complexity [for the lexicality* latent class interaction $F(3,212) = 26.21$, $p < 0.001$, for the orthographic complexity* latent class interaction Wilks' Lambda = 0.41, $F(9,511.24) = 24.90$, $p < 0.001$]. Multivariate analysis of variance for decoding efficiency showed a statistically significant

TABLE 9 Estimated means of four-profile model for grade 2 (cohort 2)**.

Variable	Profile 1 (*n* = 22)	Profile 2 (*n* = 55)	Profile 3 (*n* = 90)	Profile 4 (*n* = 49)	Overall mean
	Word decoding accuracy				
CVC (3.88)*	84.51	94.10	**97.61**	**98.60**	95.61
CC (4.84)	77.93	89.38	**96.43**	**98.24**	93.16
Bisyl (9.90)	51.17	77.12	**90.00**	**94.64**	83.82
Poly (13.03)	51.94	70.87	**85.41**	**90.91**	79.55
	Word decoding efficiency				
CVC (11.33)	38.38	46.74	**63.33**	**84.55**	61.41
CC (11.04)	23.16	29.98	**50.61**	**75.13**	48.15
Bisyl (8.19)	9.50	15.78	**28.85**	**49.43**	28.25
Poly (6.11)	6.62	11.13	**20.33**	**34.44**	19.81
	Pseudoword decoding accuracy				
CVC (5.52)	72.18	88.42	**94.78**	**96.26**	91.20
CC (8.02)	58.79	80.47	**90.32**	**94.39**	85.53
Bisyl (10.10)	31.31	61.86	**78.25**	**86.59**	71.20
Poly (13.46)	20.53	51.40	**69.26**	**78.11**	61.76
	Pseudoword decoding efficiency				
CVC (8.75)	23.98	32.65	**47.10**	**64.05**	44.93
CC (7.23)	13.07	19.99	30.62	**50.83**	30.74
Bisyl (4.43)	5.10	10.14	17.24	**30.98**	17.34
Poly (3.43)	2.50	6.53	11.74	**21.66**	11.74

*Within-class standard deviation between parentheses. The standard deviation was restricted to be equal for all latent classes. **Class means above the overall mean are in boldface type.

three-way interaction [Wilks' Lambda = 0.77, $F(9,511.24) = 6.42$, $p < 0.001$]. Thus, the lexicality*orthographic complexity interaction for decoding efficiency differed significantly for the four latent classes in Grade 2 of Cohort 2. The lower part of **Figure 5**, for Cohort 2, depicts the three-way interaction for both accuracy and efficiency.

Profile 1 denotes students scoring on average low on accuracy, especially for the longer words, see **Table 9** and part (b) of **Figure 5**. Students with Profile 2 scored on average relatively high on decoding accuracy for single-syllable words and pseudowords, close to the overall average. But they scored low on accuracy for decoding bi- and polysyllable pseudowords. Students with Profile 3 scored high on accuracy for all orthographic complexity levels, except for polysyllabic pseudowords, though still a bit above the overall average. Students with Profile 4 were on average also relatively accurate with the long pseudowords, though still less than with the shorter words. In Profile 1, word decoding efficiency was far below average, especially for the longer words and pseudowords. In Profile 2, decoding accuracy was at an acceptable level for shorter words and decoding efficiency was higher than in Profile 1 but still (far) below average. In Profile 3, (pseudo)word decoding efficiency was about average or just above. Students with Profile 4 scored considerably higher than others on both word and pseudoword decoding efficiency for both shorter and longer words.

The five latent classes identified for the first-grade students of Cohort 1 were followed up in the next two grade levels to see how their performance changed when they progressed through elementary school. The upper part of **Figure 6** concerns latent classes in the first cohort and shows the average profiles for decoding accuracy and decoding efficiency of the latent classes from Grade 1 at the end of Grade 2 and Grade 3. The figure shows the three-way interaction Lexicality * Orthographic Complexity * Latent Class at Grade 2 and at Grade 3. For decoding accuracy in Cohort 1 this three-way interaction was statistically significant [Wilks' Lambda = 0.80, $F(8,310) = 4.58$, $p < 0.001$], while the four-way interaction with Time was not [Wilks' Lambda = 0.99, $F(8,310) = 0.195$, $p = 0.991$]. Accuracy clearly increased and approached 100% for reading real Dutch words in all but the two lowest latent classes. The accuracy differences between latent classes depended upon both lexicality and orthographic complexity of the words. In both grades the small Latent Class 1 differed strongly in decoding accuracy from the other latent classes. Students in the other classes improved a lot in accuracy compared with Grade 1 (see **Figure 5**). The differences in accuracy between latent classes 2–5 became smaller by grade level. For bisyllabic pseudowords, though, the accuracy still lagged behind in Latent Class 2, even in Grade 3. For decoding efficiency, the four-way interaction Lexicality * Orthographic Complexity * Latent Class * Time was statistically significant [Wilks Lambda = 0.88, $F(8,310) = 2.44$, $p = 0.014$]. This interaction is graphed in the upper part of **Figure 6** as two three-way interactions at each of the two follow-up grade levels. The different profiles of the latent classes from Grade 1 continued to exist at the later grade levels, but the differences in decoding efficiency between latent classes became larger by grade level.

The same follow-up analyses were executed for the four latent classes found in Grade 2 (Cohort 2), following up through Grades 3 and 4. Again, for decoding accuracy the four-way interaction involving Latent Class, Lexicality, Orthographic Complexity, and Time was not statistically significant [Wilks' Lambda = 0.92, $F(9,340.87) = 1.25$, $p = 0.262$]. The lower part of **Figure 6** shows a graph of the three-way interactions at each of the two follow-up grade levels. This three-way interaction Lexicality * Orthographic Complexity * Latent Class was statistically significant [Wilks' Lambda = 0.649, $F(9,340.87) = 7.38$, $p < 0.001$]. As before, the accuracy differences between latent classes depended upon both lexicality and orthographic complexity. The accuracy differences between latent classes were large for decoding pseudowords and small for reading Dutch words. Accuracy was particularly low for decoding the longer pseudowords (bisyllabic and polysyllabic pseudowords). The accuracy differences between latent classes became smaller by grade level. But there were still large differences between latent classes for accuracy of decoding the bi- and polysyllabic pseudowords. Like in Cohort 1, for decoding efficiency the four-way interaction was statistically significant [Wilks' Lambda = 0.852, $F(9,340.87) = 2.58$, $p = 0.007$]. See in the upper part of **Figure 6** the three-way interactions plotted for Grade 3 and Grade 4. The profiles for latent Class 3 and 4 hardly differed in terms of accuracy, but there were clear differences in efficiency at both follow-up grade levels. For real Dutch words the efficiency averages differed by orthographic complexity in all latent classes, while accuracies only differed a bit between short and long words (CVC and CC versus Bi- and Polysyllabic).

Discussion

The present study shows that the development of print tuning in Dutch as a transparent orthography is largely a matter of growing decoding efficiency. From first grade on, children were highly competent in accurately decoding words, and to a lesser extent also in decoding pseudowords. Apparently, their self-teaching device of (pseudo)word decoding as proposed by Share (2004), which is basically taught in first grade, was sufficient to foster children's decoding words with growing accuracy. Beyond the early stage of learning to read, children made very few errors and became faster in word decoding resulting in a prolonged growth of word decoding efficiency. Throughout the primary grades, children seem to make a progression from slow and sequential grapheme-by-grapheme based decoding to parallel phonology-based orthographic processing (cf. Castles and Coltheart, 2004). We found

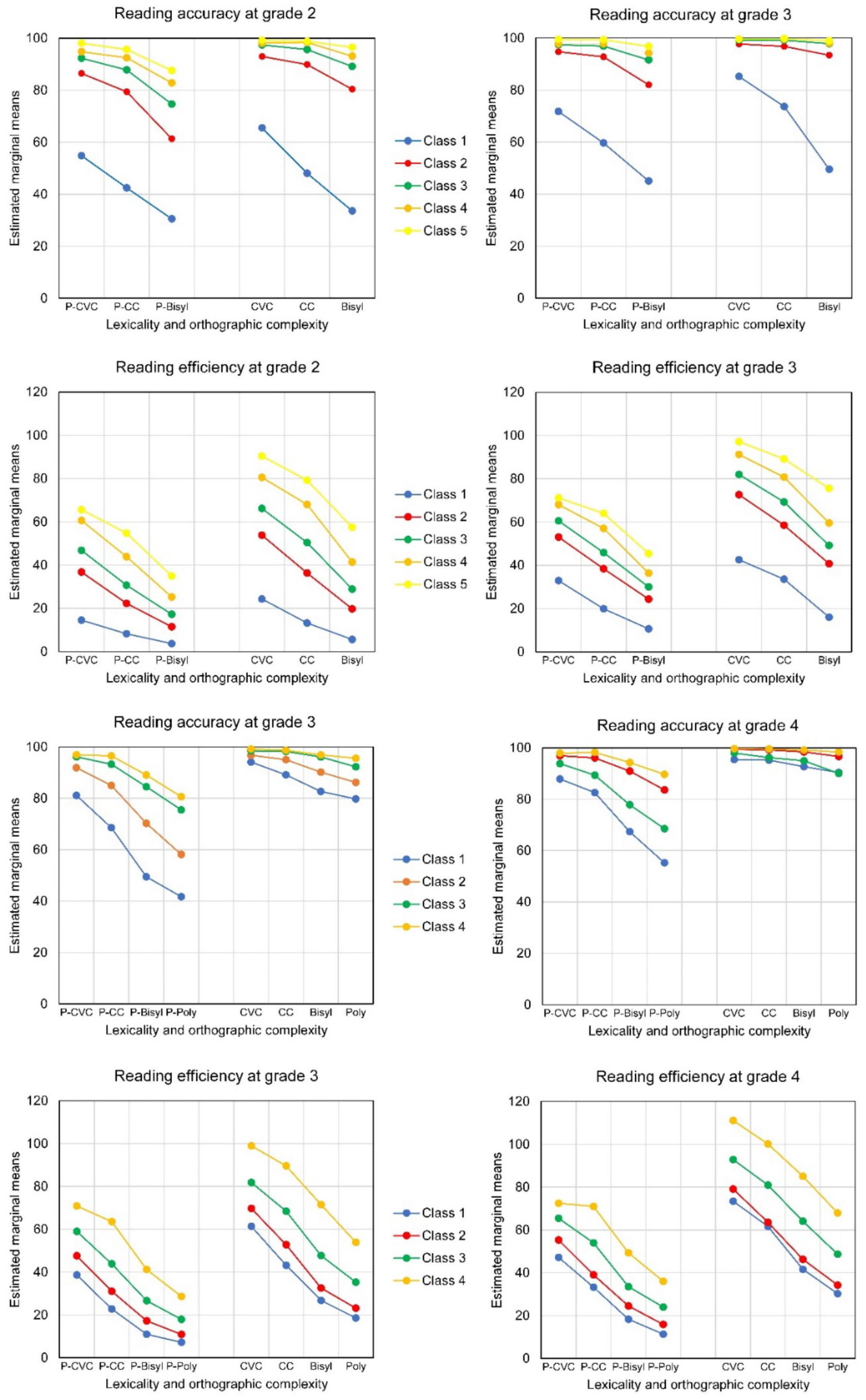


FIGURE 6
Average profiles of Reading Accuracy (% correct) and Efficiency (number of words per minute) of 2-year follow-up of the Latent Classes in Cohort 1 (Grade 2 and 3) and 2 (Grade 3 and 4). In each plot pseudowords to the left, real words to the right.

significant moderating effects for both orthographic complexity and lexical status. The orthographic complexity effect applied to both the accuracy and efficiency of pseudowords and words of beginning readers but tended to be smaller with progression of grades, which is in line with Juphard et al. (2004). Word reading efficiency appeared to develop during elementary education in about the same way for words and pseudowords and for (pseudo)words of different orthographical complexities. Development could be described as a quadratic function of grade level, with random intercepts. Average intercepts differed strongly by both lexical status and lexical complexity. The difference in difficulty due to lexical status shown by the intercepts was about the same for both types of single-syllable words (CVC and words with consonant clusters). For bisyllabic patterns, the lexicality effect was smaller, and for polysyllabic patterns, it was smallest. The average intercepts decreased by increasing orthographic complexity, both for reading real Dutch words and for reading pseudowords, which is in line with findings by Caravolas (2018). The pattern of development (linear change, and negative acceleration) was largely the same for the four levels of orthographic complexity of both words and pseudowords. The learning curves that initially increase exponentially and later level off are similar to logistic learning functions featured in neural networks posited for the reading process under conditions of supervised learning (cf. Seidenberg, 2017).

It is important to note that the results of the present study are fully commensurate with neural findings. To begin with, they are in line with the finding that print tuning starts with an indexation of visual expertise for print as evidenced by an early N1 component in the visual event-related potential which is followed by an inverted U-shape development across the grades (Maurer et al., 2011). It can be assumed that the development of decoding accuracy in our study can be seen as the product of this emergence of visual expertise in the early grades. Furthermore, in line with the postulation of a hierarchical visual word form system being tuned for print and containing local combination detectors with sensitivity to larger fragments of orthographic representations resulting in a posterior-to-anterior gradient of word form specificity (Dehaene et al., 2005), we found the development of decoding in middle and higher grades to change from inefficient (inaccurate, slow) to efficient (accurate, fast) in both typical and poor readers. Finally, in line with neural evidence for the impact of orthographic complexity (Borleffs et al., 2019) and lexicality (Maurer et al., 2006; Weiss and Booth, 2017), we found children development of accuracy and efficiency of decoding to be a function of orthographic complexity and lexicality.

Developmental changes in print tuning

With respect to the development of print tuning, there was clear evidence that decoding accuracy preceded decoding efficiency. A high level of (pseudo)word decoding accuracy appeared necessary before decoding efficiency could develop. It was shown that students reaching an 80% accuracy threshold more strongly developed decoding efficiency than students scoring below this threshold. The effects of this decoding accuracy threshold on the developmental curves of (pseudo)word decoding efficiency were restricted to the early grade levels, for single-syllabic words grades 1–3, and for words of two or more syllables these effects extended to Grade 4. For the efficiency of reading pseudowords the accuracy effects continued to exist longer, to Grade 4 for CVC pseudowords, to Grade 5 for CC pseudowords, and to Grade 6 for bisyllabic and polysyllabic pseudowords. These data extend previous findings by Juul et al. (2014) and Karageorgos et al. (2020) and provide empirical evidence for the proposition put forth by Wise Younger et al. (2017), namely that decoding automaticity should be conceptualized as a continuum and not a dichotomy. It can thus be assumed that growth in word decoding entails children establishing strong connections between letters and sounds for a growing variety of (pseudo)words but also frequent retrieval of word forms, which fosters increased reading fluency and automaticity of word decoding. With this development and practice, children thus proceed from partially specified to fully specified representations of written words with the strength of the association between print and sound (or sound and print, for that matter) becoming increasingly automated. And it can be assumed that with sufficient reading practice, words may become perceptual representations which are recognized by sight and the direct ventral route to word decoding without the need for the dorsal route of letter-sound conversion (Coltheart et al., 2001; Das et al., 2011).

Effects of lexicality

Regarding the second research question, we studied the developmental relations between word decoding efficiency and pseudoword decoding efficiency at each of the four levels of orthographic complexity, using bivariate latent change score analysis. It was shown that the development of word decoding efficiency was strongly associated with the level of pseudoword decoding efficiency. The effects of pseudoword decoding efficiency on growth in word decoding efficiency were strongest in the early grades, especially for the higher levels of orthographic complexity. For CVC and CC words the level of pseudoword decoding efficiency was associated with increase in word decoding efficiency in all grades up to Grade 5. But for the longer words (bisyllabic and polysyllabic), there was no significant effect of pseudoword decoding on word decoding efficiency after Grade 4. It can tentatively be concluded that the capacity of decoding pseudowords that children develop as a self-teaching mechanism in the early grades helps them to store and retrieve orthographic word representations in subsequent grades (Share, 2004). As was shown by Takashima et al. (2014) sublexical parts of pseudowords can be stored and retrieved in memory during orthographic learning. Interestingly, effects were also noted the other way around, indicating that preceding levels of word decoding efficiency contributed to the amount of change in pseudoword decoding efficiency. For CVC words this was evidenced only in the early grades (1–3), while the growth of complex pseudoword decoding efficiency was found to be dependent on the level of word decoding efficiency achieved in the previous school year. It can be assumed that in decoding orthographically complex words also constituent parts such as onsets, rimes and syllables become stored in memory and may thus help children to become better in pseudoword decoding. This finding is fully commensurate with the finding by Pugh et al. (2013) that phonological and orthographic processing may lead to bidirectional connectivity patterns in the beginning reader's brain.

Differentiation of print tuning problems

Regarding our third research question, we conducted Latent Profile Analyses to search for student profiles of accuracy and efficiency decoding and word decoding performances in the data in the first two grades. In grade 1, we identified five latent profiles based on fit statistics. The profiles varied by accuracy on shorter versus longer words and by lexicality differences in accuracy followed by similar differences in decoding efficiency. The first profile represented students with low scores on accuracy and efficiency, especially for the longer words. Students with the second profile scored on average on decoding accuracy for monosyllabic words and pseudowords but low for multisyllabic words and pseudowords. Students with the third profile scored on average on accuracy but stayed behind in pseudoword decoding efficiency. Students with the fourth profile were above average on decoding accuracy, also relatively accurate with the long pseudowords, though still less than with the shorter words. In students with the fifth profile, all accuracy and efficiency scores were high. In second grade, four latent profiles were identified. The first profile referred to students with both word and pseudoword decoding accuracy and efficiency far below average, even more so for the longer (pseudo)words. Students with the second profile showed decoding accuracy just below the means with decoding efficiency staying behind. Students in the third profile showed relatively high scores on word and pseudoword decoding accuracy and efficiency scores just above average, whereas the students in the fourth profile showed high accuracy and efficiency scores across all decoding measures. It can thus be concluded that in the early primary grades, there are students having a hard time in becoming fully accurate in decoding both pseudowords and words (see Castles, 2006). There is also evidence that accuracy precedes efficiency across these profiles and that decoding problems are a function of orthographic complexity and lexicality. They are greater in words with complex syllables than in CVC words and in polysyllabic words as compared to monosyllabic words. This is in line with previous findings by De Luca et al. (2002), Zoccolotti et al. (2005), and Caravolas (2018).

Implications

The results of the present study make it clear that print tuning can be explained in terms of a single associative network, and that its development departs from relatively simple toward more complex structures. As has also been emphasized by Verhoeven and Perfetti (2017), the transitions during the process of learning to read may often reflect the adoption of increasingly sophisticated sublexical decoding strategies such as the search for units already available within a phonological domain (e.g., rimes, syllables, and morphemes). It can be further assumed that with continued reading instruction and practice, children learn to apply such strategies more proficiently and thereby extend their decoding abilities. Along these lines, it can be argued that – across orthographies – children learning to read need to overcome the consistency problem reflecting the fact that orthographic units may have multiple pronunciations, and the granularity problem reflecting the fact that the efficiency of using smaller grain sizes may facilitate the decoding process of more complex orthographic patterns (see Ziegler and Goswami, 2005). Thanks to the (re)structuring and increasing awareness of the phonological infrastructure of spoken language, and because of a learned specialization to recognize and extend orthographic codes, visual word forms are stored in memory which increase in number, specificity and redundancy through reading exposure (see Verhoeven and Perfetti, 2022). Thanks to continuous print tuning, written words can become familiar perceptual objects that are then recognized more quickly. Highly fluent word reading results into an effortless perceptual response that can include the automatization of word decoding, familiarity-based memory retrieval, and the attainment of fluent skilled reading (Dehaene et al., 2015).

The results of the present study have important practical implications. In general, the monitoring and promotion of children's word and pseudoword decoding skills throughout elementary school appears to be of utmost importance. Word and pseudoword reading can be considered related abilities fundamental to reading development in alphabetic orthographies. Word decoding assessment may index children's orthographic representations of words, which are strengthened by the underlying "self-teaching mechanism" of alphabetic pseudoword decoding. Given the strong relationships between decoding skills over the grades, a strong focus on word decoding in the early grades can be emphasized. This can be accomplished by designing kindergarten instruction to provide practice with the sound structures of words, recognition and writing of letters, and an understanding of the alphabetic principle.

Children's attention should be directed to the phonological structure of their language and to the connections between phonemes and spellings. Initial reading instruction should focus on the sublexical structure of words given the nature of the orthographic system in question. Explicit instruction and practice should be arranged to help children sounding out written words, uncovering the orthographic representations of new words, and identifying words primarily via the recoding of constituent grapheme-phoneme relationships (Share, 2004). Our data on children with decoding deficits suggest that they may make up for their lack in word decoding. Therefore, it is of utmost importance to identify poor readers, including children with developmental dyslexia, as early as possible and to combine phonological awareness and reading accuracy training in early intervention and reading efficiency training in follow-up interventions (see Snowling and Hulme, 2012). In addition, they should be given abundant opportunities to read to achieve fluency. It is only by providing access to a wide range of well-written and graded text materials that children can make the transition from the slow cognitively demanding attack of individual words to the effortless identification of words in context (cf. Castles et al., 2018).

To conclude, the development of print tuning in a relatively transparent orthography involves the adoption of a self-teaching mechanism that allows children to familiarize with a script to become competent in the correct phonological recoding of novel words or pseudowords. Every encounter with a real word may lead to a phonological recoding of that word which may then be fed back to the orthographic representation of the word in memory as the initial step of word-specific word identification. The complexity of turning the unfamiliar word form in a familiar orthographic representation is found to be dependent on orthographic complexity as indicated by word length and syllabic complexity. Via ongoing reading exposure, written words may become perceptual objects that can be recognized accurately and with growing speed. Learning to read thus builds on an increased ability of reading pseudowords and a growing capacity of storing and retrieving orthographic representations from memory. Children start out by becoming fully accurate in decoding after which they may become efficient in word decoding as is needed in order to be able to comprehend written text. In the case of reading problems, children must step-by-step learn to become accurate and fast in phonological recoding and in the retrieval of words from memory.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, to any qualified researcher.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

LV initiated the research design and data collection, and was the overall main author and leading author of the introduction and discussion sections. MV coordinated the data analysis and was the overall second author and leading author of the results section. JK contributed to the design and data analysis and was the overall third author. All authors contributed to the article and approved the submitted version.

Conflict of interest

JK was employed by Cito, The Netherlands, but affiliated with Stichting Cito (Foundation Cito), the not-for-profit part of Cito which is dedicated to applied scientific research into educational measurement.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Balass, M., Nelson, J. R., and Perfetti, C. A. (2000). Word learning: an ERP investigation of word experience effects on recognition and word processing. *Contemp. Educ. Psychol.* 35, 126–140.

Burnham, K. P., and Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociol. Methods Res.* 33, 261–304. doi: 10.1177/0049124104268644

Borleffs, E., Maassen, B. A. M., Lyytinen, H., and Zwarts, F. (2017). Measuring orthographic transparency and morphological-syllabic complexity in alphabetic orthographies: A narrative review. *Read. Writ.* 30, 1617–1638. doi: 10.1007/s11145-017-9741-5

Borleffs, E., Maassen, B. A. M., Lyytinen, H., and Zwarts, F. (2019). Cracking the code: The impact of orthographic transparency and morphological-syllabic complexity on reading and developmental dyslexia. *Front. Psychol.* 9:2534. doi: 10.3389/fpsyg.2018.02534

Caravolas, M. (2018). Growth of word and pseudoword reading efficiency in alphabetic orthographies: impact of consistency. *J. Learn. Disabil.* 51, 422–433. doi: 10.1177/0022219417718197

Castles, A. (2006). The dual route model and the developmental dyslexias. *Lond. Rev. Educ.* 4, 49–61. doi: 10.1080/13603110600574454

Castles, A., and Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition* 91, 77–111. doi: 10.1016/S0010-0277(03)00164-1

Castles, A., Rastle, K., and Nation, K. (2018). Ending the reading wars: reading acquisition from novice to expert. *Psychol. Sci. Public Interest* 19, 5–51. doi: 10.1177/1529100618772271

Christodoulou, J. A., Del Tufo, S. N., Lymberis, J., Saxler, P. K., Ghosh, S. S., Triantafyllou, C., et al.(2014). Brain bases of reading fluency in typical reading and impaired fluency in dyslexia. *PLoS One* 9:e100552. doi: 10.1371/journal.pone.0100552

Chyl, K., Fraga-González, G., Brem, S., and Jednoróg, K. (2021). Brain dynamics of (a)typical reading development–A review of longitudinal studies. *NPJ Sci. Learn.* 6:1. doi: 10.1038/s41539-020-00081-5

Cohen, L., and Dehaene, S. (2009). "Ventral and dorsal contribution to word reading," in *Cognitive Neuroscience*, ed. M. S. Gazzaniga (Cambridge, MA: MIT Press), 789–804.

Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M., et al. (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.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., and Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychol. Rev.* 108, 204–256. doi: 10.1037/0033-295X.108.1.204

Das, T., Padakannaya, P., Pugh, K. R., and Singh, N. C. (2011). Neuroimaging reveals dual routes to reading in simultaneous proficient readers of two orthographies. *NeuroImage* 54, 1476–1487. doi: 10.1016/j.neuroimage.2010.09.022

De Luca, M., Borelli, M., Judica, A., Spinelli, D., and Zoccolotti, P. (2002). Reading words and pseudowords: an eye movement study of developmental dyslexia. *Brain Lang.* 80, 617–626. doi: 10.1006/brln.2001.2637

De Luca, M., Di Pace, E., Judica, A., Spinelli, D., and Zoccolotti, P. (1999). Eye movement patterns in linguistic and non-linguistic tasks in developmental surface dyslexia. *Neuropsychologia* 37, 1407–1420. doi: 10.1016/S0028-3932(99)00038-X

Dehaene, S., and Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron* 22, 384–398. doi: 10.1016/j.neuron.2007.10.004

Dehaene, S., Cohen, L., Morais, J., and Kolinsky, R. (2015). Illiterate to literate: behavioural and cerebral changes induced by reading acquisition. *Nat. Rev. Neurosci.* 16, 234–244. doi: 10.1038/nrn3924

Dehaene, S., Cohen, L., Sigman, M., and Vinckier, F. (2005). The neural code for written words: a proposal. *Trends Cogn. Sci.* 9, 335–341. doi: 10.1016/j.tics.2005.05.004

Eberhard-Moscicka, A. K., Jost, L. B., Raith, M., and Maurer, U. (2015). Neurocognitive mechanisms of learning to read: print tuning in beginning readers related to word-reading fluency and semantics but not phonology. *Dev. Sci.* 18, 106–118. doi: 10.1111/desc.12189

Ferguson, S. L., Moore, E. W. G., and Hull, D. M. (2020). Finding latent groups in observed data: a primer on latent profile analysis in Mplus for applied researchers. *Int. J. Behav. Dev.* 44, 458–468. doi: 10.1177/0165025419881721

Gniewosz, B., and Gniewosz, G. (2018). How to model intraindividual change in cohort data using Mplus' multi-group approach. *Int. J. Behav. Dev.* 42, 373–380. doi: 10.1177/0165025417750282

Graves, W. W., Desai, R., Humphries, C., Seidenberg, M. S., and Binder, J. R. (2010). Neural systems for reading aloud: a multiparametric approach. *Cerebr. Cortex* 20, 1799–1815. doi: 10.1093/cercor/bhp245

Juphard, A., Carbonnel, S., and Valdois, S. (2004). Length effect in reading and lexical decision: evidence from skilled readers and a developmental dyslexic participant. *Brain Cogn.* 55, 332–340. doi: 10.1016/j.bandc.2004.02.035

Juul, H., Poulsen, M., and Elbro, C. (2014). Separating speed from accuracy in beginning reading development. *J. Educ. Psychol.* 106, 1096–1106. doi: 10.1037/a0037100

Karageorgos, P., Müller, B., and Richter, T. (2019). Modelling the relationship of accurate and fluent word recognition in primary school. *Learn. Individ. Diff.* 76:101779. doi: 10.1016/j.lindif.2019.101779

Karageorgos, P., Richter, T., Haffmans, M.-B., Schindler, J., and Naumann, J. (2020). The role of word-recognition accuracy in the development of word-recognition speed and reading comprehension in primary school: a longitudinal examination. *Cogn. Dev.* 56:100949. doi: 10.1016/j.cogdev.2020.100949

Lemons, C. J., Key, A. P., Fuchs, D., Yoder, P. J., Fuchs, L. S., Compton, D. L., et al. (2010). Predicting reading growth with event-related potentials: thinking differently about indexing "responsiveness". *Learn. Individ. Diff.* 20, 158–166. doi: 10.1016/j.lindif.2009.11.006

Lo, Y., Mendell, N. R., and Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika* 88, 767–778. doi: 10.1093/biomet/88.3.767

Masyn, K. E. (2013). "Latent class analysis and finite mixture modeling," in *The Oxford Handbook of Quantitative Methods*, ed. T. D. Little (New York, NY: Oxford University Press), 551–611.

Maurer, U., Brandeis, D., and McCandliss, B. (2005a). Fast, visual specialization for reading in English revealed by the topography of the N170 ERP response. *Behav. Brain Funct.* 1:13. doi: 10.1186/1744-9081-1-13

Maurer, U., Brem, S., Bucher, K., and Brandeis, D. (2005b). Emerging neurophysiological specialization for letter strings. *J. Cogn. Neurosci.* 17, 1532–1552. doi: 10.1162/089892905774597218

Maurer, U., Brem, S., Kranz, F., Bucher, K., Benz, R., Halder, P., et al. (2006). Coarse neural tuning for print peaks when children learn to read. *Neuroimage* 33, 749–758. doi: 10.1016/j.neuroimage.2006.06.025

Maurer, U., Schulz, E., Brem, S., van der Mark, S., Bucher, K., Martin, E., et al. (2011). The development of print tuning in children with dyslexia: evidence from longitudinal ERP data supported by fMRI. *Neuroimage* 57, 714–722. doi: 10.1016/j.neuroimage.2010.10.055

McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annu. Rev. Psychol.* 60, 577–605. doi: 10.1146/annurev.psych.60.110707.163612

McCandliss, B. D., and Noble, K. G. (2003). The development of reading impairment: a cognitive neuroscience model. *Ment. Retard. Dev. Disabil. Res. Rev.* 9, 196–204. doi: 10.1002/mrdd.10080

Muthén, L. K., and Muthén, B. O. (1998–2012). *Mplus User's Guide*, 7th Edn. Los Angeles, CA: Muthén & Muthén.

Norris, D. (2013). Models of visual word recognition. *Trends Cogn. Sci.* 17, 517–524. doi: 10.1016/j.tics.2013.08.003

Nunn, A. (1998). *Dutch Orthography*. Utrecht: Center for Language Studies.

Perfetti, C. (2007). Reading ability: lexical quality to comprehension. *Sci. Stud. Read.* 11, 357–383. doi: 10.1080/10888430701530730

Pugh, K. R., Landi, N., Preston, J. L., Mencl, W. E., Austin, A. C., Sibley, D., et al. (2013). The relationship between phonological and auditory processing and brain organization in beginning readers. *Brain Lang.* 125, 173–183. doi: 10.1016/j.bandl.2012.04.004

Rahbari, N., and Sénéchal, M. (2010). Learning to read and spell in Persian: a cross-sectional study from Grades 1 to 4. *Dev. Psychol.* 46, 1514–1527. doi: 10.1037/a0020377

Rasbash, J., Charlton, C., Browne, W. J., Healy, M., and Cameron, B. (2016). *MLwiN Version 2.36*. Bristol: Centre for Multilevel Modelling, University of Bristol.

Rau, A. K., Moll, K., Snowling, M. J., and Landerl, K. (2015). Effects of orthographic consistency on eye movement behavior: Germanand english children and adults process the same words differently. *J. Exp. Child Psychol.* 130, 92–105. doi: 10.1016/j.jecp.2014.09.012

Reichle, E. D., Liversedge, S. P., Drieghe, D., Blythe, H. I., Joseph, H. S., White, S. J., et al. (2013). Using E-Z Reader to examine the concurrent development of eye-movement control and reading skill. *Dev. Rev.* 33, 110–149. doi: 10.1016/j.dr.2013.03.001

Reynolds, C. R., Altmann, R. A., and Allen, D. N. (2021). *Mastering Modern Psychological Testing: Theory and Methods*, 2nd Edn. Berlin: Springer.

Richlan, F., Kronbichler, M., and Wimmer, H. (2013). Structural abnormalities in the dyslexic brain: a meta-analysis of voxel-based morphometry studies. *Hum. Brain Mapp.* 34, 3055–3065. doi: 10.1002/hbm.22127

Seidenberg, M. (2017). *Language at the Speed of Sight: How We Read, Why So Many Can't, and What Can Be Done About It*. New York, NY: Basic Books.

Sereno, S. C., and Rayner, K. (2003). Measuring word recognition in reading: eye movements and event-related potentials. *Trends Cogn. Sci.* 7, 489–493. doi: 10.1016/j.tics.2003.09.010

Seymour, P. H. K., Aro, M., and Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *Br. J. Psychol.* 94, 143–174. doi: 10.1348/000712603321661859

Share, D. L. (2004). Orthographic learning at a glance: on the time course and developmental onset of self-teaching. *J. Exp. Child Psychol.* 87, 267–298. doi: 10.1016/j.jecp.2004.01.001

Snowling, M. J., and Hulme, C. (2012). Interventions for children's language and literacy difficulties. *Int. J. Lang. Commun. Disord.* 47, 27–34. doi: 10.1111/j.1460-6984.2011.00081.x

Spinelli, D., De Luca, M., Di Filippo, G., Mancini, M., Martelli, M., and Zoccolotti, P. (2005). Length effect in word naming in reading: role of reading experience and reading deficit in Italian readers. *Dev. Neuropsychol.* 27, 217–235. doi: 10.1207/s15326942dn2702_2

Takashima, A., Wagensveld, B., van Turennout, M., Zwitserlood, P., Hagoort, P., and Verhoeven, L. (2014). Training-induced neural plasticity in visual-word decoding and the role of syllables. *Neuropsychologia* 61, 299–314. doi: 10.1016/j.neuropsychologia.2014.06.017

Taylor, J. S. H., Duff, F. J., Woollams, A. M., Monaghan, P., and Ricketts, J. (2015). How word meaning influences word reading. *Curr. Direct. Psychol. Sci.* 24, 322–328. doi: 10.1177/0963721415574980

Tiffin-Richards, S., and Schroeder, S. (2015). Word length and frequency effects on children's eye movements during silent reading. *Vis. Res.* 113, 33–43. doi: 10.1016/j.visres.2015.05.008

van den Boer, M., and de Jong, P. F. (2015). Parallel and serial reading processes in children's word and nonword reading. *J. Educ. Psychol.* 107, 141–151. doi: 10.1037/a0037101

Verhoeven, L. (2017). "Learning to read Dutch," in *Learning to Read across Languages and Writing Systems*, eds L. Verhoeven and C. Perfetti (Cambridge, MA: Cambridge University Press), 323–346.

Verhoeven, L., and Keuning, J. (2018). The nature of developmental dyslexia in a transparent orthography. *Sci. Stud. Read.* 22, 7–23. doi: 10.1080/10888438.2017.1317780

Verhoeven, L., Keuning, J., Horsels, L., and van Boxtel, H. (2013). *Wetenschappelijke Verantwoording Testinstrumentarium Taalontwikkelingsstoornissen (Scientific Validation of Test for Language Impairment)*. Arnhem: Cito.

Verhoeven, L., and Perfetti, C. (eds) (2017). *Learning to Read across Languages and Writing Systems*. Cambridge, MA: Cambridge University Press.

Verhoeven, L., and Perfetti, C. (2022). Universals in learning to read across languages and writing systems. *Sci. Stud. Read.* 26, 150–164. doi: 10.1080/10888438.2021.1938575

Verhoeven, L., Perfetti, C., and Pugh, K. (2019). *Developmental Dyslexia across Languages and Writing Systems*. Cambridge, MA: Cambridge University Press.

Verhoeven, L., and van Leeuwe, J. (2009). Modeling the growth of word decoding skills: evidence from Dutch. *Sci. Stud. Read.* 13, 205–223. doi: 10.1080/10888430902851356

Weiss, Y., and Booth, J. R. (2017). Neural correlates of the lexicality effect in children. *Brain Lang.* 175:64–70. doi: 10.1016/j.bandl.2017.09.006

Wise Younger, J., Tucker-Drob, E., and Booth, J. (2017). Longitudinal changes in reading network connectivity related to skill improvement. *NeuroImage* 158, 90–98. doi: 10.1016/j.neuroimage.2017.06.044

Zhang, Z., and Peng, P. (2021). Reading real words vs pseudowords: a meta-analysis of research in developmental dyslexia. *Dev. Psychol.* doi: 10.1037/dev0001340 [Epub ahead of print].

Ziegler, J. C., and Goswami, U. (2005). Reading acquisition, developmental dyslexia and skilled reading across languages: a psycholinguistic grain size theory. *Psychol. Bull.* 131, 3–29. doi: 10.1037/0033-2909.131.1.3

Ziegler, J. C., Perry, C., and Zorzi, M. (2014). Modelling reading development through phonological decoding and self-teaching: implications for dyslexia. *Philos. Trans. R. Soc. B Biol. Sci.* 369:20120397. doi: 10.1098/rstb.2012.0397

Zoccolotti, P., De Luca, M., Di Pace, E., Gasperini, F., Judica, A., and Spinelli, D. (2005). Word length effect in early reading and in developmental dyslexia. *Brain Lang.* 93, 369–373. doi: 10.1016/j.bandl.2004.10.010

frontiers | Frontiers in Neuroscience

TYPE Original Research
PUBLISHED 29 September 2022
DOI 10.3389/fnins.2022.920150

OPEN ACCESS

EDITED BY
Kazuhiko Sawada,
Tsukuba International University, Japan

REVIEWED BY
Elizabeth Huber,
University of Washington, United States
Sabrina Turker,
Max Planck Institute for Human
Cognitive and Brain Sciences,
Germany

*CORRESPONDENCE
Johanna Liebig
johanna.liebig@fu-berlin.de

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to
Neurodevelopment,
a section of the journal
Frontiers in Neuroscience

RECEIVED 14 April 2022
ACCEPTED 18 August 2022
PUBLISHED 29 September 2022

CITATION
Beyer M, Liebig J, Sylvester T, Braun M, Heekeren HR, Froehlich E, Jacobs AM and Ziegler JC (2022) Structural gray matter features and behavioral preliterate skills predict future literacy – A machine learning approach.
Front. Neurosci. 16:920150.
doi: 10.3389/fnins.2022.920150

Structural gray matter features and behavioral preliterate skills predict future literacy – A machine learning approach

Moana Beyer[1†], Johanna Liebig[1,2*†], Teresa Sylvester[1,2], Mario Braun[3], Hauke R. Heekeren[2,4], Eva Froehlich[5], Arthur M. Jacobs[1,2] and Johannes C. Ziegler[6]

[1]Department of Education and Psychology, Freie Universität Berlin, Berlin, Germany, [2]Center for Cognitive Neuroscience Berlin, Freie Universität Berlin, Berlin, Germany, [3]Centre for Cognitive Neuroscience, Universität Salzburg, Salzburg, Austria, [4]Department of Biological Psychology and Cognitive Neuroscience, Freie Universität Berlin, Berlin, Germany, [5]Department of Decision Neuroscience and Nutrition, German Institute of Human Nutrition Potsdam-Rehbrücke, Nuthetal, Germany, [6]Laboratoire de Psychologie Cognitive, Aix-Marseille Université and Centre National de la Recherche Scientifique, Marseille, France

When children learn to read, their neural system undergoes major changes to become responsive to print. There seem to be nuanced interindividual differences in the neurostructural anatomy of regions that later become integral parts of the reading network. These differences might affect literacy acquisition and, in some cases, might result in developmental disorders like dyslexia. Consequently, the main objective of this longitudinal study was to investigate those interindividual differences in gray matter morphology that might facilitate or hamper future reading acquisition. We used a machine learning approach to examine to what extent gray matter macrostructural features and cognitive-linguistic skills measured before formal literacy teaching could predict literacy 2 years later. Forty-two native German-speaking children underwent T1-weighted magnetic resonance imaging and psychometric testing at the end of kindergarten. They were tested again 2 years later to assess their literacy skills. A leave-one-out cross-validated machine-learning regression approach was applied to identify the best predictors of future literacy based on cognitive-linguistic preliterate behavioral skills and cortical measures in *a priori* selected areas of the future reading network. With surprisingly high accuracy, future literacy was predicted, predominantly based on gray matter volume in the left occipito-temporal cortex and local gyrification in the left insular, inferior frontal, and supramarginal gyri. Furthermore, phonological awareness significantly predicted future literacy. In sum, the results indicate that the brain morphology of the large-scale reading network at a preliterate age can predict how well children learn to read.

KEYWORDS

MRI, machine learning, reading acquisition, precursors, prediction, children

Introduction

Literacy is a multidimensional concept that comprises the ability to read (derive meaning from written symbols) and write (encode meaning through written symbols). Literacy is a key competence in today's information-driven society. Not surprisingly, difficulties in reading can have far-reaching consequences ranging from low academic achievement (Bruck, 1987; Fletcher and Vaughn, 2009) to emotional problems or even psychiatric disorders (Schulte-Körne, 2010; Livingston et al., 2018). To better understand why some children fail to acquire literacy skills successfully, there has been quite some effort in identifying factors that potentially facilitate or hamper reading acquisition (Hoeft et al., 2007; Perry et al., 2019). This has led to the identification of several precursors of literacy that predict future reading and spelling at the end of kindergarten.

To date, the most robust predictors of literacy prior to formal instruction at school are cognitive-linguistic preliterate skills. The two most reliable skills are rapid automatized naming (RAN) and phonological awareness (PA). RAN tasks assess a child's speed and accuracy in naming familiar stimuli such as digits, letters, and colors. RAN and fluent reading share many subprocesses, such as saccadic eye movements, lexical access, and the mapping of visual objects onto language representations (Norton and Wolf, 2012). PA refers to the ability to represent, recognize, access, and manipulate any phonological unit within a word. Thus, PA is essential to map orthography onto phonology and hence bootstrap reading acquisition (Ziegler et al., 2014, 2020). The strong link between these two variables and reading acquisition has been repeatedly shown in large-scale cross-linguistic studies both at the concurrent (Ziegler et al., 2010; Landerl et al., 2013) and the longitudinal level (Caravolas et al., 2012; Landerl et al., 2019, 2022). In recent years, there has also been an increasing effort to study neurofunctional (Lohvansuu et al., 2018; Liebig et al., 2020, 2021) predictors of future reading proficiency before the onset of literacy (see Chyl et al., 2021 for a recent review). The overlap of the functional and anatomical neural architecture of reading suggests a close link between brain morphology and function.

Consequently, multiple studies have examined reading-related macrostructural features of the cortex (Linkersdörfer et al., 2012; Richlan et al., 2013; Eckert et al., 2016). However, the results of different studies are far from converging into a uniform picture (Ramus et al., 2018; Chyl et al., 2021). Thus, in the present study, we aimed to compare and weigh the effects of several gray matter macrostructural brain maturation features and behavioral measures of cognitive-linguistic preliterate skills, which were gathered at the end of kindergarten to predict individual literacy skills 2 years later.

A large-scale reading network has been identified in skilled adult readers that can roughly be characterized by two posterior and one anterior stream. The ventral stream is linked to the occipito-temporal cortex and hosts the visual word form area. It is associated with direct orthographic reading strategies (Cohen and Dehaene, 2004; Dehaene and Cohen, 2011). The dorsal stream, located in the temporo-parietal cortex, is primarily devoted to auditory-phonological recoding (Pugh et al., 2000). Both streams converge in the frontal stream linked to the inferior frontal gyrus and insular cortex, among others (Binder et al., 2009; Price, 2012; Martin et al., 2015; Froehlich et al., 2018). However, in the last decades, the modularized view of reading has been increasingly challenged and replaced by a unified view of reading (Price and Devlin, 2011; Braun et al., 2019). According to this view, reading is orchestrated by the large-scale network in a highly distributed and interactive way (Hofmann and Jacobs, 2014; Ziegler et al., 2018). This partly pre-existing network already devoted to language and sensory-motor processing must be fundamentally reorganized during reading acquisition to become responsive to print (Dehaene et al., 2015; Liebig et al., 2017; Dehaene-Lambertz et al., 2018).

Until today, most work examining reading-related brain morphology has focused on altered gray matter volume, hereafter referred to as cortical volume, in impaired compared to neurotypical readers (Ramus et al., 2018). The underlying reason might be the cortical morphology's potential to be a promising early biomarker of future literacy as it is primarily determined by neurodevelopmental processes *in utero* and is partly confined by genetic heritability (Gilmore et al., 2018). Yet, brain structure undergoes continuing changes that are highly intertwined with changes in cognitive abilities, resulting in interindividual variability (Raznahan et al., 2011; Frangou et al., 2022) and might thus be very well suited to identify subtle differences in the cortical morphology that will affect future for reading acquisition.

However, the results of three coordinate-based meta-analyses showed little consistency across studies (Linkersdörfer et al., 2012; Richlan et al., 2013; Eckert et al., 2016). Nonetheless, they point to decreased cortical volume in all three reading streams. More specifically, bilateral temporo-parietal, left ventral occipito-temporal, left frontal, and bilateral cerebellar regions show volumetric differences in developmental dyslexia. A handful of studies focused on the cortical folding pattern in relation to reading acquisition and dyslexia. Impaired readers seem to exhibit abnormal gyrification in the left occipito-temporal and temporo-parietal cortices (Im et al., 2016; Williams et al., 2018), i.e., the ventral and the dorsal stream. Focusing on single structural features generally bears the risk of overseeing interactions between different anatomical measures. Therefore, Płoński et al. (2017) tested several macrostructural features using a cross-validated (CV) classification algorithm in a large cross-linguistic sample of 8- to 13-year-old children and adolescents. Children with dyslexia displayed increased folding and curvature in left temporo-parietal regions, lower surface area in the prefrontal cortex, i.e., in all three reading streams. This comprehensive analysis revealed the benefit of machine learning approaches and the combination of neuroanatomical measures to identify cortical differences more closely and with greater specificity.

In summary, children and adults with impaired reading show a pattern of decreased cortical volume and surface area paired with abnormal gyrification in diverse regions of the large-scale reading network. The neuroanatomical differences seemed to precede the onset of reading instruction at school. Prereaders later diagnosed with dyslexia showed reduced surface area in all three future reading streams, i.e., the bilateral fusiform gyri (Beelen et al., 2019), the left supramarginal gyrus, and the left inferior frontal gyrus (Hosseini et al., 2013). In contrast, cortical thickness did not differ between groups (Hosseini et al., 2013; Beelen et al., 2019). However, neuroanatomical differences in all three reading streams have not been comprehensively investigated yet. In a longitudinal study examining native German-speaking children, Kuhl et al. (2020) reported mixed results concerning the relationship between macrostructural features and reading. Only abnormal gyrification in the left auditory cortex dissociated preliterate children who developed dyslexia from their typically developing peers. Although a uniform picture is still missing, these studies provide a crucial foundation for characterizing the neural basis of reading difficulties.

All of the above-described studies compared typical and impaired reading acquisition. However, group contrasts can be problematic because there is no consistent definition of impaired reading or dyslexia, and thus different methods and thresholds are utilized to classify children across studies (Francis et al., 2005; Fletcher, 2009). Furthermore, reading performance is a continuous variable and splitting the sample into two categorical groups loses valuable information (Button et al., 2013). A different approach is to directly target the relationship between anatomy and reading in typically developing children using continuous sampling and spanning the entire range of reading proficiency. However, studies investigating macrostructural features of typical reading acquisition are scarce and yield mixed results. Longitudinal studies in emergent to intermediate readers show that decreases in cortical volume in different regions linked to the dorsal and frontal stream (e.g., left inferior parietal cortex, superior temporal gyrus, and precentral gyrus) correlate positively with literacy skills (Houston et al., 2014; Linkersdörfer et al., 2014; Jednoróg et al., 2015). These results indicate that an age-appropriate maturation of the large-scale reading network facilitates reading acquisition from early on. In contrast, Torre and Eden (2019) did not find any relationships between cortical volume and word reading in pre-defined regions of the reading network, neither in a large sample of 404 typical readers (6- to 22-year-old) nor in a subsample of 6- to 9-year-old children.

Similarly, Perdue et al. (2020) reported mixed results. They identified a positive relationship between cortical thickness in the left superior temporal gyrus and word and pseudoword reading in typically developing children (4- to 9-year-old) but did not find a relationship between reading skills and surface area in their whole-brain based analysis. These results were supported by a study that tested Chinese-speaking children and showed that word reading was positively correlated with cortical thickness in bilateral superior temporal gyri, the left inferior temporal gyrus, and the left supramarginal gyrus (Xia et al., 2018).

In all of these studies, neuroanatomical features were correlated with literacy-related skills. Correlational frameworks, however, do not allow for generalization to unseen individuals (Dubois and Adolphs, 2016). Furthermore, correlational approaches with a small sample size are prone to over-fitting both signal and noise (Vul et al., 2009; Dubois and Adolphs, 2016; Elliott et al., 2020; Sui et al., 2020). This limitation can be tackled by using CV methods, in which a training sample predicts performance in an independent data set. Until today, there are only a few landmark studies using CV methods to predict literacy-related skills in continuous samples. For example, Skeide et al. (2016) used a whole-brain kernel ridge regression to test individual differences in cortical volume in several reading-related regions to predict reading speed in 5- to 12-year-old native German-speaking children. Bilateral middle frontal gyri, the left superior temporal gyrus, and the left occipito-temporal cortex were positively associated with reading skills. In addition, clusters in the visual word form area and the left visual cortex were negatively associated with reading speed (Skeide et al., 2016). Thus, crucial regions of the dorsal and ventral stream predicted reading with high precision. Choosing a similar approach (Cui et al., 2018), used large datasets of the Human Connectome Project to predict individual reading comprehension and decoding skills in young adults. More specifically, they performed an elastic net penalized linear regression to predict individual literacy scores based on whole-brain cortical volume. The most critical predictive clusters were located in frontal and subcortical regions. The generalizability of the prediction model was then tested in an independent sample of Chinese children (8- to 13-years-old) with mixed results.

One could summarize the mixed results of the relationship between the macrostructural features and typical reading development in the following way: Firstly, all studies that found a significant relationship between macrostructural brain measures and reading report positive relationships. Secondly, cortical volume, particularly in the temporo-parietal areas (Houston et al., 2014; Linkersdörfer et al., 2014), i.e., in the ventral and dorsal stream, is not only robustly correlated with but also predicts reading performance in children (but see, Torre and Eden, 2019, for an exception). Cortical thickness in all three reading streams, i.e., superior temporal gyrus, supramarginal gyrus, and inferior frontal gyrus, showed a positive relationship in typical reading children (Xia et al., 2018; Perdue et al., 2020). Until today, there is only one study testing the effect of the surface area on typical reading acquisition, which did not find any significant correlation (Perdue et al., 2020). However, replication is still outstanding. Thus, the goal of the present study was to add new evidence to the still preliminary data on the relationship between macrostructural features and reading acquisition. More specifically, we aimed at predicting future literacy in a sample of German-speaking kindergarten children using continuous sampling and CV prediction modeling. Several aspects distinguish the present research from the two studies summarized above. Firstly, we analyzed longitudinal data to predict the literacy skills of kindergarten children 2 years later instead of examining concurrent brain-behavior relationships (see Ramus et al., 2018, for promises and pitfalls). Secondly, we included several macrostructural features and compared their relative importance. Thirdly, we used both structural and behavioral information (RAN and PA tested at the end of kindergarten) in the CV models to investigate whether or not, and if so, to what extent adding neuroanatomic data would improve the prediction of future literacy over and above cognitive-linguistic preliterate skills. For that, we obtained cortical volume, surface area, and local gyrification (lGI) from structural scans of preliterate children at the end of kindergarten. The reasoning behind our choice was the following: Firstly, we included cortical volume as one of the most widely tested cortical measures to show the validity of our data and see if we could replicate the well-established pattern using a CV-algorithm. Secondly, we aimed to re-test surface area as this feature was so far tested only once in a continuous sample. Thirdly, we added the lGI that was previously only tested in group comparisons (Williams et al., 2018), where it yielded new and promising insights. Thus, we decided to incorporate this relatively new measure into the present analysis.

The cohort of children was then tested again at the end of the second year of primary school to assess their literacy skills. An elastic net regularized regression was applied to predict future literacy ability. The model was based on cognitive-linguistic preliterate skills and anatomical markers of the cortical surface in pre-defined regions of the "future" reading network. To the best of our knowledge, this is the first study to apply a continuous machine-learning approach to predict future literacy abilities.

Based on previous pediatric neuroimaging, we hypothesized that cortical volume, surface area, and lGI in all three (future) reading streams gathered at a preliterate age would predict literacy 2 years later. More specifically, we expected a crucial contribution of the cortical volume of the occipito-temporal (Skeide et al., 2016) and temporo-parietal cortices (Houston et al., 2014; Linkersdörfer et al., 2014; Jednoróg et al., 2015) to the prediction of future reading skills. Our hypothesis regarding the lGI was less specific, as this feature has not yet been tested with a continuous approach. However, referring to the promising results in group-based approaches (Williams et al., 2018), we expected that the gyrification pattern in the occipito-temporal cortex might predict future literacy. Similarly, reduced surface area in the temporo-parietal (Hosseini et al., 2013; Beelen et al., 2019) areas, as well as the frontal cortex (Hosseini et al., 2013; Płoński et al., 2017), has been associated with dyslexia. Thus, we were interested in finding out, if interindividual differences in the surface area of these regions also predict future literacy skills.

Materials and methods

Study participants

Eighty-six German-speaking preliterate children were recruited in their last year of kindergarten on a voluntary basis throughout the city of Berlin. Advertisements in newsletters, kindergartens, and social media platforms were the main recruitment channels. Initial screening ensured that participants had no history of neurological diseases and normal hearing and visual acuity. All participants scored above the 85th percentile on the non-verbal part of the German adaption of the Wechsler Intelligence Scale for Children (WISC-IV; Petermann and Petermann, 2011) tested in the second grade of primary school. Furthermore, children were screened for reading expertise to ensure true preliteracy using a custom-made screening test (see **Supplementary material** and Liebig et al., 2021 for a detailed description). Both parents and children were carefully briefed about the longitudinal study design and the functional magnetic resonance imaging (MRI) constraints. Parents gave written informed consent and received compensation for their travels. All children gave their informed consent to participate in the study and were given age-appropriate education gifts. The Ethics Committee of the German Association for Psychology (DGPs) approved the study.

At the first appointment, nine children refused to participate in the MRI training session (mock-scanner) and were thus excluded from the study. All children who successfully participated in the actual functional MRI session were reinvited 2 years later at the end of the second grade. Ten participants could not be reinvited for the second appointment, and two children had to repeat the second grade and were tested 1 year later. Twenty-five children were excluded from the T1-weighted image analysis pipeline due to poor image quality or insufficient cortex reconstruction (discussed in section "T1-weighted imaging analysis"). The final sample consisted of 42 children, as summarized in **Table 1**. Seven of them had at least one first or second-degree relative with diagnosed developmental dyslexia stipulated by a parental questionnaire (Landerl and Moll, 2010).

Psychometric assessment

This study applied an extensive battery of psychometric testing at the two aforementioned developmental time points.

TABLE 1 Demographic and psychometric information of the final pediatric sample before (TP1) and after literacy acquisition (TP2).

Descriptive data	Test	Raw scores (mean ± SD)	Range of raw scores	Percentile ranks (mean ± SD)
Demographic information				
Age at TP1		5.58 ± 0.48	5.01–6.09	
Age at TP2		8.25 ± 0.53	7.41–8.92	
Female/male		24/18		
Family history of dyslexia		7		
Right-handed/left-handed		38/4		
Monolingual/bilingual		37/5		
Non-verbal intelligence at TP1	CPM	23.26 ± 5.37	13–35	
Non-verbal intelligence at TP2	WISC	115.48 ± 12.71	90–147	
Dyslexia at TP2		10		
Literacy precursor abilities (at TP1)				
Rapid naming	BISC	14.81 ± 3.79	5–20	n.a.
Phonological awareness	BISC	36.17 ± 3.41	24–40	n.a.
Literacy abilities (at TP2)				
Reading fluency	SLRT-II	36.76 ± 15.69	6–71	52.95 ± 34.52 ^{lp} 54.94 ± 33.91 ^{hp}
Reading comprehension	ELFE 1-6	54.00 ± 22.71	6–90	49.49 ± 32.14
Spelling accuracy	DERET 1-2 +	17.36 ± 11.56	1–50	40.24 ± 30.74
Literacy ability	SLRT-II, ELFE 1-6, DERET 1-2 +	0.00 ± 0.98	−2.60 to 1.84	n.a.

Dyslexia was defined as performance below the 16th percentile rank of the reference population in either spelling accuracy or real word reading fluency or in both based on Kuhl et al.'s (2020) classification criteria. CPM, Colored Progressive Matrices; WISC, Wechsler Intelligence Scale for Children; BISC, Bielefelder Screening zur Früherkennung von Lese-Rechtschreibschwierigkeiten; SLRT-II, Salzburger Lese- und Rechtschreibtest; ELFE 1-6, Ein Leseverständnistest für Erst- bis Sechstklässler; DERET 1-2+, Deutscher Rechtschreibtest für das erste und zweite Schuljahr; n.a., age-standardized scores are not available for subtests and literacy ability overall; lp, percentile lower boundary; hp, percentile higher boundary.

Only tests used for the analyses of the present paper are described in this section [for a detailed description of all assessments, see Liebig et al. (2021) and the **Supplementary material**).

At the first assessment (TP1), i.e., prior to reading acquisition, PA and RAN were assessed with the *Bielefelder Screening zur Früherkennung von Lese-Rechtschreibschwierigkeiten* (BISC; Jansen, 2002). PA was calculated using several subtests of the BISC: syllable segmentation, rhyme recognition, word synthetization, and sound-to-word comparisons. RAN was operationalized by the time needed to name the color of objects printed in black and white and in an incongruent color. Errors made were sanctioned with a penalty time, i.e., incorrect responses were penalized with a longer reaction time. Non-verbal intelligence was measured using the *Raven's Colored Progressive Matrices* (CPM; Raven and Court, 1998).

At the second assessment (TP2), i.e., after 2 years of schooling, children were tested on reading fluency and accuracy using two subtests of the *Salzburger Lese- und Rechtschreibtest* (SLRT-II) that focused on word and pseudoword reading speed and accuracy (Moll and Landerl, 2010). Reading comprehension was quantified using the *Ein Leseverständnistest für Erst- bis Sechstklässler* (ELFE 1-6; Lenhard and Schneider, 2006). This test captures reading comprehension on three levels with increasing complexity: word comprehension (decoding and synthesis), sentence comprehension (understanding of syntax), and text comprehension (understanding information and drawing conclusions). Spelling accuracy was assessed by writing from dictation using continuous text and gapped sentences using the German spelling test *Deutscher Rechtschreibtest für das Erste und Zweite Schuljahr* (DERET 1-2+; Stock and Schneider, 2008). Descriptive statistics for these psychometric tests are provided in **Table 1**.

T1-weighted magnetic resonance image acquisition

T1-weighted images were acquired at TP1, i.e., at the end of kindergarten. A few days before the actual image acquisition, children had undergone a training session at the Max Planck Institute for Human Development Berlin using a mock scanner to familiarize them with the MRI procedure. The actual MRI session took place at the Centre for Cognitive Neuroscience Berlin (CCNB). Both sessions were adapted for young children. Their heads were cushioned with foam to ensure head stability and comfort, and age-appropriate earplugs were provided to attenuate scanner noise. Whole-brain anatomical images were gathered for each participant on a 3.0 Tesla Magnetom MRI system (Siemens Healthineers, Erlangen, Germany), equipped with a 12-channel head coil (repetition time: 2,000 ms, echo time: 30 ms, flip angle = 70°, 176 sagittal sections, voxel size: 2 mm × 2 mm × 2 mm, and field of view: 256 × 256 voxel matrix). Acquisition of the T1-weighted images followed a brief experiment in the scanner (described in Liebig et al., 2021) and lasted 4.5 min. During this time, a child-friendly video was played.

T1-weighted imaging analysis

First, the T1-weighted images were visually inspected by two independent raters using Freeview 3.0, FreeSurfer's visualization tool (Fischl, 2012), and MANGO 4.1, a multi-image analysis graphical user interface (Lancaster and Martinez, 2006). Additionally, image quality was assessed using the Computational Anatomical Toolbox (CAT) 12, an extension to the statistical parametric mapping (SPM) 12 software (Welcome Department of Cognitive Neurology; Ashburner et al., 2021). Fourteen participants were excluded due to severe ringing and blurring artifacts in the MRI scans caused by head motion. Children with moderate rigid body movement were marked and treated with special care in the subsequent visual inspection step, i.e., after reconstructing the surfaces during preprocessing. The cortices of eleven children were insufficiently reconstructed and excluded from subsequent analyses.

A fully automated pipeline of the FreeSurfer 7.1.1 software package (Fischl, 2012) was utilized to preprocess the T1-weighted MRI scans, which included removal of non-brain tissues, transformation, and intensity normalization, segmentation of white and deep gray matter, correction of topological errors, and reconstruction of the cortical surface. Surface area and cortical volume were extracted from the T1-weighted image. In FreeSurfer, surface area is quantified as the sum of the areas of adjacent triangle faces on the surface mesh, computed in each participant's native space, allowing for individual variations in the area of each triangle. Cortical volume is defined as the amount of gray matter between the gray/white and pial boundary. These features were modeled for each hemisphere separately.

After completing all preprocessing steps, the segmentation of each participant's cortex was visually inspected in Freeview 3.0 (Fischl, 2012) to ensure accurate classification of gray–white matter boundaries, correct skull stripping, and true separation between brain and non-brain matter. All surfaces were checked and edited in the coronal, sagittal, and axial planes to ensure optimal results. All edits strictly followed the guidelines provided by FreeSurfer.[1] The editor was blind to the participants' degrees of literacy. Subsequently, the adjusted images were reprocessed *via* the automated reconstruction pipeline and checked a second time for accurate reconstruction by the editor.

Next, the three-dimensional lGI proposed by Schaer et al. (2008) was computed in FreeSurfer and the Image Processing Toolbox of Matrix Laboratory (MATLAB) 2020b (The Math Works Inc, 2020) to measure the regional folding of the cortex using a spherical kernel of 25 mm at each vertex. Compared to other metrics of cortical folding such as curvature, sulcal depth measurement, and the classical two-dimensional gyrification index (Zilles et al., 1988), the lGI takes the inherent three-dimensional nature of the cortical surface into account and makes it robust against slice orientation and the presence of buried sulci. The automated construction of the lGI was validated against manual measurement and manifested as a reliable measure of gyrification (Schaer et al., 2012).

For the subsequent region of interest (ROI) analyses, eight left-hemispheric ROIs spanning all three reading streams were selected *a priori* based on the functional meta-analysis of Richlan et al. (2009) and previous research in children (Płoński et al., 2017; Beelen et al., 2019; Perdue et al., 2020). For the ventral stream, these included (1) the fusiform gyrus, (2) the inferior temporal gyrus, and (3) the middle temporal gyrus; the dorsal stream was represented by (4) the superior temporal gyrus, (5) the inferior parietal cortex, consisting of the inferior parietal and the angular gyrus, and (6) the supramarginal gyrus; finally, (7) the insular cortex, and (8) the inferior frontal gyrus, a result of combining pars opercularis and pars triangularis, linked to the frontal stream were selected. ROIs were taken from the Desikan-Killiany atlas, which subdivides the cortex into 34 gyral regions based on curvature and sulcal information on the inflated cortex for each hemisphere (Desikan et al., 2006). This automatic labeling has been discussed as having higher accuracy than manual parcelation (Desikan et al., 2006). Next, the mean metrics for all ROIs were extracted from the FreeSurfer output and imported into MATLAB. Although the FreeSurfer average participant is adult-based, it is frequently used in pediatric samples, and surface-based registration has been validated in children ages 1–11 with good alignment of cortical landmarks (Ghosh et al., 2010). No smoothing was applied to the lGI data as it is already intrinsically smoothed on the individual level as defined by the algorithms employed during the lGI procedure (averaging across a 25-diameter circle). Surface area and cortical thickness metrics were smoothed at full width half maximum (FWHM) of 20 mm to approximate the intrinsic smoothing of the lGI algorithms and increase the signal-to-noise ratio (mri_surf2surf). In contrast to volumetric smoothing, surface-based smoothing only averages data from nearby vertices on the cortical surface, preventing the merging of signals from different tissue types and resulting in higher spatial specificity (Greve and Fischl, 2018).

Statistical analyses

Demographic and psychometric data were assessed in MATLAB 2020b. Literacy ability was calculated as the mean of reading fluency of words and pseudowords (SLRT-II), reading comprehension of words, sentences, and text (ELFE 1-6), and spelling accuracy scores from dictation (DERET 1-2+) to match other attempts of reading predictions in German-speaking samples (e.g., Kuhl et al., 2020). Instead of age-normed standard

---

[1] http://surfer.nmr.mgh.harvard.edu/fswiki/FsTutorial/TroubleshootingData

scores, raw scores for all psychometric measures were utilized as we were interested in the within-subject association between measures at TP1 and TP2 and not in comparisons between peers.

A CV elastic net linear regression (Zou and Hastie, 2005) was used to perform a variable selection of the best preliterate brain and behavioral predictors of literacy ability, as introduced above. Furthermore, previous studies have shown that there may be subtle differences in literacy skills between females and males as well as an association with non-verbal intelligence (Flannery et al., 2000; Rutter et al., 2004; Liederman et al., 2005; Cotton and Crewther, 2009; Halpern, 2013; Quinn and Wagner, 2015). Therefore, sex and non-verbal intelligence measured at TP1 were added as additional prediction variables to the model. All analyses steps outlined below were implemented in R 4.1.2 (R Core Team, 2021) using the packages caret 6.0-90 (Kuhn, 2015) and glmnet 4.1-3 (Friedman et al., 2010).

Regularized analysis methods such as elastic net regressions are better suited for handling neuroimaging data than classical linear regression models because of their superiority in dealing with inter-correlated predictors (Carroll et al., 2009). In regularized linear models, a penalty term is added to the least-squares objective function (Hoerl and Kennard, 1970). The amount of penalization is governed by smoothing parameters. The penalty, in turn, controls the bias-variance trade-off by reducing variance at the cost of deliberately introducing some bias into the resulting estimators (Hastie et al., 2009). The elastic net penalty (Zou and Hastie, 2005) combines the power of a least-absolute-shrinkage-and-selection-operator (LASSO) regularization (Tibshirani, 1996) to select relevant variables in the model, i.e., set the weights of certain coefficients to zero, with a Ridge penalty (Hoerl and Kennard, 1970), which takes correlation between prediction variables better into account (Cho et al., 2010). Thus, highly correlated predictors are retained or discarded from the model as variables, making it an ideal regression approach for brain data with high ratios of features to cases (Zou and Hastie, 2005; Whelan and Garavan, 2014).

The elastic net aims at minimizing the following loss function:

$$L_{enet}\left(\hat{\beta}\right) = \frac{\sum_{i=1}^{n}(y_i - x_i^T\hat{\beta})^2}{2n} + \lambda\left(\frac{1-\alpha}{2}\sum_{j=1}^{m}\hat{\beta}_j^2 + \alpha\sum_{j=1}^{m}\left|\hat{\beta}_j\right|\right),$$

where $\sum_{j=1}^{m}\hat{\beta}_j^2$ represents the LASSO (or L1), and $\sum_{j=1}^{m}\left|\hat{\beta}_j\right|$ the Ridge (or L2) penalty, λ is the tuning parameter that determines the weight of the composite regularization term, i.e., the bias-variance trade-off, and α the hyperparameter that controls the balance between the two types of penalties. The former parameter ranges from 0 to infinity, with $\lambda = 0$ resulting in the ordinary least squares solution due to eliminating the penalty from the equation. The latter hyperparameter may take on values from 0 to 1. If $\alpha = 0$, the regression is identical to the Ridge regression; if $\alpha = 1$, the L2 term cancels out, and the penalty corresponds to the LASSO penalty.

Here, both α and λ were estimated within the inner loop of a nested leave-one-out cross-validation (LOOCV). LOOCV is the most extreme form of *k*-fold CV (Allen, 1974; Stone, 1974; Geisser, 1975). When using LOOCV, the number of folds equals the number of observations ($k = n$), which is especially valuable when the sample size is small (Allen, 1974; Stone, 1974). A nested LOOCV method was chosen to avoid a biased, overly optimistic estimate of the true generalization error, which may be the case if observations are part of both the training and test dataset (Varma and Simon, 2006). This framework is visualized in **Figure 1**.

The best model in the inner loop was selected based on the lowest prediction error, quantified by the root mean squared error (RMSE). The resulting model was then utilized in the outer loop to predict literacy ability. In turn, the vector of predicted test observations over iterations was entered into several formulas to calculate the following goodness-of-fit measures: the mean absolute error (MAE), the RMSE, and R^2 based on the test set observations. Top prediction variables were identified based on their variable importance as calculated by caret (Kuhn, 2015).

An ordinary least squares linear regression was computed using only the two cognitive-linguistic skills, i.e., PA and RAN. The resulting RMSE was then compared to the RMSE from the elastic net model to determine if adding the gray matter macrostructural prediction variables and covariates would improve the prediction of literacy ability.

For comparison, the prediction model was recomputed using a nested 10-fold CV procedure. This process was repeated 50 times to enhance the estimate of the true unknown underlying mean model performance by fitting and evaluating more models and thereby controlling for potential biases caused by the pseudorandom split of the data (Vehtari et al., 2017).

Frequently, estimated total intracranial volume (eTIV) is used as a covariate in similar research paradigms. Therefore, the LOOCV model was recomputed with eTIV as an additional prediction variable. However, this variable was not part of the final model because past research has suggested that controlling for eTIV may overcorrect for differences in head volume and may reduce individual differences in continuous regression approaches (Westman et al., 2013; Wierenga et al., 2014). Due to this incongruity, we decided to focus on the more parsimonious model, in line with Occam's razor (Blumer et al., 1987).

Furthermore, partial correlations between the gray matter features within the ROIs were calculated to investigate the associations between the different indices. The resulting *p*-values were Holm–Bonferroni family-wise error corrected. A correlational whole-brain analysis was conducted to identify potential areas associated with literacy ability but was not captured by the selected ROIs. The methodology and results are described in **Supplementary material**.

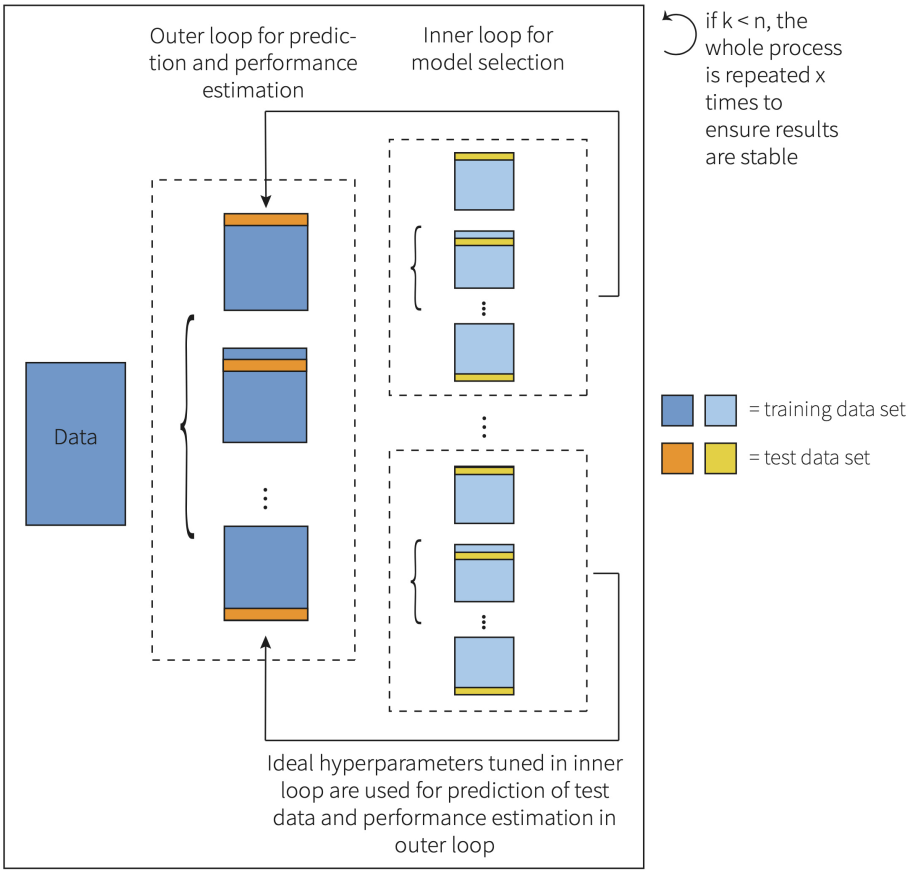


FIGURE 1
Representation of the nested *k*-fold cross-validation framework.

Results

A LOOCV elastic net regression model was computed to identify the behavioral and gray matter features measured at a preliterate age that were the strongest predictors of literacy ability 2 years later. The predictive strength of literacy ability was improved when gray matter macrostructural features were added as prediction variables on the top of the cognitive-linguistic preliterate skills variables, i.e., PA and RAN. More precisely, the RMSE of 0.82 decreased by 29% to 0.58 in the elastic net model that included the gray matter macrostructural features. The same elastic net regression produced highly accurate estimations of literacy ability as testified by a high correlation coefficient between predicted and observed values of $r = 0.80$ (see **Figure 2**).

On average, the LOOCV models were reduced to 20 prediction variables, i.e., dropping approximately eight predictors in each iteration. This was affected by a low mean value of lambda, $\bar{\lambda} = 0.01$, and a high mean value of alpha, $\bar{\alpha} = 0.85$, which led to a small weight of the composite penalty term with a stronger contribution of the LASSO penalty. The models were approximately stable over iterations, as indicated in **Table 2**.

The features with the greatest contribution to the prediction of literacy ability were lGI in the insular cortex and cortical volume in the fusiform gyrus. Additionally, lGI in the supramarginal and posterior inferior frontal gyrus and cortical volume in the inferior temporal gyrus were also important variables in predicting reading and writing skills. Both cognitive-linguistic prediction variables were in the final model, with PA ($\bar{b} = 0.41$) explaining twice as much variance in literacy ability as RAN ($\bar{b} = 0.18$). Moreover, sex contributed unique variance to the model ($\bar{b} = 0.15$): Females showed a greater probability of slightly higher literacy scores than males. Non-verbal intelligence ($\bar{b} = 0.01$) was excluded from the model in most iterations of the LOOCV procedure.

The 10-fold CV elastic net model revealed results comparable to the LOOCV model, as documented in **Tables 2–4**. A list of the central gray matter and behavioral prediction variables of the leave-one-out and 10-fold CV models is provided in **Table 4** and visualized in **Figure 3**. A complete listing of all predictors is provided in **Supplementary material**.

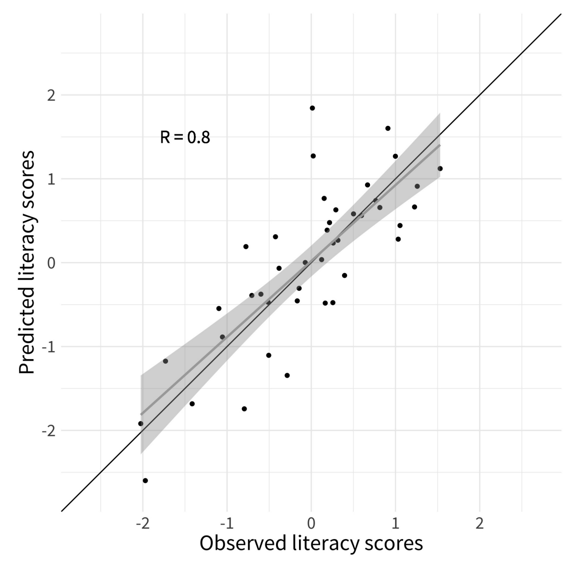


FIGURE 2
Predicted literacy ability from leave-one-out cross-validated (LOOCV) elastic net regression model. Scatterplots depict the observed literacy ability (*y*-axis) by predicted literacy skills (*x*-axis). The black line represents the line of identity; the gray line is the regression line of literacy ability on predicted literacy, with the shaded area representing a 95% pointwise confidence interval.

Model predictions, mean coefficients and feature ranks remained almost unchanged when intracranial volume was added as a covariate.

Correlational analyses revealed a high, positive association of cortical volume and surface area across ROIs ($r = 0.87$–0.96, $p < 0.001$ Holm–Bonferroni corrected). In contrast, lGI correlated only low to moderately with cortical volume ($r = 0.03$–0.41) and surface area ($r = 0.06$–0.45).

Discussion

The present study aimed to predict future literacy in preliterate children using a continuous sampling approach. We successfully applied a linear regression approach to predict future reading acquisition of children with measures gathered at the end of kindergarten (i.e., before formal reading instruction). The CV model captured individual differences in future literacy based on gray matter macrostructural features and cognitive-linguistic preliterate skills measured at the end of kindergarten. More specifically, the elastic net regularized linear regression models predicted approximately 65% of the variance in literacy 2 years later (**Figure 2**). Intriguingly, the top five features contributing to the prediction are part of the three major reading streams. These were lGI in the insular cortex, the inferior frontal gyrus, the supramarginal gyrus, and cortical volume in the fusiform gyrus and inferior temporal gyrus. This pattern emphasizes that reading readiness in all crucial parts of the large-scale reading network differs among children and that these individual variations significantly impact reading acquisition.

TABLE 3 Model performance of the leave-one-out (LOO-) and 10-fold cross-validated (CV) elastic net regressions.

	RMSE	MAE	R^2	R
LOOCV	0.575	0.459	0.652	0.807
10-fold CV	0.579	0.438	0.646	0.804

R, coefficient of determination; RMSE, root mean squared error; MAE, mean absolute error.

Individualized prediction of future literacy in cross-validated frameworks

The identification and validation of biomarkers for the early detection of children at risk of developing reading difficulties have been of major interest over the past years. A better knowledge of the prerequisites facilitating or hampering reading acquisition improves our understanding of the multifaceted learning process and forms the basis for developing specific preventive treatment strategies. Recently, it has been advocated to replace the traditional correlational approach with prediction frameworks to generalize the observed patterns to independent data sets (Gabrieli et al., 2015; Dubois and Adolphs, 2016). In line with this, we applied CV regression models to assess the composition of neural and behavioral markers of future reading acquisition. We were able to predict future literacy with high precision. Compared to classical linear regressions, the algorithm prevents over-fitting by adding additional constraints to the model. Automatic feature selection in the training phase results in sparse predictive models, making it an optimal algorithm for neuroimaging, characterized by many features and

TABLE 2 Tuned hyperparameters and selected non-zero coefficients of the leave-one-out (LOO-) and 10-fold cross-validated (CV) elastic net regressions.

	Alpha (mean ± SD)	Lambda (mean ± SD)	Number of non-zero coefficients (mode, range)
LOOCV	0.853 ± 0.328	0.012 ± 0.006	20, 16–27
10-fold CV	0.658 ± 0.392	0.022 ± 0.016	19, 13–27

The count of regression variables does not include the model's intercept. Twenty-eight variables were entered into the model. SD, standard deviation.

TABLE 4 The top ten prediction variables of literacy ability (mean coefficient >0.15) based on the leave-one-out (LOO) and 10-fold cross-validated (CV) elastic net linear regressions.

Rank	Selection frequency (LOOCV)	Mean coefficient		Gray matter feature or psychometric variable	Region of the left hemisphere
		LOOCV	10-fold CV		
1	76.19%	0.71	0.61	Local gyrification	Insular cortex
2	71.43%	0.69	0.59	Cortical volume	Fusiform gyrus
3	71.43%	−0.50	−0.43	Local gyrification	Supramarginal gyrus
4	64.29%	−0.47	−0.43	Cortical volume	Inferior temporal gyrus
5	57.14%	−0.40	−0.33	Local gyrification	Inferior frontal gyrus
6	59.52%	0.41	0.41	Phonological awareness	
7	76.19%	−0.24	−0.24	Surface area	Inferior temporal gyrus
8	61.90%	0.21	0.19	Cortical volume	Middle temporal gyrus
9	38.10%	0.18	0.18	Rapid naming	
10	35.71%	0.18	0.15	Local gyrification	Middle temporal gyrus

Predictors are listed according to their average rank. The rank displays the variable importance as defined by caret (Kuhn, 2015), i.e., how much unique variance of the response variable can be explained by this variable. Compared to the mean correlation coefficient, this metric is more stable against outlier models. The selection frequency shows how often the variable was chosen at this rank for the LOOCV regression. All prediction variables were standardized before being entered into the model.

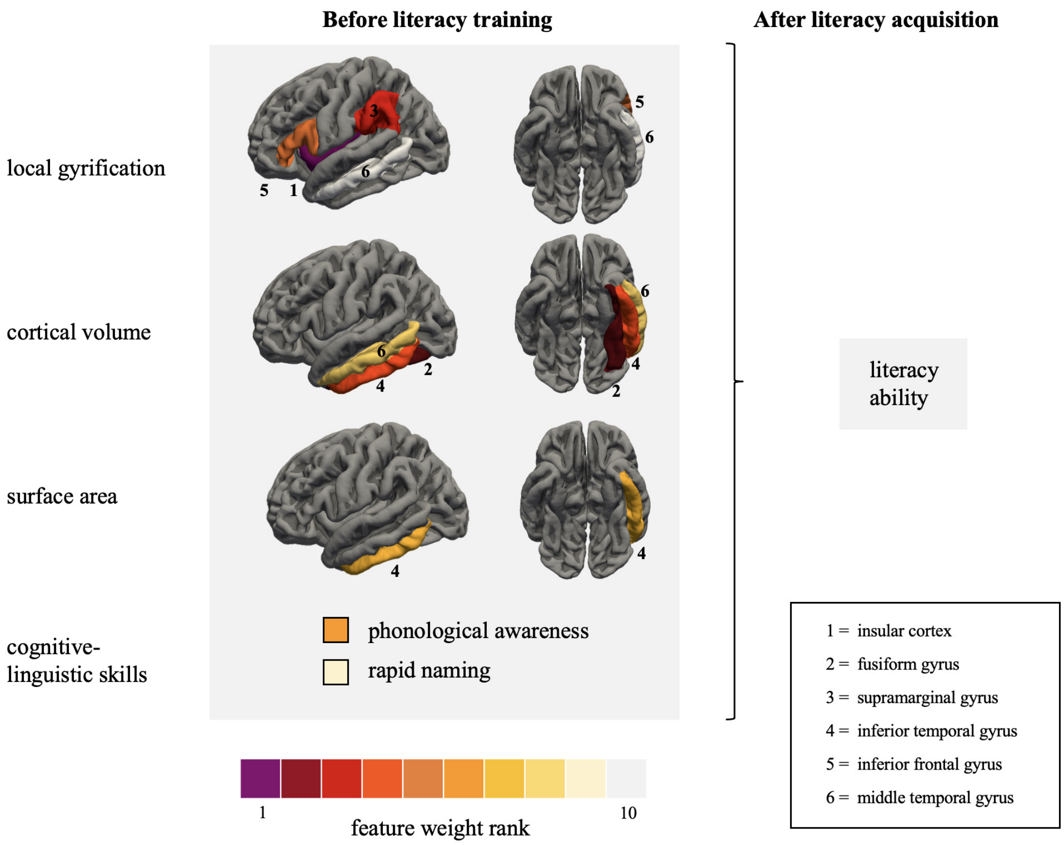


FIGURE 3
Visualization of the primary features collected at the end of kindergarten predicting literacy ability measured at the end of the second grade. Features with the greatest contribution to the prediction of literacy ability are coded purple and red. The regions are depicted on the left pial surface of the FreeSurfer template based on the Desikan-Killiany atlas (Desikan et al., 2006). Literacy ability is defined as the summary score of reading fluency, reading comprehension and spelling accuracy, measured with the *Salzburger Lese- und Rechtschreibtest*, *Ein Leseverständnistest für Erst- bis Sechstklässler*, and *Deutscher Rechtschreibtest für das erste und zweite Schuljahr*, respectively.

small sample sizes (Cui and Gong, 2018). Computationally, our results replicate and thus further demonstrate the effectiveness of elastic net penalized linear regressions for gray matter-based reading prediction as reported by Cui et al. (2018) for a large sample size (N = 870). We further generalize its suitability to the longitudinal prediction of literacy skills based on macrostructural and behavioral information. Similar to Cui et al. (2018), applying the previously built models to new cohorts of preliterate children in cross-linguistic studies would be interesting. If the biomarkers identified here could be replicated across languages, this would strengthen the generality of the approach and the validity of the predictors thus making it possible to test for orthography-specific effects.

The predictive power of reading streams

The three central reading streams did not equally contribute to predicting literacy. Compared to the ventral and frontal streams, individual variations in the macrostructural features linked to the dorsal stream played a minor role. Indeed, out of three *a priori* defined ROIs associated with the dorsal stream, only the supramarginal gyrus was consistently selected during the CV prediction. Interestingly, Linkersdörfer et al. (2012) identified a link between gray matter reduction and functional underactivation in the supramarginal gyrus in dyslexia. The present findings strongly suggest that individual variations in the supramarginal gyrus apply to impaired reading and affect future literacy in a wide range of typically developing children. In a previous study, we examined possible early neurofunctional literacy markers in the same cohort of children (Liebig et al., 2021). We observed a correlation between RAN and neural functioning in the supramarginal gyrus. Likewise, a cluster in the angular gyrus, extending to the supramarginal gyrus, predicted future reading fluency. The latter, however, did not survive additional rigorous correction for the number of regression models (Liebig et al., 2021). Taken together, the present results converge with our previous findings to suggest that variations in both the functional and structural architecture of the supramarginal gyrus might be a promising biomarker for predicting future reading acquisition.

In contrast, neither the superior temporal gyrus nor the inferior parietal cortex explained significant amounts of unique variance in literacy skills. On the one hand, the subordinate role of the dorsal stream stands in contrast with the classical model of reading acquisition, according to which initial decoding relies on the dorsal stream. In contrast, parallel automatized word recognition relies on the ventral stream and emerges only later during reading acquisition (Pugh et al., 2000, 2013). On the other hand, recent functional (Kronbichler et al., 2007; Price and Devlin, 2011; Richlan et al., 2011; Liebig et al., 2017) and structural (Richlan et al., 2013; Williams et al., 2018) brain imaging studies on reading acquisition strongly emphasize the crucial role of the ventral stream not only in beginning readers but already in preliterate children (Hoeft et al., 2011; Liebig et al., 2021). The fact that all *a priori* defined regions of the ventral stream were reliably selected during the prediction iterations supports the idea that neuroanatomical characteristics of the ventral occipito-temporal cortex are essential determinants of successful reading acquisition. This finding does not question the general importance of the dorsal stream for initial decoding (Martin et al., 2015; Zhou et al., 2016; Liebig et al., 2017; Braun et al., 2019), but it raises the question of whether the transition from serial decoding to rapid parallel access to written word forms (automatization) requires the integrity of specific neuroanatomical properties related to the ventral stream that can be assessed even prior to reading.

Our results thus extend the fast-growing knowledge about the early importance of the ventral stream in several ways: Firstly, macrostructural features of the ventral stream do not only distinguish between children and adults with and without dyslexia (Linkersdörfer et al., 2012; Richlan et al., 2013) but individual differences in the morphology of gray matter features significantly contribute to the individualized prediction of future reading acquisition. This has been observed in the functional data of Liebig et al. (2021), who showed that neural activity in the ventral stream correlated with RAN and predicted future reading fluency in the same cohort of children. Taken together, we observed highly similar relationships in the same cohort of children on the functional and structural levels when applying different computational approaches (CV prediction vs. classical correlational analysis) both on the whole-brain level and in an ROI-based study. This convergence across the two studies clearly supports the plausibility of our effects. Similarly, the results converge with a recent finding to suggest that increased neural plasticity of temporo-parietal regions in emergent readers supports reading acquisition (Phan et al., 2021) paving the way for early identification and targeted intervention of children at-risk of encountering difficulties during reading development.

Like the ventral stream, all parts of the frontal stream significantly contributed to the prediction models. This result is in line with the interactive account of reading (Price and Devlin, 2011), according to which reading acquisition is marked by top-down influences from frontal to ventral occipito-temporal regions. Hence, individual differences in the gyrification pattern of the frontal stream should be seen in concert with gray matter features in the ventral stream.

Gray matter macrostructural features underlying future literacy

In the present study, we compared different macrostructural indices (lGI, cortical volume, and surface area) to evaluate their suitability for the individualized prediction of literacy. We found that lGI and cortical volume had better predictive power than surface area. Regional specificities seem to drive the individualized prediction when looking at the distributional pattern. Cortical volume was the decisive feature in the ventral stream, whereas the lGI was the strongest feature in the frontal and dorsal stream, which makes it a promising macrostructural feature in relation to reading. Computationally, the lGI allows a more reliable calculation of the cortical folding than previous measures because it utilizes three-dimensional surface properties to fully capture the patterns of the cortical mantle (Schaer et al., 2008). Using lGI as a measure, it has already been shown that developmental dyslexics exhibited a thinner and more gyrified left occipito-temporal cortex (Williams et al., 2018) and a more gyrified primary auditory cortex (Kuhl et al., 2020). The present results suggest that the lGI is also suitable to detect subtle individual differences in continuous sampling. However, future studies need to replicate and thus validate the suitability of lGI in relation to literacy skills.

The considerable importance of cortical volume in the ventral stream is in line with previous results. Several meta-analyses confirm that cortical volume in the ventral occipito-temporal cortex distinguish children and adults with and without dyslexia (Linkersdörfer et al., 2012; Richlan et al., 2013) and is generally associated with reading skills (Eckert et al., 2016; Skeide et al., 2016).

Surface area only played a minor role in the individualized prediction of literacy ability, which might be explained by the high correlation of surface area and cortical volume across regions of interest ($r = 0.87$–0.96). This is in line with the notion that cortical volume is the product of cortical thickness and surface area (Winkler et al., 2010). Instead, lGI and cortical volume were only moderately correlated ($r = 0.03$–0.42). When aiming to capture different aspects of variance in the gray matter, it might thus be advisable to focus on not too strongly correlated features and integrate these into the prediction models.

In the present study, we observe both negative and positive relationships between literacy and the macrostructural features depending on the ROIs. The associations can be characterized as follows: previously, different macrostructural features, i.e., cortical volume and thickness of the supramarginal gyrus, were reliably associated with reading skills. Both positive relationships between cortical volume (Jednoróg et al., 2015; Xia et al., 2018) and longitudinal volume reductions (Houston et al., 2014; Linkersdörfer et al., 2014) were associated with reading skills. In the present study, we observed a negative relationship between literacy and the lGI in the supramarginal gyrus, which was frequently selected as the third most important predictor of literacy. This finding strongly suggests that the gyrification pattern in the dorsal stream also affects literacy.

We observed both positive and negative associations in the ventral stream: while cortical volume in the fusiform gyrus and the middle temporal gyrus was positively associated with future literacy, both cortical volume and surface area exhibited a negative relationship with reading and writing. The positive association partly contradicts previous findings in children (Simon et al., 2013; Skeide et al., 2016). However, the operationalization of literacy and the age of the samples differ substantially from the present study. While we examined literacy ability on different levels to better account for this multifaceted nature of reading, Simon et al. (2013) and Skeide et al. (2016) focused on reading speed. However, positive relationships have been reported when also looking at cortical thickness (Xia et al., 2018). In sum, there is cumulative evidence that different regions of the ventral occipito-temporal cortex crucially relate to individual differences in reading ability and distinguish between children with and without dyslexia (Płoński et al., 2017; Beelen et al., 2019). In the present study, the cortical volume of the fusiform gyrus and the inferior temporal gyrus could explain twice as much unique variance as the two literacy precursory skills. This robust finding aligns with the increasingly recognized importance of the ventral stream in the first steps of reading acquisition (e.g., Hoeft et al., 2011; Liebig et al., 2021).

Local gyrification in the frontal stream also showed both directions, i.e., positive in the insular cortex and negative in the inferior frontal gyrus. Although the insular cortex is an integral part of the language and reading network (Price, 2012) its macrostructural features have seldomly been examined concerning reading acquisition. The lGI of the insular cortex was selected as the strongest predictor of future literacy. The insular cortex has previously been associated with diverse aspects of language and reading (Price, 2012). Most interesting for the present study, the insula might be a crucial part of the phonological network in reading acquisition, which is delayed in children with developmental dyslexia (Łuniewska et al., 2019). Similarly, the insula seems to be more strongly involved in pre-readers compared to readers emphasizing its importance during the first steps of reading acquisition (Monzalvo and Dehaene-Lambertz, 2013; Chyl et al., 2018). With the present results, we provide first evidence that the gyral folding pattern of the insular cortex might be a promising early biomarker of future literacy acquisition in native German-speaking children. However, future research needs to refine this ample evidence and disentangle the contribution of different aspects of gray matter morphology and possible sensitive phases of cortical plasticity. The idea of the interaction of neural plasticity and reading acquisition was recently endorsed by a structural neuroimaging study showing a gray matter volume increase in decisive regions of the ventral and dorsal reading network during the earliest phases of reading acquisition (Phan et al., 2021). Whether this holds for the insular cortex as well needs to be tackled in future research.

In general, the observed pattern of regional-specific directions of the relationships found in the same cohort of participants is in line with previous continuous approaches (Jednoróg et al., 2015; Skeide et al., 2016).

Cognitive-linguistic preliteracy skills

The cognitive-linguistic preliterate skills were among the top ten features that were frequently selected in the CV approach with PA being a stronger predictor of literacy than RAN. This pattern does not entirely align with previous behavioral results in German-speaking children. It has been reported that preliterate RAN reliably predicts reading while PA only becomes significant in beginning readers (Landerl et al., 2019). However, in the present study, we combine neural and behavioral information in a CV predictive framework and operationalized literacy on different levels of complexity, which might have led to the observed differences. Our results suggest that a combination of both behavioral and macrostructural features makes it possible to predict reading outcomes with high accuracy even before the onset of literacy instruction.

Limitations and conclusions

Prior work linking brain anatomy to reading ability was primarily based on groups with and without dyslexia (Linkersdörfer et al., 2012; Richlan et al., 2013). Only a few studies investigated this link with a continuous approach (e.g., Houston et al., 2014; Jednoróg et al., 2015; Torre and Eden, 2019). However, developmental trajectories might differ between individuals within each group of typical and dyslexic readers (Chyl et al., 2021). Thus, continuous sampling and group-based approaches should be combined to track both group differences and interindividual differences. Such an integrated approach allows identifying both general neural makers applicable to the entire range of reading acquisition and abnormal patterns related explicitly to impaired reading.

Furthermore, examining very young children in the MRI scanner led to a greater motion and thus lower image quality than studies with older children. However, we thoroughly controlled the images and applied rather strict dropout criteria to control for the pitfalls of pediatric neuroimaging. From a theoretical perspective, we only provide preliminary insights into the prerequisites of reading acquisition. We systematically targeted possible predictors of future literacy by comparing several macrostructural and behavioral measures. However, even in a CV predictive framework, in which the number of features may expand the number of observations, the maximal number of features has to be limited when aiming to obtain interpretable results. Thus, we utilized an ROI approach rather than a whole-brain analysis and limited the number of gray matter structural features. *A priori* selection of regions and features may risk overseeing relationships beyond the targeted areas and features.

Apart from these limitations, the present study is a further step in applying CV models to examine biomarkers of typical reading acquisition in pediatric neuroimaging. Individual variations in several macrostructural gray matter features in crucial parts of the large-scale left-hemispheric reading network predicted literacy skills 2 years later with high precision. In the predictive framework, the ventral and frontal streams showed considerable importance. Thus, from a theoretical perspective, our results support recent arguments about the importance of the ventral stream in reading acquisition (Hoeft et al., 2011; Richlan, 2012; Liebig et al., 2017) in concert with top-down modulation of the frontal stream (Price and Devlin, 2011).

From a clinical perspective, the present results might also have implications for education and therapy. We provide evidence that children might come to the task of learning to read with different initial conditions at the neuroanatomical and behavioral level that might well impact how quickly and efficiently they will be able to learn at school (see also Liebig et al., 2021). With the increasing number of longitudinal structural and functional studies conducted at the end of kindergarten pointing in the same direction (Chyl et al., 2021), our results clearly favor an early diagnosis of future reading difficulties. Structural neuroimaging might be a promising tool, given that gray matter features are far easier to acquire than functional neuroimaging. Firstly, no task is required and thus, imaging time notably drops compared to functional imaging. Secondly, the requirements put upon the children in terms of attention and compliance decrease substantially in structural neuroimaging, making the features gathered even more objective. Thus, structural imaging seems more feasible in a clinical routine and daily practice than functional neuroimaging. Notably, we observed a substantial overlap of those regions contributing to the prediction in the structural and functional analysis in the same cohort of children (see Liebig et al., 2021, for the functional analysis), emphasizing the potential of gray matter features to be become early biomarkers of normal and impaired reading development.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the German Association for Psychology. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

JL, EF, JZ, and AJ designed the study. MBe analyzed the data with advice from JL and TS. MBe and JL drafted the manuscript. All authors contributed to the final version of the article.

Funding

This work was supported by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF) with grants FKZ 01GJ1010A and 01GJ1404. JZ was supported by the Alexander-von-Humboldt Stiftung and the Institute of Language, Communication and the Brain (ANR-16-CONV-0002).

Acknowledgments

We thank all research assistants of the Department of Experimental and Neurocognitive Psychology at the Freie Universität Berlin and the members of the CCNB for their help with the experiment.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2022.920150/full#supplementary-material

References

Allen, D. M. (1974). 'The Relationship Between Variable Selection and Data Agumentation and a Method for Prediction'. *Technometrics* 16, 125–127. doi: 10.1080/00401706.1974.10489157

Ashburner, J., Barnes, G., Chen, C.-C., Daunizeau, J., Flandin, G., Friston, K., et al. (2021). *'SPM12 Manual'*. London: Functional Imaging Laboratory Wellcome Centre for Human Neuroimaging.

Beelen, C., Vanderauwera, J., Wouters, J., Vandermosten, M., and Ghesquière, P. (2019). 'Atypical gray matter in children with dyslexia before the onset of reading instruction'. *Cortex* 121, 399–413.

Binder, J. R., Desai, R. H., Graves, W. W., and Conant, L. L. (2009). 'Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies'. *Cereb. Cortex* 19, 2767–2796. doi: 10.1093/CERCOR/BHP055

Blumer, A., Ehrenfeucht, A., Haussler, D., and Warmuth, M. K. (1987). 'Occam's Razor'. *Inf. Process. Lett.* 24, 377–380. doi: 10.1016/0020-0190(87)90114-1

Braun, M., Kronbichler, M., Richlan, F., Hawelka, S., Hutzler, F., and Jacobs, A. M. (2019). 'A model-guided dissociation between subcortical and cortical contributions to word recognition'. *Sci. Rep.* 9:4506. doi: 10.1038/s41598-019-41011-9

Bruck, M. (1987). 'The adult outcomes of children with learning disabilities'. *Ann. Dyslexia* 37, 252–263. doi: 10.1007/BF02648071

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., et al. (2013). 'Power failure: Why small sample size undermines the reliability of neuroscience'. *Nat. Rev. Neurosci.* 14, 365–376. doi: 10.1038/nrn3475

Caravolas, M., Lervåg, A., Mousikou, P., Efrim, C., Litavský, M., Onochie-Quintanilla, E., et al. (2012). 'Common Patterns of Prediction of Literacy Development in Different Alphabetic Orthographies'. *Psychol. Sci.* 23, 678–686. doi: 10.1177/0956797611434536

Carroll, M. K., Cecchi, G. A., Rish, I., Garg, R., and Rao, A. R. (2009). 'Prediction and interpretation of distributed neural activity with sparse models'. *Neuroimage* 44, 112–122. doi: 10.1016/j.neuroimage.2008.08.020

Cho, S., Kim, K., Kim, Y. J., Lee, J. K., Cho, Y. S., Lee, J. Y., et al. (2010). 'Joint Identification of Multiple Genetic Variants via Elastic-Net Variable Selection in a Genome-Wide Association Analysis'. *Ann. Hum. Genet.* 74, 416–428. doi: 10.1111/J.1469-1809.2010.00597.X

Chyl, K., Fraga-González, G., Brem, S., and Jednoróg, K. (2021). 'Brain dynamics of (a)typical reading development—a review of longitudinal studies'. *npj Sci. Learn.* 6:4. doi: 10.1038/s41539-020-00081-5

Chyl, K., Kossowski, B., Dębska, A., Łuniewska, M., Banaszkiewicz, A., Łelechowska, A., et al. (2018). 'Prereader to beginning reader: Changes induced by reading acquisition in print and speech brain networks'. *J. Child Psychol. Psychiatry Allied Disc.* 59, 76–87. doi: 10.1111/jcpp.12774

Cohen, L., and Dehaene, S. (2004). 'Specialization within the ventral stream: The case for the visual word form area'. *Neuroimage* 22, 466–476.

Cotton, S. M., and Crewther, S. G. (2009). The Relationship Between Reading and Intelligence in Primary School Aged Children: Implications for Definitional Models of Dyslexia. *Open Educ. J.* 2, 42–50.

Cui, Z., and Gong, G. (2018). 'The effect of machine learning regression algorithms and sample size on individualized behavioral prediction with functional connectivity features'. *Neuroimage* 178, 622–637. doi: 10.1016/j.neuroimage.2018.06.001

Cui, Z., Su, M., Li, L., Shu, H., and Gong, G. (2018). 'Individualized Prediction of Reading Comprehension Ability Using Gray Matter Volume'. *Cereb. Cortex* 28, 1656–1672. doi: 10.1093/cercor/bhx061

Dehaene, S., and Cohen, L. (2011). 'The unique role of the visual word form area in reading'. *Trends Cogn. Sci.* 15, 254–262. doi: 10.1016/j.tics.2011.04.003

Dehaene, S., Cohen, L., Morais, J., and Kolinsky, R. (2015). 'Illiterate to literate: Behavioural and cerebral changes induced by reading acquisition'. *Nat. Rev. Neurosci.* 16, 234–244. doi: 10.1038/nrn3924

Dehaene-Lambertz, G., Monzalvo, K., and Dehaene, S. (2018). 'The emergence of the visual word form: Longitudinal evolution of category-specific ventral visual areas during reading acquisition'. *PLoS Biol.* 16:e2004103. doi: 10.1371/journal.pbio.2004103

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., et al. (2006). 'An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest'. *Neuroimage* 31, 968–980. doi: 10.1016/j.neuroimage.2006.01.021

Dubois, J., and Adolphs, R. (2016). 'Building a Science of Individual Differences from fMRI'. *Trends Cogn. Sci.* 20, 425–443. doi: 10.1016/j.tics.2016.03.014

Eckert, M. A., Berninger, V. W., Vaden, K. I., Gebregziabher, M., and Tsu, L. (2016). 'Gray matter features of reading disability: A combined meta-analytic and direct analysis approach'. *eNeuro* 3, 11296–11301. doi: 10.1523/ENEURO.0103-15.2015

Elliott, M. L., Knodt, A. R., Ireland, D., Morris, M. L., Poulton, R., Ramrakha, S., et al. (2020). 'What Is the Test-Retest Reliability of Common Task-Functional MRI Measures? New Empirical Evidence and a Meta-Analysis'. *Psychol. Sci.* 31, 792–806. doi: 10.1177/0956797620916786

Fischl, B. (2012). 'FreeSurfer'. *Neuroimage* 62, 774–781. doi: 10.1016/j.neuroimage.2012.01.021

Flannery, K. A., Liederman, J., Daly, L., and Schultz, J. (2000). 'Male prevalence for reading disability is found in a large sample of Black and White children free from ascertainment bias'. *J. Int. Neuropsychol. Soc.* 6, 433–442. doi: 10.1017/S1355617700644016

Fletcher, J. M. (2009). 'Dyslexia: The evolution of a scientific concept'. *J. Int. Neuropsychol. Soc.?* 15, 501–508. doi: 10.1017/S1355617709090900

Fletcher, J. M., and Vaughn, S. (2009). 'Response to Intervention: Preventing and Remediating Academic Difficulties'. *Child Dev. Perspect.* 3, 30–37. doi: 10.1111/j.1750-8606.2008.00072.x

Francis, D., Fletcher, J., Stuebing, K., Lyon, G., Shaywitz, B., and Shaywitz, S. (2005). 'Psychometric Approaches to the Identification of LD IQ and Achievement Scores Are Not Sufficient'. *J. Learn. Disabil.* 38, 98–108. doi: 10.1177/00222194050380020101

Frangou, S., Modabbernia, A., Williams, S. C. R., Papachristou, E., Doucet, G. E., Agartz, I., et al. (2022). 'Cortical thickness across the lifespan: Data from 17,075 healthy individuals aged 3–90 years'. *Hum. Brain Mapp.* 43, 431–451. doi: 10.1002/hbm.25364

Friedman, J., Hastie, T., and Tibshirani, R. (2010). 'Regularization paths for generalized linear models via coordinate descent'. *J. Stat. Softw.* 33, 1–22. doi: 10.18637/jss.v033.i01

Froehlich, E., Liebig, J., Morawetz, C., Ziegler, J. C., Braun, M., Heekeren, H. R., et al. (2018). 'Same Same But Different: Processing Words in the Aging Brain'. *Neuroscience* 371, 75–95. doi: 10.1016/j.neuroscience.2017.11.042

Gabrieli, J. D. E., Ghosh, S. S., and Whitfield-Gabrieli, S. (2015). 'Prediction as a Humanitarian and Pragmatic Contribution from Human Cognitive Neuroscience'. *Neuron* 85, 11–26. doi: 10.1016/j.neuron.2014.10.047

Geisser, S. (1975). 'The predictive sample reuse method with applications'. *J. Am. Stat. Assoc.* 70, 320–328. doi: 10.1080/01621459.1975.10479865

Ghosh, S. S., Kakunoori, S., Augustinack, J., Nieto-Castanon, A., Kovelman, I., Gaab, N., et al. (2010). 'Evaluating the validity of volume-based and surface-based brain image registration for developmental cognitive neuroscience studies in children 4 to 11years of age'. *Neuroimage* 53, 85–93. doi: 10.1016/j.neuroimage.2010.05.075

Gilmore, J. H., Santelli, R. K., and Gao, W. (2018). 'Imaging structural and functional brain development in early childhood'. *Nat. Rev. Neurosci.* 19, 123–137. doi: 10.1038/nrn.2018.1

Greve, D. N., and Fischl, B. (2018). 'False positive rates in surface-based anatomical analysis'. *Neuroimage* 171, 6–14. doi: 10.1016/j.neuroimage.2017.12.072

Halpern, D. F. (2013). *Sex Differences in Cognitive Abilities*. London: Psychological Press.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Berlin: Springer.

Hoeft, F., McCandliss, B. D., Black, J. M., Gantman, A., Zakerani, N., Hulme, C., et al. (2011). 'Neural systems predicting long-term outcome in dyslexia'. *Proc. Natl. Acad. Sci. U.S.A.* 108, 361–366. doi: 10.1073/pnas.1008950108

Hoeft, F., Ueno, T., Reiss, A. L., Meyler, A., Whitfield-Gabrieli, S., Glover, G. H., et al. (2007). 'Prediction of Children's Reading Skills Using Behavioral. Functional, and Structural Neuroimaging Measures'. *Behav. Neurosci.* 121, 602–613. doi: 10.1037/0735-7044.121.3.602

Hoerl, A. E., and Kennard, R. W. (1970). 'Ridge Regression: Biased Estimation for Nonorthogonal Problems'. *Technometrics* 12, 55–67. doi: 10.1080/00401706.1970.10488634

Hofmann, M. J., and Jacobs, A. M. (2014). 'Interactive activation and competition models and semantic context: From behavioral to brain data'. *Neurosci. Biobehav. Rev.* 46, 85–104. doi: 10.1016/j.neubiorev.2014.06.011

Hosseini, S. M. H., Black, J. M., Soriano, T., Bugescu, N., Martinez, R., Raman, M. M., et al. (2013). 'Topological properties of large-scale structural brain networks in children with familial risk for reading difficulties'. *Neuroimage* 71, 260–274. doi: 10.1016/j.neuroimage.2013.01.013

Houston, S. M., Lebel, C., Katzir, T., Manis, F. R., Kan, E., Rodriguez, G. R., et al. (2014). 'Reading skill and structural brain development'. *Neuroreport* 25, 347–352. doi: 10.1097/WNR.0000000000000121

Im, K., Raschle, N. M., Smith, S. A., Ellen Grant, P., and Gaab, N. (2016). 'Atypical Sulcal Pattern in Children with Developmental Dyslexia and At-Risk Kindergarteners'. *Cereb. Cortex* 26, 1138–1148. doi: 10.1093/cercor/bhu305

Jansen, H. (2002). *BISC: Bielefelder Screening zur Früherkennung von Lese-Rechtschreibschwierigkeiten*. Göttingen: Hogrefe.

Jednoróg, K., Marchewka, A., Altarelli, I., Monzalvo Lopez, A. K., van Ermingen-Marbach, M., Grande, M., et al. (2015). 'How reliable are gray matter disruptions in specific reading disability across multiple countries and languages? Insights from a large-scale voxel-based morphometry study'. *Hum. Brain Mapp.* 36, 1741–1754. doi: 10.1002/HBM.22734

Kronbichler, M., Bergmann, J., Hutzler, F., Staffen, W., Mair, A., Ladurner, G., et al. (2007). 'Taxi vs. taksi: On orthographic word recognition in the left ventral occipitotemporal cortex'. *J. Cogn. Neurosci.* 19, 1584–1594. doi: 10.1162/jocn.2007.19.10.1584

Kuhl, U., Neef, N. E., Kraft, I., Schaadt, G., Dörr, L., Brauer, J., et al. (2020). 'The emergence of dyslexia in the developing brain'. *Neuroimage* 211:116633. doi: 10.1016/j.neuroimage.2020.116633

Kuhn, M. (2015). *'Caret: Classification and Regression Training'*. Houghton: Astrophysics Source Code Library.

Lancaster, J. L., and Martinez, M. J. (2006). *'Multi-image Analysis GUI'*. Houston: University of Texas Health Science Center.

Landerl, K., Castles, A., and Parrila, R. (2022). 'Cognitive Precursors of Reading: A Cross-Linguistic Perspective'. *Sci. Stud. Read.* 26, 111–124. doi: 10.1080/10888438.2021.1983820

Landerl, K., Freudenthaler, H. H., Heene, M., De Jong, P. F., Desrochers, A., Manolitsis, G., et al. (2019). 'Phonological Awareness and Rapid Automatized Naming as Longitudinal Predictors of Reading in Five Alphabetic Orthographies with Varying Degrees of Consistency'. *Sci. Stud. Read.* 23, 220–234. doi: 10.1080/10888438.2018.1510936

Landerl, K., and Moll, K. (2010). 'Comorbidity of learning disorders: Prevalence and familial transmission'. *J. Child Psychol. Psychiatry Allied Disc.* 51, 287–294. doi: 10.1111/j.1469-7610.2009.02164.x

Landerl, K., Ramus, F., Moll, K., Lyytinen, H., Leppänen, P. H. T., Lohvansuu, K., et al. (2013). 'Predictors of developmental dyslexia in European orthographies with varying complexity'. *J. Child Psychol. Psychiatry* 54, 686–694. doi: 10.1111/JCPP.12029

Lenhard, W., and Schneider, W. (2006). *Ein Leseverständnistest für Erst-und Sechstklässler: ELFE 1-6*. Göttingen: Hogrefe.

Liebig, J., Friederici, A. D., Neef, N. E., Friederici, A. D., Emmrich, F., Brauer, J., et al. (2020). 'Auditory brainstem measures and genotyping boost the prediction of literacy: A longitudinal study on early markers of dyslexia'. *Dev. Cogn. Neurosci.* 46:100869. doi: 10.1016/j.dcn.2020.100869

Liebig, J., Froehlich, E., Morawetz, C., Braun, M., Jacobs, A. M., Heekeren, H. R., et al. (2017). 'Neurofunctionally dissecting the reading system in children'. *Dev. Cogn. Neurosci.* 27, 45–57. doi: 10.1016/j.dcn.2017.07.002

Liebig, J., Froehlich, E., Sylvester, T., Braun, M., Heekeren, H. R., Ziegler, J. C., et al. (2021). 'Neural processing of vision and language in kindergarten is associated with prereading skills and predicts future literacy'. *Hum. Brain Mapp.* 42, 3517–3533. doi: 10.1002/hbm.25449

Liederman, J., Kantrowitz, L., and Flannery, K. (2005). 'Male vulnerability to reading disability is not likely to be a myth: A call for new data.'. *J. Learn. Disabil.* 38, 109–129. doi: 10.1177/00222194050380020201

Linkersdörfer, J., Jurcoane, A., Lindberg, S., Kaiser, J., Hasselhorn, M., Fiebach, C. J., et al. (2014). 'The association between gray matter volume and reading proficiency: A longitudinal study of beginning readers'. *J. Cogn. Neurosci.* 27, 308–318. doi: 10.1162/jocn_a_00710

Linkersdörfer, J., Lonnemann, J., Lindberg, S., Hasselhorn, M., and Fiebach, C. J. (2012). 'Grey Matter Alterations Co-Localize with Functional Abnormalities in Developmental Dyslexia: An ALE Meta-Analysis'. *PLoS One* 7:e43122. doi: 10.1371/journal.pone.0043122

Livingston, E., Siegel, L., and Ribary, U. (2018). 'Developmental dyslexia: Emotional impact and consequences'. *Aust. J. Learn. Diffic.* 23, 1–29. doi: 10.1080/19404158.2018.1479975

Lohvansuu, K., Hämäläinen, J. A., Ervast, L., Lyytinen, H., and Leppänen, P. H. T. (2018). 'Longitudinal interactions between brain and cognitive measures on reading development from 6 months to 14 years'. *Neuropsychologia* 108, 6–12. doi: 10.1016/j.neuropsychologia.2017.11.018

Łuniewska, M., Chyl, K., Dębska, A., Banaszkiewicz, A., Żelechowska, A., Marchewka, A., et al. (2019). 'Children With Dyslexia and Familial Risk for Dyslexia Present Atypical Development of the Neuronal Phonological Network'. *Front. Neurosci.* 13:1287. doi: 10.3389/fnins.2019.01287

Martin, A., Schurz, M., Kronbichler, M., and Richlan, F. (2015). 'Reading in the brain of children and adults: A meta-analysis of 40 functional magnetic resonance imaging studies'. *Hum. Brain Mapp.* 36, 1963–1981. doi: 10.1002/HBM.22749

Moll, K., and Landerl, K. (2010). *'SLRT-II: Lese-und Rechtschreibtest; Weiterentwicklung des Salzburger Lese-und Rechtschreibtests (SLRT)'*. Bern: Huber.

Monzalvo, K., and Dehaene-Lambertz, G. (2013). 'How reading acquisition changes children's spoken language network'. *Brain Lang.* 127, 356–365. doi: 10.1016/j.bandl.2013.10.009

Norton, E. S., and Wolf, M. (2012). 'Rapid Automatized Naming (RAN) and Reading Fluency: Implications for Understanding and Treatment of Reading Disabilities'. *Ann. Rev. Psychol.* 63, 427–452. doi: 10.1146/annurev-psych-120710-100431

Perdue, M. V., Mednick, J., Pugh, K. R., and Landi, N. (2020). 'Gray Matter Structure Is Associated with Reading Skill in Typically Developing Young Readers'. *Cereb. Cortex* 30, 5449–5459. doi: 10.1093/cercor/bhaa126

Perry, C., Zorzi, M., and Ziegler, J. C. (2019). 'Understanding Dyslexia Through Personalized Large-Scale Computational Models'. *Psychol. Sci.* 30, 386–395. doi: 10.1177/0956797618823540

Petermann, F., and Petermann, U. J. (2011). *Wechsler Intelligence Scale for Children§-Fourth Edition - Manual 1: Grundlagen, Testauswertung und Interpretation*. Frankfurt am Main: Pearson.

Phan, T. V., Sima, D., Smeets, D., Ghesquière, P., Wouters, J., and Vandermosten, M. (2021). 'Structural brain dynamics across reading development: A longitudinal MRI study from kindergarten to grade 5'. *Hum. Brain Mapp.* 42, 4497–4509. doi: 10.1002/hbm.25560

Płoński, P., Gradkowski, W., Altarelli, I., Monzalvo, K., van Ermingen-Marbach, M., Grande, M., et al. (2017). 'Multi-parameter machine learning approach to the neuroanatomical basis of developmental dyslexia'. *Hum. Brain Mapp.* 38, 900–908. doi: 10.1002/hbm.23426

Price, C. J. (2012). 'A review and synthesis of the first 20years of PET and fMRI studies of heard speech, spoken language and reading'. *Neuroimage* 62, 816–847. doi: 10.1016/j.neuroimage.2012.04.062

Price, C. J., and Devlin, J. T. (2011). 'The Interactive Account of ventral occipitotemporal contributions to reading'. *Trends Cogn. Sci.* 15, 246–253. doi: 10.1016/j.tics.2011.04.001

Pugh, K. R., Einar Mencl, W., Jenner, A. R., Katz, L., Frost, S. J., Lee Ren, J., et al. (2000). 'Functional neuroimaging studies of reading and reading disability (developmental dyslexia). *Mental Retard. Dev. Disabil. Res. Rev.* 6, 207–213. doi: 10.1002/1098-277920006:3<207::AID-MRDD8<3.0.CO;2-P

Pugh, K. R., Landi, N., Preston, J. L., Mencl, W. E., Austin, A. C., Sibley, D., et al. (2013). 'The relationship between phonological and auditory processing and brain organization in beginning readers'. *Brain Lang.* 125, 173–183. doi: 10.1016/j.bandl.2012.04.004

Quinn, J. M., and Wagner, R. K. (2015). 'Gender Differences in Reading Impairment and in the Identification of Impaired Readers: Results From a Large-Scale Study of At-Risk Readers'. *J. Learn. Disabil.* 48, 433–445.

R Core Team (2021). *'R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing'.

Ramus, F., Altarelli, I., Jednoróg, K., Zhao, J., and Scotto Di Covella, L. (2018). 'Neuroanatomy of developmental dyslexia: Pitfalls and promise'. *Neurosci. Biobehav. Rev.* 84, 434–452. doi: 10.1016/j.neubiorev.2017.08.001

Raven, J. C., and Court, J. H. (1998). *Raven's Coloured Progressive Matrices And Vocabulary Scales*. Oxford: Oxford pyschologists Press.

Raznahan, A., Shaw, P., Lalonde, F., Stockman, M., Wallace, G. L., Greenstein, D., et al. (2011). 'How Does Your Cortex Grow?'. *J. Neurosci.* 31, 7174–7177. doi: 10.1523/JNEUROSCI.0054-11.2011

Richlan, F. (2012). Developmental dyslexia: Dysfunction of a left hemisphere reading network. *Front. Hum. Neurosci.* 6:120. doi: 10.3389/fnhum.2012.00120

Richlan, F., Kronbichler, M., and Wimmer, H. (2009). 'Functional abnormalities in the dyslexic brain: A quantitative meta-analysis of neuroimaging studies'. *Hum. Brain Mapp.* 30, 3299–3308. doi: 10.1002/hbm.20752

Richlan, F., Kronbichler, M., and Wimmer, H. (2011). 'Meta-analyzing brain dysfunctions in dyslexic children and adults'. *Neuroimage* 56, 1735–1742. doi: 10.1016/j.neuroimage.2011.02.040

Richlan, F., Kronbichler, M., and Wimmer, H. (2013). 'Structural abnormalities in the dyslexic brain: A meta-analysis of voxel-based morphometry studies'. *Hum. Brain Mapp.* 34, 3055–3065. doi: 10.1002/hbm.22127

Rutter, M., Caspi, A., Fergusson, D., Horwood, L. J., Goodman, R., Maughan, B., et al. (2004). 'Sex Differences in Developmental Reading Disability: New Findings from 4 Epidemiological Studies'. *J. Am. Med. Assoc.* 291, 2007–2012. doi: 10.1001/jama.291.16.2007

Schaer, M., Bach Cuadra, M., Tamarit, L., Lazeyras, F., Eliez, S., and Thiran, J. (2008). 'A Surface-based approach to quantify local cortical gyrification'. *IEEE Trans. Med. Imaging* 27, 161–170. doi: 10.1109/TMI.2007.903576

Schaer, M., Cuadra, M. B., Schmansky, N., Fischl, B., Thiran, J., and Eliez, S. (2012). 'How to measure cortical folding from mr images: A step-by-step tutorial to compute local gyrification index'. *J. Vis. Exp.* 59:e3417. doi: 10.3791/3417

Schulte-Körne, G. (2010). 'Diagnostik und Therapie der Lese-Rechtschreib-Störung'. *Dtsch. Arztebl.* 107, 718–727.

Simon, G., Poirel, N., Rossi, S., Lie Lubin, A., Pineau, A., and Houdé, O. (2013). 'Dynamics of the Anatomical Changes That Occur in the Brains of Schoolchildren as They Learn to Read'. *PLoS One* 8:e81789. doi: 10.1371/journal.pone.0081789

Skeide, M. A., Kraft, I., Müller, B., Schaadt, G., Neef, N. E., Brauer, J., et al. (2016). 'NRSN1 associated grey matter volume of the visual word form area reveals dyslexia before school'. *Brain* 139, 2792–2803. doi: 10.1093/brain/aww153

Stock, C., and Schneider, W. (2008). *Deutscher Rechtschreibtest für das erste und zweite Schuljahr: Deret 1-2.* Göttingen: Hogrefe.

Stone, M. (1974). 'Cross-Validatory Choice and Assessment of Statistical Predictions'. *J. R. Stat. Soc.* 36, 111–133. doi: 10.1111/j.2517-6161.1974.tb00994.x

Sui, J., Jiang, R., Bustillo, J., and Calhoun, V. (2020). 'Neuroimaging-based Individualized Prediction of Cognition and Behavior for Mental Disorders and Health: Methods and Promises'. *Biol. Psychiatry* 88, 818–828. doi: 10.1016/j.biopsych.2020.02.016

The Math Works Inc (2020). *'The Math Works, Inc. MATLAB. Version 2020b'. Computer Software*. Natick: The Math Works, Inc.

Tibshirani, R. (1996). 'Regression Shrinkage and Selection Via the Lasso'. *J. R. Stat. Soc.* 58, 267–288. doi: 10.1111/j.2517-6161.1996.tb02080.x

Torre, G.-A. A., and Eden, G. F. (2019). 'Relationships between gray matter volume and reading ability in typically developing children, adolescents, and young adults'. *Dev. Cogn. Neurosci.* 36:100636. doi: 10.1016/j.dcn.2019.100636

Varma, S., and Simon, R. (2006). 'Bias in error estimation when using cross-validation for model selection'. *BMC Bioinform.* 7:91. doi: 10.1186/1471-2105-7-91

Vehtari, A., Gelman, A., and Gabry, J. (2017). 'Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC'. *Stat. Comput.* 27, 1413–1432. doi: 10.1007/s11222-016-9696-4

Vul, E., Harris, C., Winkielman, P., and Pashler, H. (2009). 'Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition'. *Perspect. Psychol. Sci.* 4, 274–290. doi: 10.1111/j.1745-6924.2009.01125.x

Westman, E., Aguilar, C., Muehlboeck, S. J., and Simmons, A. (2013). 'Regional Magnetic Resonance Imaging Measures for Multivariate Analysis in Alzheimer's Disease and Mild Cognitive Impairment'. *Brain Topogr.* 26, 9–23. doi: 10.1007/s10548-012-0246-x

Whelan, R., and Garavan, H. (2014). 'When Optimism Hurts: Inflated Predictions in Psychiatric Neuroimaging'. *Biol. Psychiatry* 75, 746–748. doi: 10.1016/j.biopsych.2013.05.014

Wierenga, L. M., Langen, M., Oranje, B., and Durston, S. (2014). 'Unique developmental trajectories of cortical thickness and surface area'. *Neuroimage* 87, 120–126. doi: 10.1016/j.neuroimage.2013.11.010

Williams, V. J., Juranek, J., Cirino, P., and Fletcher, J. M. (2018). 'Cortical Thickness and Local Gyrification in Children with Developmental Dyslexia'. *Cereb. Cortex* 28, 963–973. doi: 10.1093/cercor/bhx001

Winkler, A. M., Kochunov, P., Blangero, J., Almasy, L., Zilles, K., Fox, P. T., et al. (2010). 'Cortical thickness or grey matter volume? The importance of selecting the phenotype for imaging genetics studies'. *Neuroimage* 53, 1135–1146. doi: 10.1016/j.neuroimage.2009.12.028

Xia, Z., Zhang, L., Hoeft, F., Gu, B., Gong, G., and Shu, H. (2018). 'Neural correlates of oral word reading, silent reading comprehension, and cognitive subcomponents'. *Int. J. Behav. Dev.* 42, 342–356. doi: 10.1177/0165025417727872

Zhou, W., Wang, X., Xia, Z., Bi, Y., Li, P., and Shu, H. (2016). 'Neural Mechanisms of Dorsal and Ventral Visual Regions during Text Reading'. *Front. Psychol.* 7:1399. doi: 10.3389/fpsyg.2016.01399

Ziegler, J., Montant, M., Briesemeister, B., Brink, T., Wicker, B., Ponz, A., et al. (2018). 'Do Words Stink? Neural Reuse as a Principle for Understanding Emotions in Reading'. *J. Cogn. Neurosci.* 30, 1023–1032. doi: 10.1162/jocn_a_01268

Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Faísca, L., et al. (2010). 'Orthographic Depth and Its Impact on Universal Predictors of Reading: A Cross-Language Investigation'. *Psychol. Sci.* 21, 551–559. doi: 10.1177/0956797610363406

Ziegler, J. C., Perry, C., and Zorzi, M. (2014). 'Modelling reading development through phonological decoding and self-teaching: Implications for dyslexia'. *Philos. Trans. R. Soc. B* 369:20120397. doi: 10.1098/rstb.2012.0397

Ziegler, J. C., Perry, C., and Zorzi, M. (2020). 'Learning to Read and Dyslexia: From Theory to Intervention Through Personalized Computational Models'. *Curr. Dir. Psychol. Sci.* 29, 293–300. doi: 10.1177/0963721420915873

Zilles, K., Armstrong, E., Schleicher, A., and Kretschmann, H. J. (1988). 'The human pattern of gyrification in the cerebral cortex'. *Anat. Embryol.* 179, 173–179. doi: 10.1007/BF00304699

Zou, H., and Hastie, T. (2005). 'Regularization and variable selection via the elastic net'. *J. R. Stat. Soc.* 67, 301–320. doi: 10.1111/j.1467-9868.2005.00503.x

frontiers | Frontiers in Human Neuroscience

TYPE Original Research
PUBLISHED 30 September 2022
DOI 10.3389/fnhum.2022.919465

OPEN ACCESS

EDITED BY
Aaron J. Newman,
Dalhousie University, Canada

REVIEWED BY
Kristen Pammer,
The University of Newcastle, Australia
Sara D. Beach,
University of Wisconsin-Madison,
United States

*CORRESPONDENCE
Johannes C. Ziegler
Johannes.Ziegler@univ-amu.fr

SPECIALTY SECTION
This article was submitted to
Cognitive Neuroscience,
a section of the journal
Frontiers in Human Neuroscience

RECEIVED 13 April 2022
ACCEPTED 15 September 2022
PUBLISHED 30 September 2022

CITATION
Tan Y, Chanoine V, Cavalli E, Anton J-L
and Ziegler JC (2022) Is there evidence
for a noisy computation deficit
in developmental dyslexia?
Front. Hum. Neurosci. 16:919465.
doi: 10.3389/fnhum.2022.919465

Is there evidence for a noisy computation deficit in developmental dyslexia?

Yufei Tan[1], Valérie Chanoine[2], Eddy Cavalli[3],
Jean-Luc Anton[4] and Johannes C. Ziegler [1]*

[1]CNRS, Laboratoire de Psychologie Cognitive (UMR 7290), Aix-Marseille University, Marseille, France, [2]Institute of Language, Communication and the Brain, Aix-Marseille University, Aix-en-Provence, France, [3]Laboratoire d'Étude des Mécanismes Cognitifs (EA 3082), Université Lumière Lyon 2, Lyon, France, [4]CNRS, Institut des Neurosciences de la Timone (UMR 7289), Centre IRM-INT@CERIMED, Aix-Marseille University, Marseille, France

The noisy computation hypothesis of developmental dyslexia (DD) is particularly appealing because it can explain deficits across a variety of domains, such as temporal, auditory, phonological, visual and attentional processes. A key prediction is that noisy computations lead to more variable and less stable word representations. A way to test this hypothesis is through repetition of words, that is, when there is noise in the system, the neural signature of repeated stimuli should be more variable. The hypothesis was tested in an functional magnetic resonance imaging experiment with dyslexic and typical readers by repeating words twelve times. Variability measures were computed both at the behavioral and neural levels. At the behavioral level, we compared the standard deviation of reaction time distributions of repeated words. At the neural level, in addition to standard univariate analyses and measures of intra-item variability, we also used multivariate pattern analyses (representational similarity and classification) to find out whether there was evidence for noisier representations in dyslexic readers compared to typical readers. Results showed that there were no significant differences between the two groups in any of the analyses despite robust results within each group (i.e., high representational similarity between repeated words, good classification of words vs. non-words). In summary, there was no evidence in favor of the idea that dyslexic readers would have noisier neural representations than typical readers.

KEYWORDS

dyslexia, neural noise, repetition, fMRI, neural adaptation

Introduction

Fluent reading is a critical skill for personal and professional development in all modern societies (Beddington et al., 2008). Yet, a substantial portion of children have severe and long-lasting difficulties in learning to read despite conventional instruction, normal intelligence, and adequate socio-cultural opportunities (World Health Organisation [WHO], 2010). This neurodevelopmental disorder is called developmental dyslexia (DD), which affects between 5 and 17% of children (Snowling, 2000; Demonet et al., 2004; Norton et al., 2015).

A large number of theories have been proposed to explain the causes of DD, such as theories that highlight temporal deficits (Vandermosten et al., 2010; Goswami, 2011; Casini et al., 2018), auditory deficits (Boets et al., 2007), phonological deficits (Bradley and Bryant, 1978; Baldeweg et al., 1999; Helenius et al., 1999; Snowling, 2001; Ramus et al., 2003; Ziegler et al., 2009), attentional deficits (Facoetti et al., 2000, 2006, 2008), visual deficits (Stein and Walsh, 1997; Stein, 2014), or cerebellar deficits (Nicolson et al., 2001; Nicolson and Fawcett, 2005). There is some consensus that no single-deficit theory can currently explain the multifactorial nature of the deficits observed in DD (Pennington, 2006; Perrachione et al., 2016; Ziegler et al., 2019; O'Brien and Yeatman, 2021). Thus, a more general theory is needed to explain the variety of sensory deficits associated with DD.

One elegant theory that has the potential to explain the various facets of DD is the neural noise hypothesis (Hancock et al., 2017). Neural noise can be defined as a stochastic variability in the neural response to repeated presentations of the same stimulus. For example, a neuron that spikes at widely variable intervals in response to repeated stimulus presentations is considered to be noisier than one that spikes at nearly the same time following each presentation. The link with DD seems rather straightforward: an excessive amount of neural noise impairs the capacity of populations of neurons to maintain stable patterns of activity, which is detrimental to both forming and maintaining representations.

How can one explain that excessive amounts of neural noise specifically affect learning-to-read more than other language processes? It could be argued that in most language processing situations, our brains are used to "cleaning-up" partial or incorrect bottom-up information using context (Pitt and Samuel, 1995). This is different in a learning-to-read situation. In the initial stages of reading acquisition, children have to learn the mapping between isolated graphemes and phonemes to set up the decoding network (Ziegler and Goswami, 2005; Ziegler et al., 2014, 2020). In this situation, "noisy" letter or phoneme information is detrimental because the same letter needs to map onto a single stable phoneme representation (B- > /b/and not/p/). Indeed, it has been shown in computational modeling that small amounts of noise in phoneme representations quickly result in catastrophic learning (Harm and Seidenberg, 1999; Ziegler et al., 2014, 2020; Perry et al., 2019). Similarly, small amounts of noise will also prevent the reading system from creating stable visual representations of letters and words, which will impair orthographic learning (Ziegler et al., 2014).

Strong evidence for the neural noise theory comes from studies that investigated the consistency of auditory brainstem responses to speech syllables in normal hearing children with a wide range of reading abilities (Hornickel et al., 2009; Hornickel and Kraus, 2013). The auditory brainstem response to speech closely mimics the spectrotemporal features of the stimulus. Hornickel and Kraus (2013) found that poor readers have significantly more variable auditory brainstem responses to speech than do good readers, independent of resting neurophysiological noise levels. Liebig et al. (2020) have shown that the neural stability of the auditory brainstem response to isolated syllables (e.g., /da/) measured at kindergarten predicted reading and spelling performance 2 years later.

One way to measure neural noise is through stimulus repetition. When neural noise is excessive, the neural encoding of repeated items should be more variable and neural adaptation to repeated items should be reduced. Excessive neural noise might make it difficult to establish robust short-term perceptual representations which are the basis for neural adaptation effects (Garrido et al., 2009). Indeed, Perrachione et al. (2016) showed that adults and children with dyslexia exhibited significantly diminished neural adaptation for a wide variety of repeated stimuli (spoken words, written words, visual objects, and faces). Similarly, Gertsovski and Ahissar (2022) showed no neural adaptation to repeated sounds in auditory cortex and other higher-level regions in adults with dyslexia compared to typically developing readers. This might also explain why individuals with dyslexia may have an impairment "anchoring" to consistent stimulus statistics in order to exploit sensory history for learning (Ahissar et al., 2006; Ahissar, 2007; but see Ziegler, 2008). Zhang et al. (2021) used electroencephalography (EEG) with frequency-tagging to track the temporal evolution of speech-structure learning (a structured vs. a random stream of repeated tri-syllabic pseudowords) in children with dyslexia and found that the learning of implicit speech structures built up more slowly in children with dyslexia than in typically developing readers. Studies in the visual domain reported slower perceptual decision making in individuals with dyslexia (Stefanac et al., 2021; Manning et al., 2022), which may be related to excessive perceptual noise. However, not all studies reported greater neural variability to be associated with poorer reading. In particular, Malins et al. (2018) showed a positive relationship between trial-by-trial activation variability in the left inferior frontal gyrus (IFG) pars triangularis and reading skill suggesting that greater levels of neural variability were associated with better reading skills.

A few studies focused on the effects of repetition and prediction rather than neural variability or neural adaption. Using an odd-ball paradigm, Beach et al. (2022b) recorded magnetoencephalography (MEG) as adults with and without dyslexia were passively exposed to speech syllables. In both groups, standards generated by as few as two repetitions were distinct from deviants, indicating normal sensitivity to repetition in dyslexia. However, only in the control group did standards become increasingly different from deviants with repetition. In another study Beach et al. (2022a) focused on prediction errors by presenting repeated words or faces with a high probability of stimulus repetition vs. a high probability of stimulus change. They found that the neural prediction error (as measured by EEG) was significantly weaker in dyslexia than the control group for both faces and words. These results were taken to suggest that "many of the mechanisms that give rise to neural adaptation as well as mismatch responses are intact in dyslexia, with the possible exception of a putatively predictive mechanism that successively integrates recent sensory information into feedforward processing" (Beach et al., 2022b, p. 1). Finally, Pugh et al. (2008) used an animacy judgment task (living/non-living) to investigate the effects of stimulus repetition in normal and dyslexic readers in functional magnetic resonance imaging (fMRI). In every block, six words were repeated six times in a pseudorandom fashion and intermixed with 20 unrepeated words that served as unrepeated control words. Their results showed that repetition had a similar (facilitatory) effect on reaction time and accuracy for both normal and dyslexic readers. In the critical regions of the left-hemisphere reading network, typically developing readers showed reduced activation for repeated words while dyslexic readers showed increased activation with repetition in these same reading-related sites, suggesting that the left-hemisphere reading circuitry in adolescent dyslexics is poorly tuned but not wholly disrupted.

In the present study, we wanted to test the neural noise theory more directly by investigating whether behavioral and neural responses to repeated words are more variable across repetitions for dyslexic readers. As suggested by Hancock et al. (2017), "systems-level multimodal imaging studies that measure response variability in reading disorders, such as using phase locking measures in EEG or single-trial estimates of BOLD response (...), can provide a direct test of the basic premise of our hypothesis" (p. 445). We followed their suggestion by measuring single-trial estimates of BOLD response in an fMRI experiment for 36 words that were repeated 12 times both for a group of dyslexic and typical readers. The present study was conducted with adult dyslexics who all had a history of childhood dyslexia and were referred to us from a regional clinical reference center of dyslexia with a formal diagnosis of dyslexia. Although one could argue that testing adult university students with dyslexia is suboptimal because the neural noise deficit might have been compensated for, we believe that one can make the opposite case. That is, reading compensation strategies might affect reading outcomes but they should not alleviate neural noise. Thus, it seems to be a fair comparison to investigate neural noise differences in groups that no longer show massive behavioral differences.

Variability measures were computed for the two groups both at the behavioral and neural level. At the behavioral level, we compared the standard deviation of reaction time distributions when participants read aloud the same set of words 12 times. At the neural level, we first looked at standard univariate (whole-brain) analyses and univariate region of interest analysis. In line with the neural adaptation results (Perrachione et al., 2016), we expected to find that the key regions of the reading system should be less sensitive to repetition in adults with DD than typical readers. We then looked at measures of intra-item variability in the BOLD signal to repeated words. That is, we compared the standard deviation of the beta values for multiple repetitions of the same words between the two groups. If the brain responses to repeated items were noisier in DD, dyslexic readers would show greater amounts of variability than typical readers. We then used multivariate pattern analyses, namely representational similarity (Kriegeskorte et al., 2008) and multivariate pattern classification (Pisner and Schnyer, 2020) to further explore the predictions of the neural noise hypothesis. Again, the logic was straightforward: If neural representations of written words were noisier for adults with DD, neural similarity between repeated words should be reduced in adults with DD. Similarly, if neural responses to repeated words were noisier in adults with DD, a classifier that was trained to discriminate words from hash marks should perform less well for adults with DD than for typical readers. The various predictions of the study are illustrated and summarized in **Figure 1**.

Materials and methods

Participants

Twenty adults with dyslexia and 20 skilled readers aged between 20 and 29 years participated in the present study. All participants were university students and native speakers of French. They were recruited at Aix-Marseille University (France) from a wide variety of academic programs (i.e., in each group, 55% of the participants were enrolled in social science programs and 45% were enrolled in science programs). The university students with dyslexia were recruited through a national clinical reference center of learning disorders (*Centre de Référence des Troubles des Apprentissages* at the Salvator Hospital in Marseille) and the *Mission Handicap* (University Medical Service) of Aix-Marseille University. They were all diagnosed with dyslexia in primary school and had received remedial teaching for an average of 5.34 years (SD = 0.41). Furthermore, they reported having struggled with reading from childhood to adulthood. The group of dyslexic readers (DYS) and the group of typical readers (CTR) were matched on gender

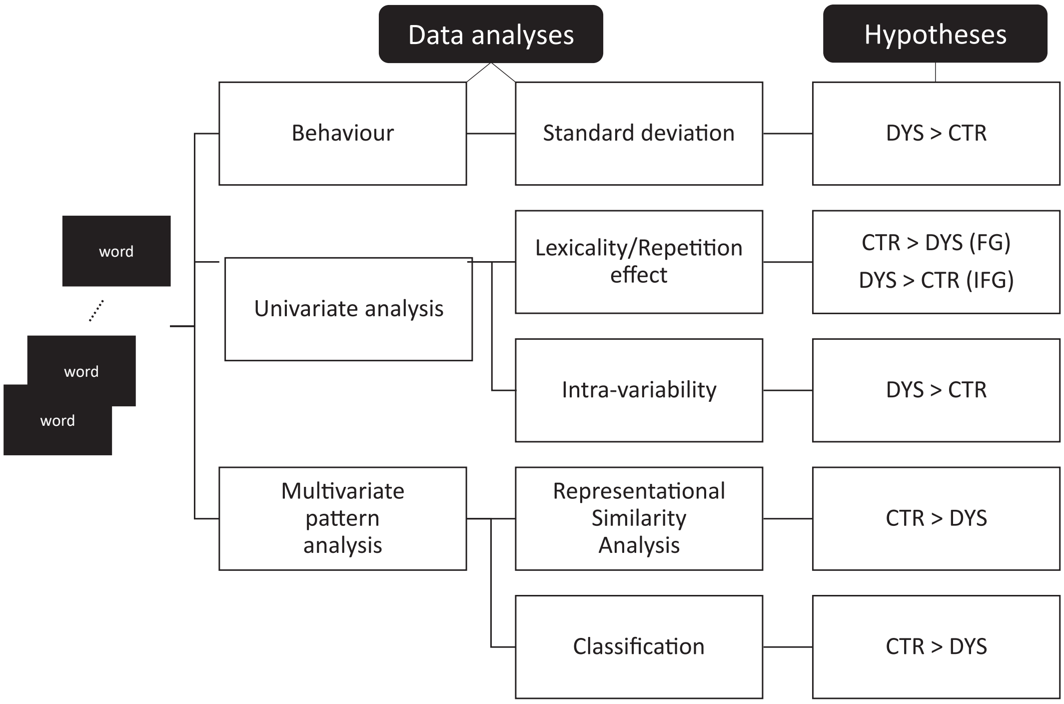


FIGURE 1
Testing the neural noise theory. Main predictions of the neural noise theory for the various analyses. DYS: dyslexic readers, CTR: typical readers, FG: fusiform gyrus, IFG: inferior frontal gyrus.

(DYS: 9 females and 11 males; CTR: 11 females, 9 males), chronological age [DYS = 22.95 ± 2.56, CTR = 23.45 ± 2.42, $t(38) = -0.64$, $p = 0.53$], verbal IQ [DYS = 38.20 ± 5.15, CTR = 39.58 ± 4.26, $t(38) = -1.11$, $p = 0.28$], and non-verbal IQ [DYS = 41.60 ± 8.33, CTR = 42.20 ± 7.45, $t(38) = -0.24$, $p = 0.81$]. The study conforms to recognized standards of the World Medical Association Declaration of Helsinki and was approved by the National Ethics Committee for Biomedical Research. All participants gave written informed consent and received €50 for their participation.

Stimuli and procedure

Reading level assessment

The reading level of the participants was assessed with two standardized reading tests. The first was the Adult Reading History Questionnaire (ARHQ, Lefly and Pennington, 2000), which is a self-report questionnaire used to diagnose the history of reading difficulties, which includes items on reading habits, reading and spelling abilities, reading speed, attitudes toward school and reading, additional assistance received, repeating grades or courses and effort required to succeed, separately from elementary school, secondary school, post-secondary education, and current life (Deacon et al., 2012). Participants answer each item on a 5-point Likert-type scale ranging from 0 to 4. The total score is divided by the maximum possible score (92), resulting in a proportion score ranging from 0 to 1. Higher scores indicate greater reading difficulties. Norms are available from an adult sample of 1,107 participants (Fichten et al., 2014). In addition, they performed the Alouette reading test (Lefavrais, 2005), which is a sensitive standardized reading fluency test for adults with dyslexia (Cavalli et al., 2017a) with excellent psychometric properties (Bertrand et al., 2010). Norms are available from an adult sample of 164 participants (Cavalli et al., 2017a). The critical variable was reading efficiency (CTL) using the following equation: $CTL = A \times 180/RT$, where A is the number of words correctly read (self-corrections included) and RT is the reading time.

Reading aloud task (in scanner)

For the reading aloud task, we selected 34 French words with frequencies ranging from 1 to 125.8 per million (Mean = 26.03, SD = 38.37) (New et al., 2004), and lengths ranging from 2 to 10 letters (Mean = 6.79, SD = 1.70). The words were presented in 40-point, Arial font. While 34 hash mark combinations were used as baseline, the number of each hash mark combination was matched with the length of each word. All stimuli were presented in white against a black background and subtended about 1.4° of visual angle for each letter.

The stimuli were presented using an in-house software developed in the NI LabVIEW environment (Bitter et al., 2017).

The software was launched and real-time synchronized with the MR acquisition using a NI-PXI 6289 digital input/output hardware, which also allowed vocal and motor answers recording. The participants lied in the MRI scanner and the stimuli were projected through a mirror onto a screen (1024 × 768), the 768 × 768 square field of view covered a 20° FOV angle. Each trial started with a fixed cross presented at the center of the screen for 340 ms, after a blank of 680 ms, a word was displayed for 680 ms, and the participants were instructed to read the words aloud while ignoring the hash marks (see **Figure 2**). The inter-trial interval jittered from 544 to 1,564 ms. There were 4 runs for each participant, each run was composed of 136 trials made of 34 words repeated three times and 34 hash mark combinations. The trials in each run were presented pseudo randomly and the order of 4 runs was counterbalanced across participants. Along with the fMRI signal, participants' answers were recorded together using the FOMRI-II microphone (Optoacoustics Ltd., Or-Yehuda, Israel).

Localizer task (in scanner)

The localizer task was adapted from a 5-min fast acquisition procedure designed by Pinel et al. (2007), which has shown to successfully activate the brain regions of auditory and visual perception, motor system, reading, language comprehension and mental calculation at an individual level. Eight types of stimuli were used: flashing horizontal checkerboards, flashing vertical checkerboards, visual motor instructions, auditory motor instructions, visual sentences, auditory sentences, visual subtraction, auditory subtractions. Each type of stimuli was presented in 10 successive trials in a randomized order for each participant. Flashing checkerboards were presented for 136 ms on each trial and participants were instructed to passively view the checkboards. For visual stimuli, each trial was composed of four successive screens of 272 ms and participants gave their responses by pressing the left or right button. For auditory stimuli, each trial lasted between 2,040 and 3,672 ms and participants also gave their responses by pressing the left or right button. The inter-trial interval ranged from 408 to 6,528 ms. The presentation of visual and auditory stimuli and the recording of responses were done in the same way as in the reading aloud task.

Data acquisition

The MRI data were acquired on a SIEMENS MAGNETOM Prisma 3T scanner with a 64-channel head coil at the Centre IRM-INT@CERIMED (INT, UMR 7289 CNRS–Aix-Marseille University). The functional images were acquired using a T2*-weighted gradient-echo planar sequence with 54 slices per volume [repetition time (TR) = 1.224 s, echo time (TE) = 30 ms, multi-band acceleration factor = 3, voxel size = 2.5 mm × 2.5 mm × 2.5 mm, flip angle = 65°, field of view (FOV) = 210 mm × 210 mm, matrix size = 84 × 84]. The anatomical image was acquired using high-resolution structural T1-weighted image with 256 slices (TR = 2.4 s, TE = 2.28 ms, voxel size = 0.8 mm × 0.8 mm × 0.8 mm, flip angle = 8°, FOV = 256 mm × 256 mm, matrix size = 320 × 320). A fieldmap acquisition (54 slices per volume, TR = 7.06 s, TE = 59 ms, voxel size = 2.5 mm × 2.5 mm × 2.5 mm, flip angle = 90°, FOV = 210 mm × 210 mm, matrix size = 84 × 84) was collected to estimate and correct the B0 inhomogeneity. A total of 1096 functional scans were acquired over 4 runs (5.59 min per run) for the reading task, and 256 functional scans were acquired in one run (5.22 min) for the localizer task.

Data analyses

Behavioral data analysis

Because the reading aloud responses were recorded in the scanner, the wave files had to be denoised in order to determine the onset time. Denoising was performed using the Wavelet Signal Denoiser toolbox of Matlab.[1] We then used the Praat software (Boersma, 2001) to determine the onset of each reading aloud response (RT) and judge whether the word pronounced correctly. The data from one participant was excluded from the behavioral data analysis because of missing data.

Then, we calculated the mean accuracy and mean RT of each group. We used a two-factorial analysis of variance (ANOVA) with group (DYS vs. CTR) and repetition (12 repetitions) as factors to test for significant differences between the two groups. The assumption of sphericity was checked by Mauchly's test, and if it was violated, the Greenhouse–Geisser correction was used to correct the *F*-test results.

The standard deviation of the RT distribution across 12 repetitions for each word was measured to analyze the variability, and a two-sample *t*-test was conducted to analyze the differences in RT variability between the two groups.

Univariate analysis of the functional magnetic resonance imaging data

Whole-brain analysis

The fMRI data was preprocessed and analyzed using SPM12 software (Wellcome Institute of Cognitive Neurology, London, UK).[2] First, we used the fieldmap images to measure field inhomogeneities, then the functional (EPI) images were realigned using the fieldmap for distortion and motion correction. The anatomical (T1) images were coregistered to the mean image of realigned EPI images. The coregistered T1 image was segmented into Gray Matter (GM), White Matter (WM), CerebroSpinal Flux (CSF), Bone tissue and Soft

---

[1] https://www.mathworks.com/help/wavelet/ref/waveletsignaldenoiser-app.html

[2] https://www.fil.ion.ucl.ac.uk

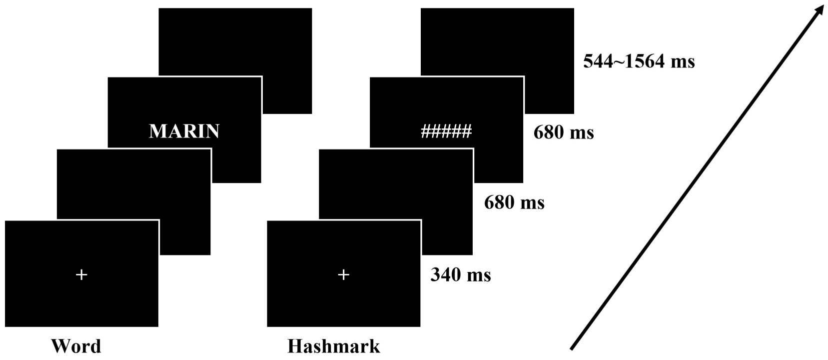


FIGURE 2
Illustration of the experimental procedure and timing in the reading aloud task that was performed in the scanner.

tissue, and normalized into standard Montreal Neurological Institute (MNI). Finally, the realigned EPI images were also normalized into MNI space using the deformation field image obtained during the anatomical normalization process, and spatially smoothed with a 5 mm full-width at half-maximum (FWHM) isotropic Gaussian kernel. The explicit masks included GM, WM, and CSF.

Prior to the first-level analysis, EPI images were denoised by GLMdenoise toolbox (Kay et al., 2013). The functional data in the first-level models were high pass filtered with a cut-off of 128 s and corrected for autocorrelation by an autoregressive model of order 1. A general linear model (GLM) in SPM12 was used to estimate the parameters. There were six regressors for the experimental conditions in each run, three for word repetitions and three for hash mark repetitions, and one regressor for runs. The duration of each event was 1.222 s. The onset and duration of each stimulus were convolved with the canonical hemodynamic response function and modeled as regressors in the design matrix.

T-contrast maps were computed separately for the lexicality effect and the word repetition effect using a voxel-based random effect analysis (RFX). The lexicality effect was obtained by subtracting activation in the control condition (hash marks) from activation in the word condition. The word repetition effect was measured by identifying regions exhibiting a change in BOLD responses across three repetitions in each of the four runs that fit a linear function (i.e., 1 × 1st repetition, 0 × 2nd repetition, −1 × 3rd repetition). The contrast maps from the first-level analyses were used to conduct the second-level one-sample t-test to test for significant group differences for the two effects. The activation areas were labeled using the Anatomical Automatic Labeling (AAL) package (Tzourio-Mazoyer et al., 2002).

Univariate region of interest analysis

For the univariate analysis on ROIs, we chose eight anatomical ROIs that are typically reported in studies of normal and impaired reading (Paulesu et al., 2014; Martin et al., 2015; Rueckl et al., 2015): left and right IFG, left and right fusiform gyrus (FG), left and right angular gyrus (AG), and left and right supramarginal gyrus (SMG). In addition, we used dorsal extrastriate cortex (hOC3d) as a purely visual control area (Kujovic et al., 2013). These ten anatomical ROIs were generated from SPM Anatomy toolbox (Eickhoff et al., 2005) and the WFU_PickAtlas.[3] They are shown in the **Supplementary Figure 1**.

The preprocessing was similar to the univariate whole-brain analysis. However, the images were not spatially normalized or smoothed to take advantage of high spatial-frequency pattern information within each participants' data (Kriegeskorte et al., 2006). They were also denoised using the GLM Denoise toolbox. All ROIs were converted into the native space of each participant using the inverse of the transformation matrix that was used to normalize the T1 image into the standard MNI space.

For a given ROI mask, we extracted each subject's percent signal change[4] using "mean" calculation across voxels. For each effect ("Lexicality" and "Repetition"), we obtained a matrix of percent signal changes per subject (n = 20) and per ROI (n = 10). The outliers (values that were greater than 2.5 standard deviations above or below the median) in a given ROI were replaced by the mean computed across subjects.

For each ROI, we performed one-tailed permutation tests[5] to compare the distribution of the percent signal changes of a given condition ("Lexicality" or "Repetition") to the null hypothesis (normal distribution) within a group of subjects or between the two groups. Statistical tests were conducted using 2000 permutations and two types of multiple comparisons, False Discovery Rate (FDR, Benjamini and Hochberg, 1995) and Bonferroni's (Bland and Altman, 1995).

---

[3] https://www.nitrc.org/projects/wfu_pickatlas/

[4] http://marsbar.sourceforge.net/faq.html#how-do-i-extract-percent-signal-change-from-my-design-using-batch

[5] https://github.com/lrkrol/permutationTest

Intra-item variability

For the intra-item variability analysis, preprocessing was identical to the univariate ROI analysis. The same words were repeated three times in each run. To make a reasonable comparison with hash marks that were always the same except that they varied in length, we selected five hash marks that repeated more than three times per run and we extracted only the first three repetitions in each run. The activation of each single trial (34 words × 3 repetitions and 5 hash marks × 3 repetitions for 4 runs, 468 trials in total) was estimated using the Least Squares All (LSA) model (all trials are estimated simultaneously in a single model) following the methodology of Mumford et al. (2014). A GLM in SPM12 was used to estimate the parameters. There were 117 regressors in each run, including 102 regressors for each word and 15 regressors for each hash mark, and also one regressor for runs. The duration of each event was 1.222 s. The onset and duration of each stimulus were convolved with the canonical hemodynamic response function and modeled as regressors in the design matrix. We therefore obtained 468 beta maps for each participant.

The individual and functional regions of interest (ROIs) were obtained from the localizer task. Because the functional data were not normalized or smoothed, we did the same for the localizer data. Participant specific contrasts of "reading sentences" vs. "flashing checkboards" were calculated to identify the reading network of each participant (Pinel et al., 2007). Only the voxels that were active at a voxel-wise statistical threshold of $p < 0.001$ (without correction for multiple comparisons) were included in the individual functional ROI. The 10 anatomical ROIs were the same as those used in the univariate ROI analysis. We then extracted the data from all the images masked with each ROI for each participant.

Given that each word was repeated for three times in four runs, we had 12 masked beta maps for each word. Because some of the beta values of some voxels were outliers, we replaced these extreme values (beta values greater than 2.5 standard deviations above or below the mean) in a masked beta map by the mean beta value of all voxels. We then calculated the standard deviation of the 12 beta maps voxel by voxel in a given ROI for each word, then averaged all the standard deviations in this ROI to obtain the mean variability of each word. Finally, the mean variabilities of the 34 words were averaged to obtain the mean variability of each participant. In order to compare the mean variability between the dyslexic and control groups, we performed a two-sample t-test. These analysis steps were repeated for each ROI.

Multivariate pattern analysis for functional magnetic resonance imaging data

Representational similarity analysis

For the representational similarity analysis, the preprocessing and first-level analysis of the fMRI data were the same as the intra-item variability analysis. We also used the same functional localizer ROI and the same anatomical ROIs. Thus, the analysis was based on 12 masked t maps for each word and hash mark.

The representational dissimilarity matrix (RDM) was obtained by measuring the correlation distance between each pair of conditions, i.e., 1 min the linear correlation between neural patterns of two conditions (Haxby et al., 2001; Aguirre, 2007; Kiani et al., 2007) which characterizes the dissimilarity between different activity patterns. In our case, we wanted to know if neural similarity between repeated words in adults with DD would be weaker than in typical readers. Thus, we measured the correlation distance between each word repetition, which resulted in a 12 × 12 repetition representational dissimilarity matrix for each word. The RDM of each word across the 12 repetitions within a given ROI was calculated using the CosMoMVPA toolbox (Oosterhof et al., 2016). In order to reduce the influence of the differences across runs, the RDM was subtracted from the mean of the entire matrix. Then the demeaned RDMs of all the words were averaged to get the mean RDM of each subject. For the statistical group analysis, we averaged all the values in the lower triangular part of the mean RDM leaving out the diagonal as the mean dissimilarity of 12 word repetitions for each subject. A two-sample t-test was performed to assess the differences between the dyslexic and the control groups. These analysis steps were repeated for each ROI.

Support vector machine classification of words and hash marks

For the classification analysis, preprocessing of fMRI data was the same as before, i.e., the images were not spatially normalized or smoothed. The explicit masks used for the statistical analysis included gray matter, white matter, and cerebrospinal flux.

The first-level analysis was different from the other analysis, because we used all the hash marks here. The activation of each single trial (34 words × 3 repetitions and 34 hash marks for 4 runs, 544 trials in total) was estimated using the Least Squares All (LSA) model (all trials are estimated simultaneously in a single model). A GLM in SPM12 was used to estimate the parameters. There were 136 regressors for all experimental conditions in each run, including 102 regressors for each word and 34 regressors for each hash mark, and also one regressor for runs. The duration of each event was 1.222 s. The onset and duration of each stimulus were convolved with the canonical hemodynamic response function and modeled as regressors in the design matrix. T-contrast maps were computed separately for each trial using a voxel-based random effect analysis (RFX). We therefore obtained 544 t-contrast maps for each participant.

We used Nilearn,[6] a Python package of machine learning for neuroimaging data (Pedregosa et al., 2011) to classify words and hash marks for each participant. We used supervised learning

---

[6] https://nilearn.github.io

and cross-validation. That is, the model was first trained with the labeled data and then tested on new unlabeled data to predict the labels. A Support Vector Machine (SVM) classifier with linear kernel was used to learn associations between data patterns and labels. In order to avoid overfitting, fourfold cross-validation was used to split data into training sets and testing sets. Because our data were imbalanced in the distribution of the target classes (408 words vs. 136 hash marks), Stratified Shuffle Split iteration was used to ensure that relative class frequencies were approximately preserved in each train and validation fold. Stratified Shuffle Split can create splits by preserving the same percentage for each target class as in the complete set. Classification performance was quantified by measuring the area under the receiver operating characteristic (ROC) curve (i.e., ROC-AUC score), which avoids inflated performance estimates for imbalanced datasets.

Permutation testing was used to evaluate the significance of the cross-validated score. The *p*-value approximates the probability that the score would be obtained by chance. It is calculated as (C + 1)/(n_permutations + 1), where C is the number of permutations whose score is greater than or equal to the true score. The n-permutation was set to 1000. Thus, the best possible *p*-value is 1/(n_permutations + 1) = 0.00099 and the worst is 1.0. We then performed a two-sample *t*-test to compare the ROC-AUC scores for the two groups.

Results

Reading level assessment

As expected, the results on the ARHQ showed that the ARHQ score of the dyslexic group was significantly higher than that of the control group [DYS = 0.58 ± 0.08, CTR = 0.33 ± 0.08, Cohen's $d = 3.12$, $t(38) = 9.57$, $p < 0.001$]. The results of the standardized reading test (Alouette) showed that dyslexic group obtained significantly lower scores than the control group [DYS = 368.79 ± 73.00, CTR = 493.03 ± 60.73, Cohen's $d = -1.85$, $t(38) = -5.85$, $p < 0.001$]. When compared to the published norms of these two tests, the scores of our sample of dyslexic readers were 2.0 standard deviations above the published norms on the ARHQ (Fichten et al., 2014) and 2.2 standard deviations below the published norms the Alouette test (Cavalli et al., 2017a).

Reading aloud task (in the scanner)

The mean accuracy was at ceiling with 99.9% for the controls and 99.4% for the dyslexic readers. We therefore analyzed only reading aloud latencies (RTs). For each participant, outliers with 2.5 standard deviation above and below the mean RT were deleted. There was no significant difference in the number of outliers between two groups [DYS = 8.65 ± 2.87, CTR = 8.74 ± 3.26, $t(37) = 0.09$, $p = 0.93$]. The mean RT for the dyslexic group was 621 ms and that for the control group was 552 ms. The results of ANOVA showed a significant main effect of group [$F_{(1,37)} = 5.373$, $p = 0.026$] and a significant main effect of repetition [$F_{(11,407)} = 2.863$, $p = 0.001$]. However, the interaction effect between group and repetition was not significant [$F_{(11,407)} = 0.810$, $p = 0.630$]. The results are plotted in **Figure 3A**. Because the strongest repetition effects were obtained in the first three repetitions (i.e., the first run), we repeated the ANOVA with the first three repetitions only. The results were identical to the previous analysis with a significant main effect of group [$F_{(1,37)} = 6.143$, $p = 0.018$], a significant main effect of repetition [$F_{(2,74)} = 9.906$, $p = 0.001$] and no significant interaction between the effects of group and repetition [$F_{(2,74)} = 1.419$, $p = 0.248$].

In order to analyze the differences in RT variability across the two groups, we calculated the standard deviation across the 12 repetitions for each word. However, because the mean RTs of two groups were significantly different, we used the coefficient of variation (i.e., the ratio of the standard deviation to the mean) instead of the standard deviation across the 12 repetitions for each word. We then averaged the coefficients of variation across the 12 repetitions of each item for each participant. As can be seen in **Figure 3B**, there was no significant difference between two groups on the mean coefficient of variation [$t(38) = -0.599$, $p = 0.553$]. The results for each participant are shown in **Figure 3C**. In sum, there is no evidence for greater variability for repeated words in adults with dyslexia as compared to typical readers.

Univariate analysis of the functional magnetic resonance imaging data

Whole-brain analysis

We first analyzed whether there were any differences between the two groups in response to words. We therefore contrasted words against hash marks for each participant using a voxel-based random effect analysis (RFX). We then performed a one-sample *t*-test to obtain the mean activation of each contrast in each group, which were labeled using the Anatomical Automatic Labeling (AAL) package. Finally, we compared the activation of each contrast between two groups with a two-sample *t*-test. All reported results use an uncorrected voxel-wise statistical threshold of $p < 0.001$, and a cluster-wise threshold of $p < 0.05$ corrected for multiple comparisons over the whole brain. Correction for multiple comparisons was based on Random Field Theory as implemented in the SPM12 software (Nichols, 2012).

For typical readers, we found word-specific activation mainly in bilateral superior temporal gyrus, left post-central gyrus, left thalamus, right rolandic operculum, right middle

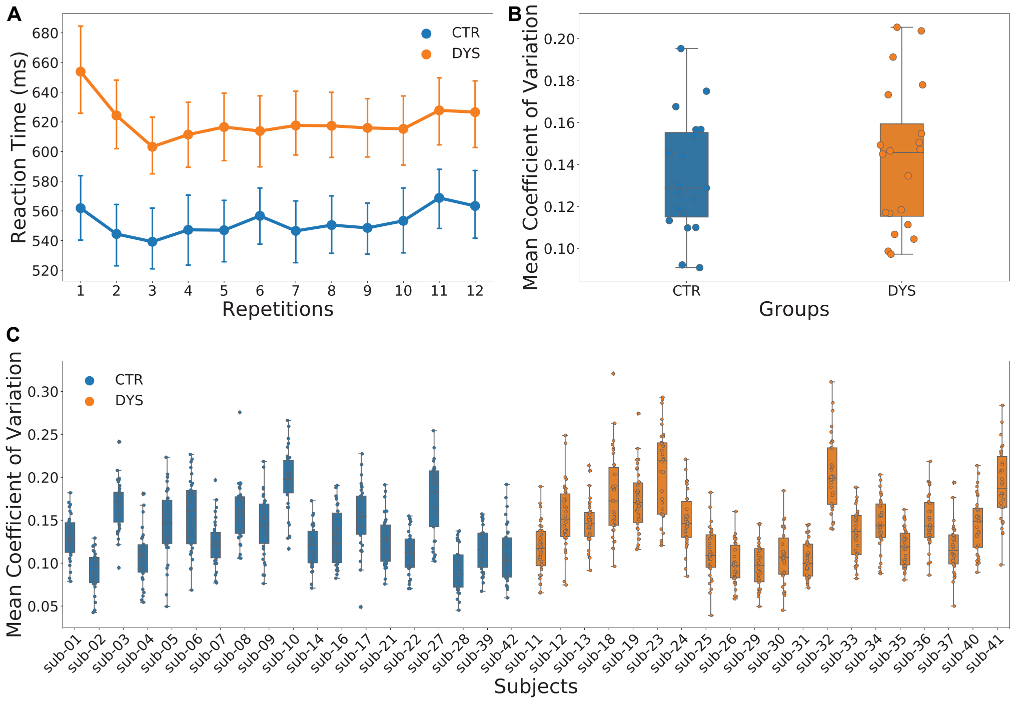


FIGURE 3

Results of reading aloud latencies. **(A)** Mean reaction times (RTs) across the twelve repetitions for both typically developing (CTR) and dyslexic readers (DYS), **(B)** scattered box plots of mean coefficients of variation of RTs for both CTR and DYS, **(C)** scattered box plots of mean coefficients of variation of RTs calculated for each word across the 12 repetitions and for each participant.

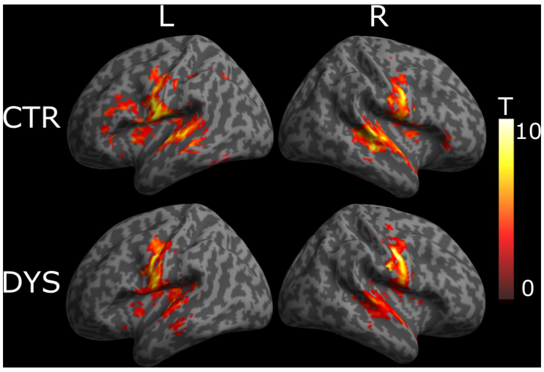


FIGURE 4

Univariate whole-brain results of the Lexicality effect (words minus hash marks). Statistical t-maps for dyslexic readers (DYS, $n = 20$) and controls (CTR, $n = 20$) are projected on **left (L)** and **right (R)** cortical surfaces (from MNI standard human cortex) using an uncorrected voxel-wise threshold of $p < 0.001$ and a cluster-wise threshold with FDR correction of $q < 0.05$.

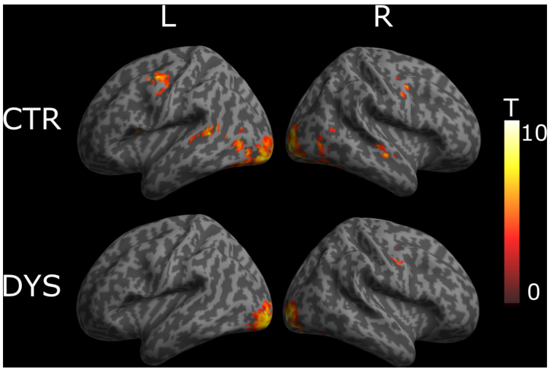


FIGURE 5

Univariate whole-brain results of the deactivation/repetition effect (a decreased activation across repetitions that fit a linear function). Statistical t-maps for dyslexic readers (DYS) and controls (CTR) are projected on **left (L)** and **right (R)** cortical surfaces (from MNI standard human cortex) using an uncorrected voxel-wise threshold of $p < 0.001$ and a cluster-wise threshold with FDR correction of $q < 0.05$.

temporal gyrus, and right cerebellum (see the upper part of Figure 4). For dyslexic readers, we found word-specific activation mainly in bilateral superior temporal gyrus, bilateral precentral gyrus, left putamen, right pallidum and bilateral cerebellum (see the lower part of Figure 4). The full list of activation clusters is presented in Supplementary Tables 1, 2 for controls and dyslexic readers, respectively. The two-sample t-test did not show any significant differences between the two groups (at a voxel-wise statistical threshold of $p < 0.001$ without

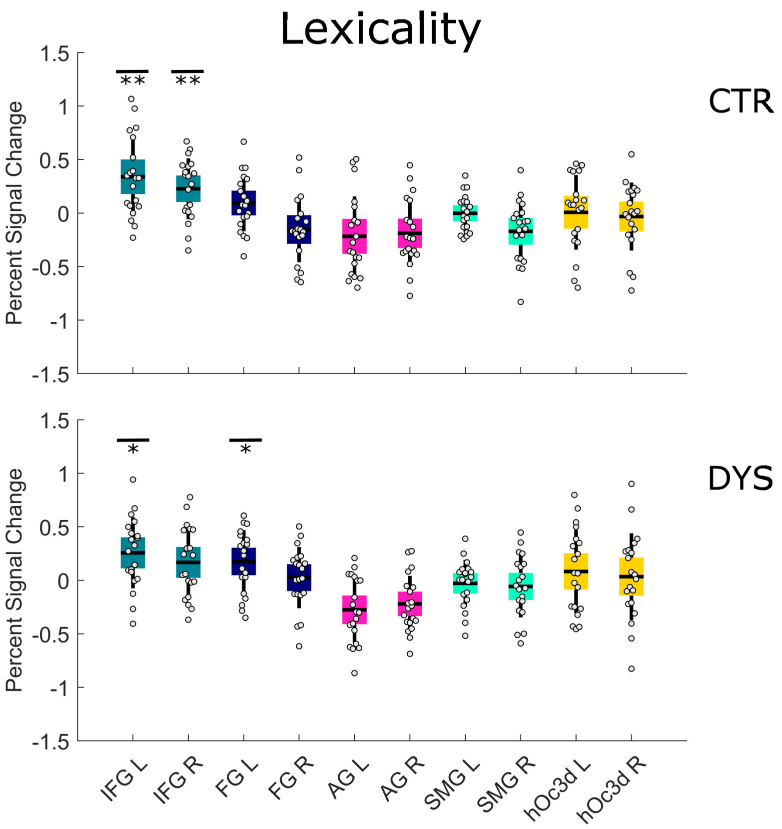


FIGURE 6
Univariate ROI results of the Lexicality effect (words minus hash marks). Percent signal change is displayed for each group (DYS and CTR) for each of the 10 predefined ROIs. The asterisks indicate significant effects *$p < 0.05$ or **$p < 0.01$ after correcting for multiple comparisons using FDR. DYS: dyslexic readers, CTR: typical readers, IFG: inferior frontal gyrus, FG: fusiform gyrus, AG: angular gyrus, SMG: supramarginal gyrus, hOc3d: Dorsal extrastriate cortex, L: left hemisphere, R: right hemisphere.

correction, and a cluster-wise threshold of $q < 0.05$ with FDR correction).

We then analyzed whether dyslexic readers showed weaker repetition (deactivation) effects than controls. For this purpose, we calculated repetition effect contrasts, which identify regions exhibiting a change in BOLD responses across three repetitions that fit a linear function. The results are shown in **Figure 5**. In typical readers, we found a linear deactivation/repetition effect in left precentral gyrus, bilateral lingual gyrus, bilateral inferior temporal gyrus, right middle occipital gyrus, left cerebellum, left inferior temporal gyrus and left middle temporal gyrus (see the upper part of **Figure 5**). For dyslexic readers, we found a linear deactivation/repetition effect in right lingual gyrus, left inferior occipital gyrus, left middle occipital gyrus, left fusiform gyrus, left cerebellum (see the lower part of **Figure 5**). The full list of activation clusters is presented in **Supplementary Tables 3, 4** for controls and dyslexic readers, respectively. The results of a two-sample *t*-test showed no significant differences between the two groups (at an uncorrected voxel-wise statistical threshold of $p < 0.001$, and a cluster-wise threshold of $q < 0.05$ with FDR correction).

Univariate region of interest analysis

We performed an additional univariate ROI analysis using the 10 ROIs mentioned before. Because the behavioral data showed larger repetition effect in the first run, we focused on a ROI analysis of the first run to make sure we would not miss a potential effect. The analysis for all runs is found in the **Supplementary material**.

For the first run, we found a significant lexicality effect in the left IFG ($p = 0.005$ with FDR correction) and the right IFG ($p = 0.005$ with FDR correction) for the control group (see the upper part of **Figure 6** and **Supplementary Table 5**). For the dyslexic group, we found a significant lexicality effect in the left IFG ($p = 0.015$ with FDR correction) and the left FG ($p = 0.037$ with FDR correction) (see the lower part of **Figure 6** and **Supplementary Table 5**). A two-sample *t*-test showed that dyslexic readers exhibited a significantly greater lexicality effect than typical readers only in the right FG ($p = 0.033$ uncorrected) (see the **Supplementary Table 5**).

As concerns the repetition deactivation effect, we found a significant deactivation in the right FG ($p = 0.030$ with FDR correction) for the control group (see the upper part of **Figure 7**

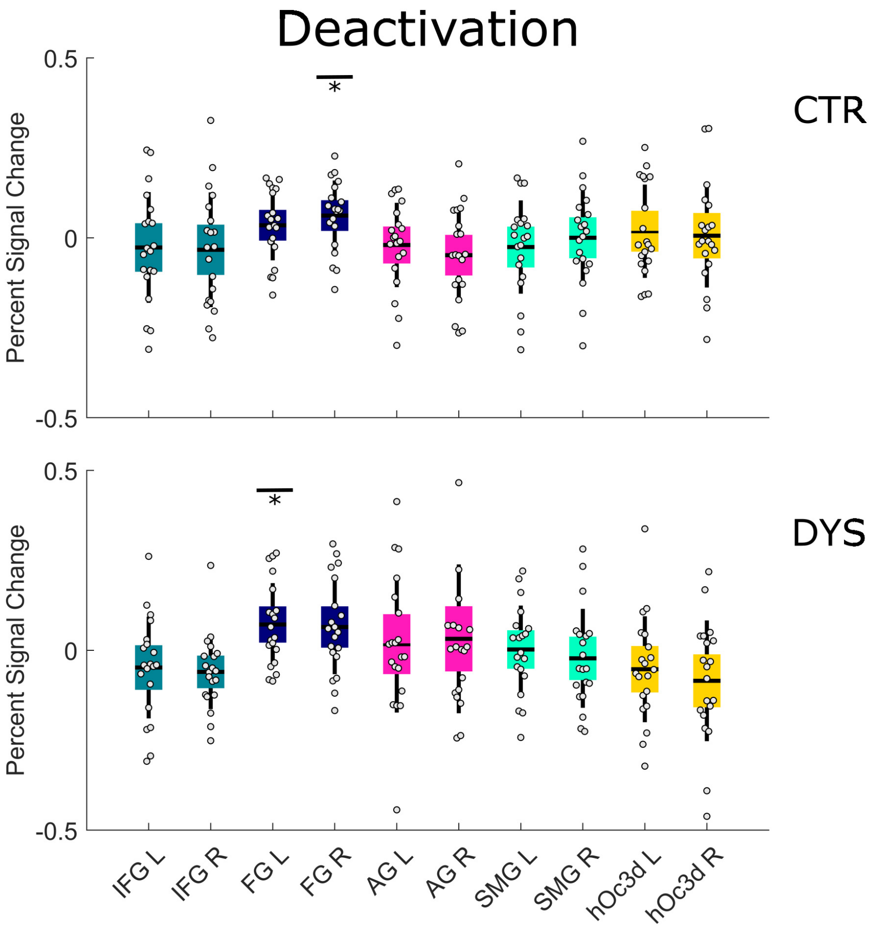


FIGURE 7
Univariate ROI results of the deactivation repetition effect (regions that show a linear decrease in activation). Percent signal change is displayed for each group (DYS and CTR) and for each of the ten predefined ROIs. The asterisks indicate significant effects *$p < 0.05$ after correcting for multiple comparisons using FDR. DYS: dyslexic readers, CTR: typical readers, IFG: inferior frontal gyrus, FG: fusiform gyrus, AG: angular gyrus, SMG: supramarginal gyrus, hOc3d: Dorsal extrastriate cortex, L: left hemisphere, R: right hemisphere.

and **Supplementary Table 6**) and in the left FG ($p = 0.050$ with FDR correction) for the dyslexic group (see the lower part of **Figure 7** and **Supplementary Table 6**). There was no significant difference between the two groups in any of the ROIs (see the **Supplementary Table 5**). The results for all runs were similar to those of the first run (see **Supplementary Tables 7**, **8**).

Intra-item variability

Because each item was repeated three times in each of the four runs (12 repetitions), the intra-item variability was obtained by calculating the standard deviation of the beta maps of twelve repetitions voxel by voxel in a given ROI for each word. Note that this approach gave rise to some extreme outlier beta values (for a similar problem, see Malins et al., 2018, p. 2983). We therefore replaced the extreme values (beta values of a voxel greater than 2.5 standard deviations above or below the mean) in a masked beta map by the mean beta value of all voxels. There was no significant difference between the two groups in the number of outlier voxels that were excluded [DYS = 41878 ± 11903 (SD), CTR = 37264 ± 8103 (SD), $t(38) = 1.43$, $p = 0.16$]. The overall percentage of trimmed outlier voxels was 2.03% for the dyslexic group and 1.98% for the control group. The mean variability was then calculated by averaging all the standard deviations in a given ROI. Finally, the mean variability of each subject was calculated by averaging the mean variability of all the words. The mean intra-word variability of two groups in each ROI is shown in **Figure 8**. Potential group differences were assessed using a two-sample *t*-test for each ROI (*p*-values were corrected for multiple comparisons using FDR). The results showed no significant differences between two groups for intra-word variability in each ROI (all $ps > 0.05$).

Interestingly, intra-word variability seemed smaller in left FG than all other regions. To test for this, we conduced paired *t*-tested of left FG variability against all other ROIs. For the control group, the activity in left FG was significantly less

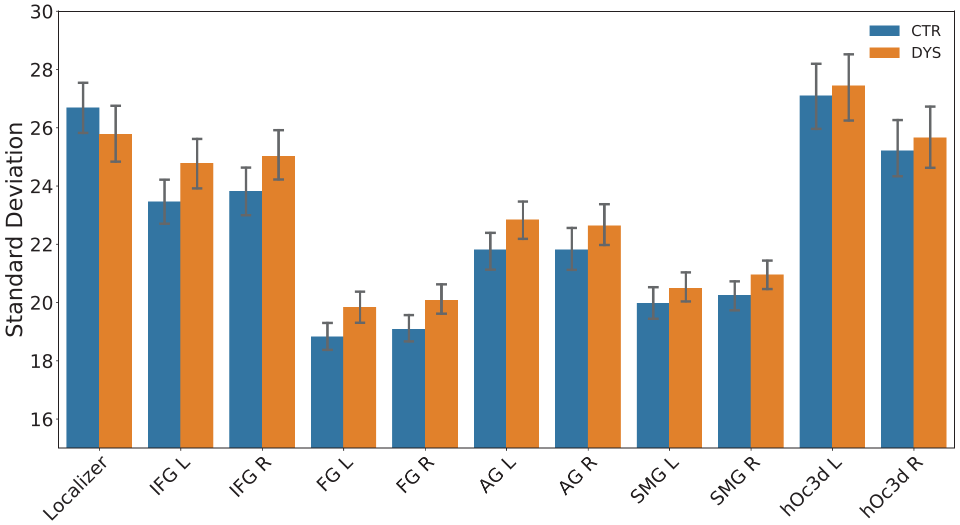


FIGURE 8
The mean intra-word variability (mean SD) across the 12 repetitions of beta values in each ROI for the control (CTR) and dyslexic (DYS) group. IFG: inferior frontal gyrus, FG: fusiform gyrus, AG: angular gyrus, SMG: supramarginal gyrus, hOc3d: dorsal extrastriate cortex, L: left hemisphere, R: right hemisphere.

variable than in all the other ROIs (all *ps* < 0.01, corrected for multiple comparisons) except right FG ($p = 0.29$). For the dyslexic group, left FG was significantly less variable than activity in the localizer, bilateral extrastriate cortex, bilateral IFG, bilateral AG, and right SMG (all *ps* < 0.01, corrected for multiple comparisons), but it was not significantly different from right FG and left SMG (*ps* > 0.1).

Multivariate pattern analysis

Representational similarity analysis

For the representational dissimilarity matrix (RDM) analyses, we calculated the RDM values across the 12 repetitions in a given ROI for each word. The mean RDM of each subject was computed by averaging the RDMs of all the words. The mean RDM of word repetitions in each ROI of two groups is displayed in **Figure 9**. For the statistical group analysis, the mean dissimilarity value for each subject was obtained by averaging all RDM values (i.e., the lower triangle of the matrix leaving out the diagonal). A two-sample t-test was used to assess the difference of dissimilarity in RDM between the dyslexic and control groups for each ROI (t-values and p-values of each ROI are shown in **Figure 9**). As above, the p-values were corrected for multiple comparisons using FDR. The results showed that there were no significant differences in the RDMs between the two groups in any of the ROIs (all *ps* > 0.05).

Support vector machine classification of words and hash marks

A Support Vector Machine (SVM) classifier was used to classify words and hash marks (see section "Materials and methods"). Classification performance was measured through the ROC-AUC score. The p-values of the permutation test were less than 0.05 in all the ROIs for most subjects. It indicates that the classifier was able to classify words and hash marks for most subjects with high accuracy. The ROC curves and mean ROC scores of the two groups in each ROI are displayed in **Figure 10**.

In order to compare the ROC-AUC scores of the two groups, a two-sample t-test was performed (t-values and p-values of each ROI are shown in **Figure 10**). It can be seen that the accuracy for the dyslexic group was significantly higher than that of the control group in left SMG ($p = 0.02$). No significant group differences were obtained in the other ROIs.

Discussion

The goal of the present study was to put the neural noise hypothesis to a direct test by investigating whether there was any evidence for excessive neural noise in adults with dyslexia when neural noise was equated with greater variability in the behavioral and neural responses to repeated presentations of the same stimulus (Hancock et al., 2017). We had participants read aloud words in an MRI scanner and the words were repeated 12 times across four runs intermixed with hash marks, which provided the baseline condition.

The behavioral and univariate results can be summarized as follows. First, the reading level assessment and the reading aloud data in the scanner clearly showed that our university students with dyslexia performed more poorly on all reading measures than the control group (weaker ARHQ scores, weaker fluency in a standardized reading test and slower RTs in reading aloud).

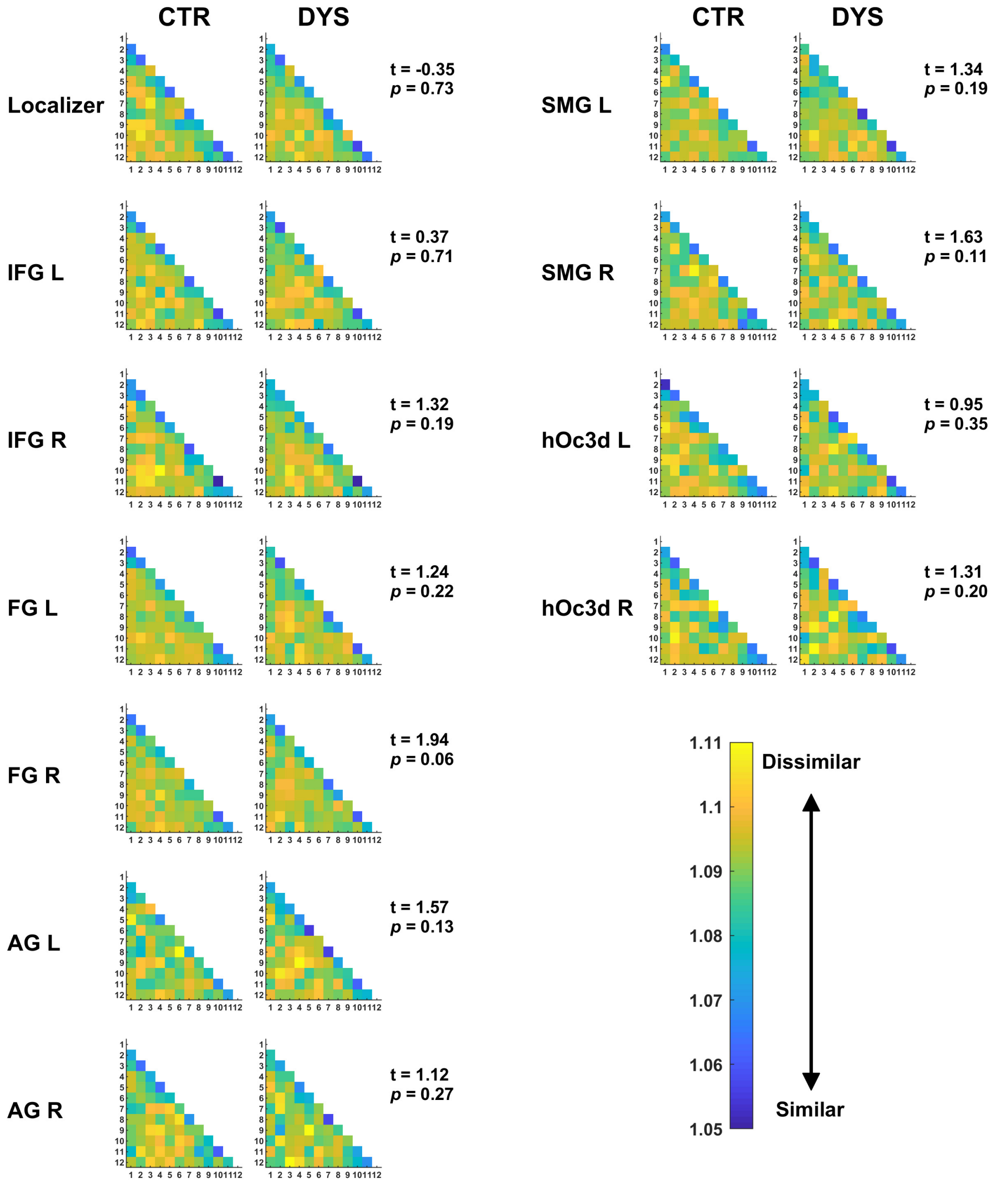


FIGURE 9

The mean RDMs for repeated words in each ROI. The values in the matrix show the dissimilarity between each pair of repetitions (the higher the value, the lower the similarity). Potential group differences were assessed with a two-sample *t*-test (see *t* and *p*-values). DYS: dyslexic readers, CTR: typical readers, IFG: inferior frontal gyrus, FG: fusiform gyrus, AG: angular gyrus, SMG: supramarginal gyrus, hOc3d: dorsal extrastriate cortex, L: left hemisphere, R: right hemisphere.

The effect sizes of the differences on the standardized tests varied between 2.1 and 3.5 standard deviations below the mean of the controls, which clearly confirms that reading performance in our group of adult dyslexic readers was still in a pathological range and this was true even when compared to normative adult samples (Fichten et al., 2014; Cavalli et al., 2017a). Thus, despite being university students, our sample of adult dyslexic readers read significantly more slowly than controls. Slow reading is a hallmark of DD in adults (Pennington et al., 1990; Lefly and Pennington, 2000; Cavalli et al., 2017b). Second, both groups showed a significant RT decrease across repetitions. However, the repetition effect was not different for the two groups, which goes against a key prediction of the neural noise hypothesis that dyslexic readers should benefit less from repetition than controls. A similar finding was reported by Pugh et al. (2008) who showed significant repetition effects but no interaction between the effects of group and repetition in a word reading paradigm. Third, in the univariate analyses, there was clear evidence for significant repetition effects (neural adaptation) in left FG in adults with dyslexia that was not different to that

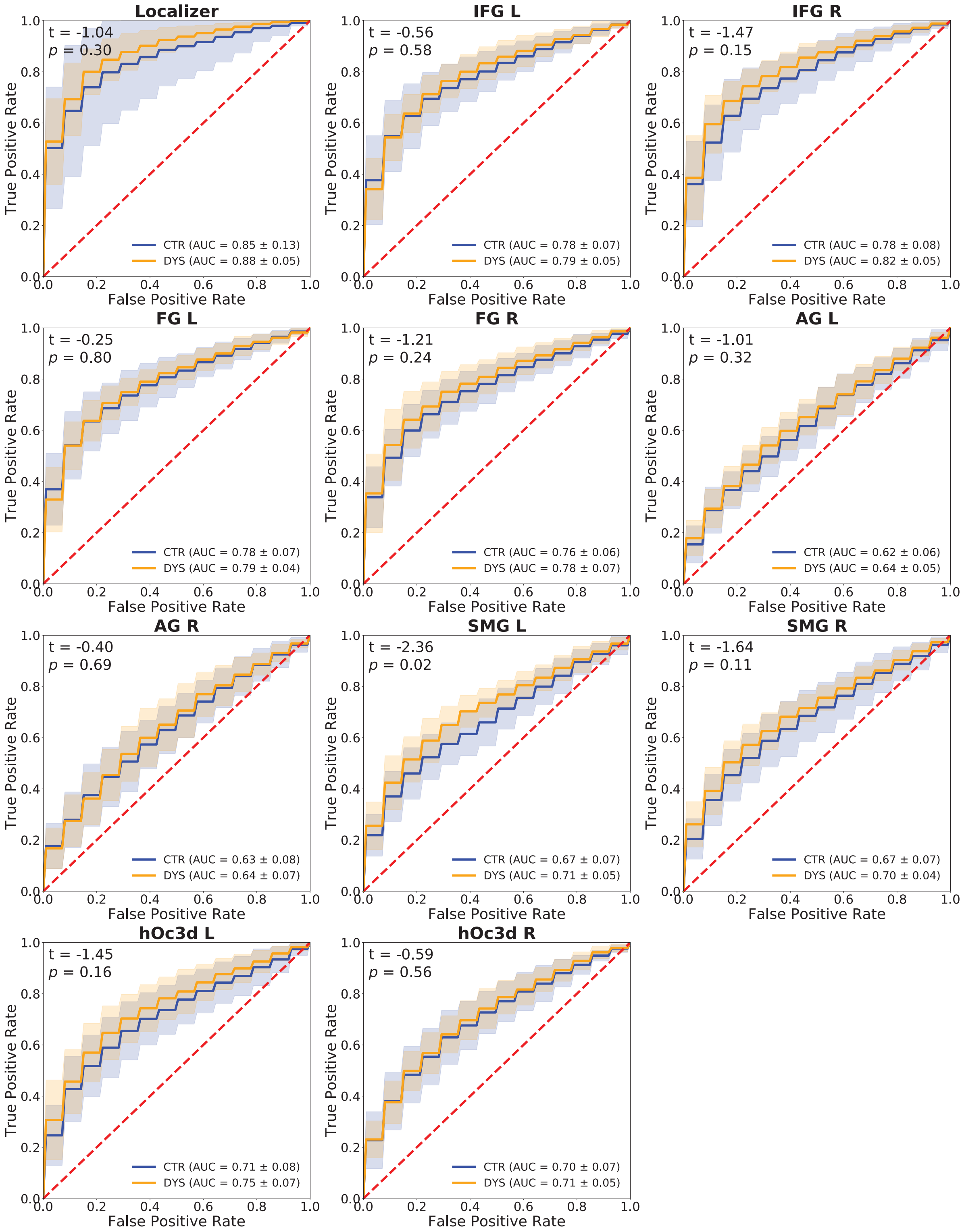


FIGURE 10

Classification performance (ROC curves) of a Support Vector Machine classifier in each ROI. The mean ROC scores of the two groups are shown in the bottom right corner, *t*-values and *p*-values of a two-sample *t*-test are shown in the upper left corner, the shaded areas correspond to ± 1 SD of the ROC curve. DYS: dyslexic readers, CTR: typical readers, IFG: inferior frontal gyrus, FG: fusiform gyrus, AG: angular gyrus, SMG: supramarginal gyrus, hOc3d: Dorsal extrastriate cortex, L: left hemisphere, R: right hemisphere.

of typical readers. The same regions also showed significant lexicality effects in dyslexics that were not different to those of controls. Similar findings have been reported by Beach et al., 2022a,b who found repetition effects in adult dyslexia that were quantitatively and qualitatively not different from those of typically developing readers.

The strongest test of the neural noise hypothesis was the intra-item variability analysis (Garrett et al., 2010; Malins et al., 2018). Because each item was repeated three times in each of the four runs (12 repetitions), the intra-item variability could be obtained by calculating the standard deviation of the beta values of 12 repetitions voxel by voxel for each word in a given ROI. Although the results showed some interesting variations of intra-item variability across the ROIs (e.g., the smallest variability was obtained in left fusiform gyrus in line with its key role as the visual word form system, see Dehaene and Cohen, 2011), there was absolutely no evidence for greater variability in the neural response to repeated words in adults with dyslexia.

Given that multiple levels of representation are involved in reading single words (visual representations of letter shape, orthographic representations of letter identity and order, phonological representations of the word's pronunciation, and semantic representations of its meaning) and they are distributed over a large reading network (Hoffman et al., 2015), we used two multivariate pattern analyses (RSA and SVM classification) that are more sensitive to the distributed nature of the information than our previous analyses (Fischer-Baum et al., 2017). The key prediction for the RSA analysis was that if neural responses to repeated words were noisier, then representational similarity across repeated words should be weaker. Although we found greater representational similarity for word repetitions that occurred with the same run than between different runs, there was again no evidence for noisier neural representations for adult dyslexic readers than controls. In our final test of the neural noise theory, we used a state-of-the art classification algorithm. In this analysis, we no longer looked at variability or similarity across repeated items, but we simply let the classifier do the classification on words vs. Hash marks on the basis of the distributed neural patterns in the data that are present in various ROIs. If the neural responses to words were noisier in adults with dyslexia, the classifier should perform more poorly for adults with dyslexia. Although group mean classification performance was good with AUC-ROC scores between 0.62 and 0.88 depending on the ROI, there were again no differences between the two groups except for superior classification of adult dyslexic readers over controls in left supramarginal gyrus. The supramarginal gyrus belongs to the dorsal route involved in phonological decoding. Superior classification performance of adult dyslexics in that ROI might suggest that they still rely to a greater extent on the less automatized dorsal route than the ventral route when reading words aloud.

A potential problem of our classification null effect is the fact that the word-hash mark classification might have been too coarse to reveal subtle differences in neural representation of word representations. That is, the differences between words and hash marks might be so big that classification performance would not be affected even if word representations were noisier in adults with dyslexia. A stronger test would have been to compute the classification of one word (all its presentations) against another word (all its presentations). However, we had too few presentations of each word to conduct this analysis. Another potential problem is that hash marks did not require a reading aloud response whereas word did. Thus, the classifier could have exploited differences in articulatory output processes to make successful classifications. However, if this were the case, we should have obtained better classification performance in Broca's area than in fusiform gyrus, which was not the case. In fact, there is little reason to believe that articulatory output processes should affect neural activation in fusiform gyrus.

Taken together, we found no evidence for increased neural noise in adults with dyslexia as defined by greater variability in the behavioral and neural responses to repeated presentations of the same stimulus. Our findings contrast with those of Perrachione et al. (2016) who found less neural adaptation in dyslexic adults for repeated words than for controls. However, the neural adaptation paradigm is very different from our paradigm because, in the neural adaptation paradigm, a single item (e.g., the word "bank") is presented eight times in a row and participants passively viewed the words. In our paradigm, words were repeated 12 times but they were intermixed with other words and hash marks and participants were asked to make an active response. It is clearly the case that our paradigm is less well suited to measure neural adaptation than that of Perrachione et al. (2016) but more work is needed to fully understand the differences between neural adaptation and repetition paradigms (Pugh et al., 2008).

At first sight, our results seem to be inconsistent with the findings of Malins et al. (2018) who showed that trial-by-trial activation variability in the left IFG pars triangularis was associated with reading skills in a sample of school-aged children. However, in their study, the correlation was positive with greater levels of neural variability being associated with better reading skills. While we did not find greater variability for typically developing readers either, it is worth noting that their results are opposite to the predictions of the neural noise hypothesis. Indeed, the authors suggest that neural variability could be beneficial in developing readers.

One explanation for why we might not have seen neural noise effects could be that the repetition of words, which was our experimental "trick" to study neural noise (variability across repetitions), actually had the opposite effect of "cleaning-up" short-term memory representations for repeated items. Thus, the massive repetition might have rendered the task too easy to tap differences in the quality of underlying representations. This is in line with the finding of Pugh et al. (2008)

who showed that after only three repetitions of the same words, the left-hemisphere reading network showed a normal response to written words in dyslexic participants. Although we cannot exclude this possibility, it is worth pointing out that our adults with dyslexia still showed slower reading aloud performance than controls even after 12 repetitions. One obvious shortcoming of the present study is that our dyslexic participants were university students, which might have compensated for their lower-level orthographic and phonological deficits by using context or higher level-linguistic information (Cavalli et al., 2017b,c). Repetition might be one of the contextual factors that is used strategically by adult dyslexic readers to compensate for their persistent low-level orthographic processing deficits. However, while it is easy to see how such compensation strategies can improve reading performance, it is less obvious to see how reading compensation could alleviate neural noise. It would be important to do a similar study with children and contrast neural adaption and neural noise paradigms. Clearly, more work is needed to put this exciting hypothesis to further test.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the Comité de Protection des Personnes (CPP). The patients/participants provided their written informed consent to participate in this study.

Author contributions

JZ and EC designed the study. VC and EC collected the data. YT and VC analyzed the data. J-LA gave advice on data analysis and interpretation. YT wrote the first draft of the manuscript. JZ, VC, and J-LA contributed to several revisions of the first manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research has been supported by grants to the Institute of Convergence ILCB (ANR-16-CONV-0002) and the Excellence Initiative of Aix-Marseille University A*MIDEX (ANR-11-IDEX-0001-02), the LABEX CORTEX (ANR-11-LABX-0042) of and the Excellence Initiative of the University of Lyon (ANR-11-IDEX-0007). YT was supported by a Ph.D scholarship of the China Scholarship Council. This scanning was performed at the IRM-INT-CERIMED Center (UMR 7289 and AMU-CNRS), which belongs to the France Life Imaging Network (grant ANR-11-INBS-0006).

Acknowledgments

We thank Stéphane Dufau for his help with denoising the wave files.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnhum.2022.919465/full#supplementary-material

References

Aguirre, G. K. (2007). Continuous carry-over designs for fmri. *NeuroImage* 35, 1480–1494. doi: 10.1016/j.neuroimage.2007.02.005

Ahissar, M. (2007). Dyslexia and the anchoring-deficit hypothesis. *Trends Cogn. Sci.* 11, 458–465. doi: 10.1016/j.tics.2007.08.015

Ahissar, M., Lubin, Y., Putter-Katz, H., and Banai, K. (2006). Dyslexia and the failure to form a perceptual anchor. *Nat. Neurosci.* 9, 1558–1564. doi: 10.1038/nn1800

Baldeweg, T., Richardson, A., Watkins, S., Foale, C., and Gruzelier, J. (1999). Impaired auditory frequency discrimination in dyslexia detected with mismatch evoked potentials. *Ann. Neurol.* 45, 495–503. doi: 10.1002/1531-8249(199904)45:4

Beach, S. D., Ozernov-Palchik, O., May, S. C., Centanni, T. M., Perrachione, T. K., Pantazis, D., et al. (2022b). The neural representation of a repeated standard stimulus in dyslexia. *Front. Hum. Neurosci.* 16:823627. doi: 10.3389/fnhum.2022.823627

Beach, S. D., Lim, S.-J., Cardenas-Iniguez, C., Eddy, M. D., Gabrieli, J. D. E., and Perrachione, T. K. (2022a). Electrophysiological correlates of perceptual prediction error are attenuated in dyslexia. *Neuropsychologia* 165:108091. doi: 10.1016/j.neuropsychologia.2021.108091

Beddington, J., Cooper, C. L., Field, J., Goswami, U., Huppert, F. A., Jenkins, R., et al. (2008). The mental wealth of nations. *Nature* 455, 1057–1060. doi: 10.1038/4551057a

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

Bertrand, D., Fluss, J., Billard, C., and Ziegler, J. C. (2010). Efficacité, sensibilité, spécificité : comparaison de différents tests de lecture [efficiency, sensitivity, specificity: comparison of different reading tests]. *Année Psychol.* 110, 299–320.

Bitter, R., Mohiuddin, T., and Nawrocki, M. (2017). *LabVIEW^{TM} Advanced Programming Techniques*. Boca Raton: CRC Press.

Bland, J. M., and Altman, D. G. (1995). Multiple significance tests: the bonferroni method. *BMJ* 310:170. doi: 10.1136/bmj.310.6973.170

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot. Int.* 5, 341–345.

Boets, B., Wouters, J., van Wieringen, A., and Ghesquiere, P. (2007). Auditory processing, speech perception and phonological ability in pre-school children at high-risk for dyslexia: a longitudinal study of the auditory temporal processing theory. *Neuropsychologia* 45, 1608–1620. doi: 10.1016/j.neuropsychologia.2007.01.009

Bradley, L., and Bryant, P. E. (1978). Difficulties in auditory organisation as a possible cause of reading backwardness. *Nature* 271, 746–747.

Casini, L., Pech-Georgel, C., and Ziegler, J. C. (2018). It's about time: revisiting temporal processing deficits in dyslexia. *Dev. Sci.* 21:e12530. doi: 10.1111/desc.12530

Cavalli, E., Cole, P., Leloup, G., Poracchia-George, F., Sprenger-Charolles, L., and El Ahmadi, A. (2017a). Screening for dyslexia in french-speaking university students: an evaluation of the detection accuracy of the alouette test. *J. Learn. Disabil.* 51, 268–282. doi: 10.1177/0022219417704637

Cavalli, E., Cole, P., Pattamadilok, C., Badier, J. M., Zielinski, C., Chanoine, V., et al. (2017b). Spatiotemporal reorganization of the reading network in adult dyslexia. *Cortex* 92, 204–221. doi: 10.1016/j.cortex.2017.04.012

Cavalli, E., Duncan, L. G., Elbro, C., El Ahmadi, A., and Cole, P. (2017c). Phonemic-morphemic dissociation in university students with dyslexia: an index of reading compensation? *Ann. Dyslexia* 67, 63–84. doi: 10.1007/s11881-016-0138-y

Deacon, S. H., Cook, K., and Parrila, R. (2012). Identifying high-functioning dyslexics: is self-report of early reading problems enough? *Ann. Dys.* 62, 120–134. doi: 10.1007/s11881-012-0068-2

Dehaene, S., and Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends Cogn. Sci.* 15, 254–262. doi: 10.1016/j.tics.2011.04.003

Demonet, J. F., Taylor, M. J., and Chaix, Y. (2004). Developmental dyslexia. *Lancet* 363, 1451–1460. doi: 10.1016/S0140-6736(04)16106-0

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., et al. (2005). A new spm toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage* 25, 1325–1335. doi: 10.1016/j.neuroimage.2004.12.034

Facoetti, A., Paganoni, P., Turatto, M., Marzola, V., and Mascetti, G. G. (2000). Visual-spatial attention in developmental dyslexia. *Cortex* 36, 109–123. doi: 10.1016/s0010-9452(08)70840-2

Facoetti, A., Ruffino, M., Peru, A., Paganoni, P., and Chelazzi, L. (2008). Sluggish engagement and disengagement of non-spatial attention in dyslexic children. *Cortex* 44, 1221–1233. doi: 10.1016/j.cortex.2007.10.007

Facoetti, A., Zorzi, M., Cestnick, L., Lorusso, M. L., Moltenia, M., Paganoni, P., et al. (2006). The relationship between visuospatial attention and nonword reading in developmental dyslexia. *Cogn. Neuropsychol.* 23, 841–855. doi: 10.1080/02643290500441841

Fischer-Baum, S., Bruggemann, D., Gallego, I. F., Li, D. S. P., and Tamez, E. R. (2017). Decoding levels of representation in reading: A representational similarity approach. *Cortex* 90, 88–102. doi: 10.1016/j.cortex.2017.02.017

Fichten, C. S., Nguyen, M. N., King, L., Havel, A., Mimouni, Z., Barile, M., et al. (2014). How well do they read? Brief English and French screening tools for college students. *Int. J. Spec. Educ.* 29, 33–46.

Garrett, D. D., Kovacevic, N., McIntosh, A. R., and Grady, C. L. (2010). Blood oxygen level-dependent signal variability is more than just noise. *J. Neurosci.* 30, 4914–4921. doi: 10.1523/jneurosci.5166-09.2010

Garrido, M. I., Kilner, J. M., Kiebel, S. J., Stephan, K. E., Baldeweg, T., and Friston, K. J. (2009). Repetition suppression and plasticity in the human brain. *NeuroImage* 48, 269–279. doi: 10.1016/j.neuroimage.2009.06.034

Gertsovski, A., and Ahissar, M. (2022). Reduced learning of sound categories in dyslexia is associated with reduced regularity-induced auditory cortex adaptation. *J. Neurosci.* 42, 1328–1342. doi: 10.1523/jneurosci.1533-21.2021

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends Cogn. Sci.* 15, 3–10. doi: 10.1016/j.tics.2010.10.001

Hancock, R., Pugh, K. R., and Hoeft, F. (2017). Neural noise hypothesis of developmental dyslexia. *Trends Cogn. Sci.* 21, 434–448. doi: 10.1016/j.tics.2017.03.008

Harm, M. W., and Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: insights from connectionist models. *Psychol. Rev.* 106, 491–528. doi: 10.1037/0033-295X.106.3.491

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., and Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293, 2425–2430. doi: 10.1126/science.1063736

Helenius, P., Uutela, K., and Hari, R. (1999). Auditory stream segregation in dyslexic adults. *Brain* 122, 907–913. doi: 10.1093/brain/122.5.907

Hoffman, P., Lambon Ralph, M. A., and Woollams, A. M. (2015). Triangulation of the neurocomputational architecture underpinning reading aloud. *Proc. Natl. Acad. Sci. U.S.A.* 112:E3719–E3728. doi: 10.1073/pnas.1502032112

Hornickel, J., and Kraus, N. (2013). Unstable representation of sound: a biological marker of dyslexia. *J. Neurosci.* 33, 3500–3504. doi: 10.1523/JNEUROSCI.4205-12.2013

Hornickel, J., Skoe, E., Nicol, T., Zecker, S., and Kraus, N. (2009). Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13022–13027. doi: 10.1073/pnas.0901123106

Kay, K., Rokem, A., Winawer, J., Dougherty, R., and Wandell, B. (2013). Glmdenoise: a fast, automated technique for denoising task-based fmri data. *Front. Neurosci.* 7:247. doi: 10.3389/fnins.2013.00247

Kiani, R., Esteky, H., Mirpour, K., and Tanaka, K. (2007). Object category structure in response patterns of neuronal population in monkey inferior temporal cortex. *J. Neurophysiol.* 97, 4296–4309. doi: 10.1152/jn.00024.2007

Kriegeskorte, N., Goebel, R., and Bandettini, P. (2006). Information-based functional brain mapping. *Proc. Natl. Acad. Sci.U.S.A.* 103, 3863–3868. doi: 10.1073/pnas.0600244103

Kriegeskorte, N., Mur, M., and Bandettini, P. (2008). Representational similarity analysis - connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* 2:4. doi: 10.3389/neuro.06.004.2008

Kujovic, M., Zilles, K., Malikovic, A., Schleicher, A., Mohlberg, H., Rottschy, C., et al. (2013). Cytoarchitectonic mapping of the human dorsal extrastriate cortex. *Brain Struct. Funct.* 218, 157–172. doi: 10.1007/s00429-012-0390-9

Lefavrais, J. (2005). *Test de l'Alouette (Revised Version)*. Paris: ECPA.

Lefly, D. L., and Pennington, B. F. (2000). Reliability and validity of the adult reading history questionnaire. *J. Learn. Disabil.* 33, 286–296. doi: 10.1177/002221940003300306

Liebig, J., Friederici, A. D., and Neef, N. E. (2020). Auditory brainstem measures and genotyping boost the prediction of literacy: A longitudinal study on early markers of dyslexia. *Dev. Cogn. Neurosci.* 46:100869. doi: 10.1016/j.dcn.2020.100869

Malins, J. G., Pugh, K. R., Buis, B., Frost, S. J., Hoeft, F., Landi, N., et al. (2018). Individual differences in reading skill are related to trial-by-trial neural activation variability in the reading network. *J. Neurosci.* 38, 2981–2989. doi: 10.1523/JNEUROSCI.0907-17.2018

Manning, C., Hassall, C. D., Hunt, L. T., Norcia, A. M., Wagenmakers, E.-J., Snowling, M. J., et al. (2022). Visual motion and decision-making in dyslexia: reduced accumulation of sensory evidence and related neural dynamics. *J. Neurosci.* 42, 121–134. doi: 10.1523/jneurosci.1232-21.2021

Martin, A., Schurz, M., Kronbichler, M., and Richlan, F. (2015). Reading in the brain of children and adults: a meta-analysis of 40 functional magnetic resonance imaging studies. *Hum. Brain Mapp.* 36, 1963–1981. doi: 10.1002/hbm.22749

Mumford, J. A., Davis, T., and Poldrack, R. A. (2014). The impact of study design on pattern estimation for single-trial multivariate pattern analysis. *NeuroImage* 103, 130–138. doi: 10.1016/j.neuroimage.2014.09.026

New, B., Pallier, C., Brysbaert, M., and Ferrand, L. (2004). Lexique 2: a new french lexical database. *Behav. Res. Methods Instrum. Comput.* 36, 516–524. doi: 10.3758/bf03195598

Nichols, T. E. (2012). Multiple testing corrections, nonparametric methods, and random field theory. *NeuroImage* 62, 811–815. doi: 10.1016/j.neuroimage.2012.04.014

Nicolson, R. I., and Fawcett, A. J. (2005). "Developmental dyslexia, learning and the cerebellum," in *Neurodevelopmental Disorders*, eds W. W. Fleischhacker and D. J. Brooks (Vienna: Springer Vienna), 19–36.

Nicolson, R. I., Fawcett, A. J., and Dean, P. (2001). Developmental dyslexia: the cerebellar deficit hypothesis. *Trends Neurosci.* 24, 508–511. doi: 10.1016/s0166-2236(00)01896-8

Norton, E. S., Beach, S. D., and Gabrieli, J. D. (2015). Neurobiology of dyslexia. *Curr. Opin. Neurobiol.* 30, 73–78. doi: 10.1016/j.conb.2014.09.007

O'Brien, G., and Yeatman, J. D. (2021). Bridging sensory and language theories of dyslexia: Toward a multifactorial model. *Dev. Sci.* 24:e13039. doi: 10.1111/desc.13039

Oosterhof, N. N., Connolly, A. C., and Haxby, J. V. (2016). Cosmomvpa: multi-modal multivariate pattern analysis of neuroimaging data in matlab/gnu octave. *Front. Neuroinf.* 10:27. doi: 10.3389/fninf.2016.00027

Paulesu, E., Danelli, L., and Berlingeri, M. (2014). Reading the dyslexic brain: multiple dysfunctional routes revealed by a new meta-analysis of pet and fmri activation studies. *Front. Hum. Neurosci.* 8:830. doi: 10.3389/fnhum.2014.00830

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: machine learning in python. *J. Machine Learn. Res.* 12, 2825–2830.

Pennington, B. F. (2006). From single to multiple deficit models of developmental disorders. *Cognition* 101, 385–413. doi: 10.1016/j.cognition.2006.04.008

Pennington, B. F., van Orden, G. C., Smith, S. D., Green, P. A., and Haith, M. M. (1990). Phonological processing skills and deficits in adult dyslexics. *Child Dev.* 61, 1753–1778. doi: 10.1111/j.1467-8624.1990.tb03564.x

Perrachione, T. K., Del Tufo, S. N., Winter, R., Murtagh, J., Cyr, A., Chang, P., et al. (2016). Dysfunction of rapid neural adaptation in dyslexia. *Neuron* 92, 1383–1397. doi: 10.1016/j.neuron.2016.11.020

Perry, C., Zorzi, M., and Ziegler, J. C. (2019). Understanding dyslexia through personalized large-scale computational models. *Psychol. Sci.* 30, 386–395. doi: 10.1177/0956797618823540

Pinel, P., Thirion, B., Meriaux, S., Jobert, A., Serres, J., Le Bihan, D., et al. (2007). Fast reproducible identification and large-scale databasing of individual functional cognitive networks. *BMC Neurosci.* 8:91. doi: 10.1186/1471-2202-8-91

Pisner, D. A., and Schnyer, D. M. (2020). " vector machine," in *Machine Learning*, Vol. Chapter 6 - Support, eds A. Mechelli and S. Vieira (Cambridge, MA: Academic Press), 101–121.

Pitt, M. A., and Samuel, A. G. (1995). Lexical and sublexical feedback in auditory word recognition. *Cogn. Psychol.* 29, 149–188. doi: 10.1006/cogp.1995.1014

Pugh, K. R., Frost, S. J., Sandak, R., Landi, N., Rueckl, J. G., Constable, R. T., et al. (2008). Effects of stimulus difficulty and repetition on printed word identification: an fmri comparison of nonimpaired and reading-disabled adolescent cohorts. *J. Cogn. Neurosci.* 20, 1146–1160. doi: 10.1162/jocn.2008.20079

Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., et al. (2003). Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. *Brain* 126, 841–865. doi: 10.1093/brain/awg076

Rueckl, J. G., Paz-Alonso, P. M., Molfese, P. J., Kuo, W. J., Bick, A., Frost, S. J., et al. (2015). Universal brain signature of proficient reading: Evidence from four contrasting languages. *Proc. Natl. Acad. Sci. U.S.A.* 112, 15510–15515. doi: 10.1073/pnas.1509321112

Snowling, M. J. (2000). *Dyslexia*. Oxford: Blackwell.

Snowling, M. J. (2001). From language to reading and dyslexia. *Dyslexia* 7, 37–46. doi: 10.1002/dys.185

Stefanac, N. R., Zhou, S.-H., Spencer-Smith, M. M., O'Connell, R., and Bellgrove, M. A. (2021). A neural index of inefficient evidence accumulation in dyslexia underlying slow perceptual decision making. *Cortex* 142, 122–137. doi: 10.1016/j.cortex.2021.05.021

Stein, J. (2014). Dyslexia: the role of vision and visual attention. *Curr. Dev. Disord. Rep.* 1, 267–280. doi: 10.1007/s40474-014-0030-6

Stein, J., and Walsh, V. (1997). To see but not to read; the magnocellular theory of dyslexia. *Trends Neurosci.* 20, 147–152. doi: 10.1016/s0166-2236(96)01005-3

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. *NeuroImage* 15, 273–289. doi: 10.1006/nimg.2001.0978

Vandermosten, M., Boets, B., Luts, H., Poelmans, H., Golestani, N., Wouters, J., et al. (2010). Adults with dyslexia are impaired in categorizing speech and nonspeech sounds on the basis of temporal cues. *Proc. Natl. Acad. Sci. U.S.A.* 107, 10389–10394. doi: 10.1073/pnas.0912858107

World Health Organisation [WHO] (2010). *ICIDH-2, The international classification of impairments, activities and participation*. Geneva: World Health Organisation.

Zhang, M., Riecke, L., and Bonte, M. (2021). Neurophysiological tracking of speech-structure learning in typical and dyslexic readers. *Neuropsychologia* 158:107889. doi: 10.1016/j.neuropsychologia.2021.107889

Ziegler, J. C. (2008). Better to lose the anchor than the whole ship. *Trends Cogn. Sci.* 12, 244–245. doi: 10.1016/j.tics.2008.04.001

Ziegler, J. C., and Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychol. Bull.* 131, 3–29. doi: 10.1037/0033-2909.131.1.3

Ziegler, J. C., Pech-Georgel, C., George, F., and Lorenzi, C. (2009). Speech-perception-in-noise deficits in dyslexia. *Dev. Sci.* 12, 732–745. doi: 10.1111/j.1467-7687.2009.00817.x

Ziegler, J. C., Perry, C., and Zorzi, M. (2014). Modelling reading development through phonological decoding and self-teaching: implications for dyslexia. *Philos Trans. R. Soc. Lond. B. Biol. Sci.* 369:20120397. doi: 10.1098/rstb.2012.0397

Ziegler, J. C., Perry, C., and Zorzi, M. (2019). "Modeling the Variability of Developmental Dyslexia," in *Developmental Dyslexia Across Languages and Writing Systems*, eds C. Perfetti, K. Pugh, and L. Verhoeven (Cambridge: Cambridge University Press), 350–371.

Ziegler, J. C., Perry, C., and Zorzi, M. (2020). Learning to read and dyslexia: from theory to intervention through personalized computational models. *Curr. Direct. Psychol. Sci.* 29, 293–300. doi: 10.1177/0963721420915873

frontiers | Frontiers in Neuroscience

TYPE Original Research
PUBLISHED 12 October 2022
DOI 10.3389/fnins.2022.984328

EDITED BY
Li-Hai Tan,
Shenzhen Institute of Neuroscience,
China

REVIEWED BY
Yang Yang,
Chinese Academy of Sciences (CAS),
China
Anthony James Krafnick,
Dominican University, United States
Alice H. D. Chan,
Nanyang Technological University,
Singapore

*CORRESPONDENCE
Neelima Wagley
neelima.wagley@vanderbilt.edu

SPECIALTY SECTION
This article was submitted to
Neurodevelopment,
a section of the journal
Frontiers in Neuroscience

RECEIVED 01 July 2022
ACCEPTED 08 September 2022
PUBLISHED 12 October 2022

CITATION
Wagley N and Booth JR (2022) Neural pathways of phonological and semantic processing and its relations to children's reading skills.
Front. Neurosci. 16:984328.
doi: 10.3389/fnins.2022.984328

Neural pathways of phonological and semantic processing and its relations to children's reading skills

Neelima Wagley* and James R. Booth

Department of Psychology and Human Development, Vanderbilt University, Nashville, TN, United States

Behavioral research shows that children's phonological ability is strongly associated with better word reading skills, whereas semantic knowledge is strongly related to better reading comprehension. However, most neuroscience research has investigated how brain activation during phonological and semantic processing is related to word reading skill. This study examines if connectivity during phonological processing in the dorsal inferior frontal gyrus (dIFG) to posterior superior temporal gyrus (pSTG) pathway is related to word reading skill, whereas connectivity during semantic processing in the ventral inferior frontal gyrus (vIFG) to posterior middle temporal gyrus (pMTG) pathway is related to reading comprehension skill. We used behavioral and functional magnetic resonance imaging (fMRI) data from a publicly accessible dataset on OpenNeuro.org. The research hypotheses and analytical plan were pre-registered on the Open Science Framework. Forty-six children ages 8–15 years old were included in the final analyses. Participants completed an in-scanner reading task tapping into phonology (i.e., word rhyming) and semantics (i.e., word meaning) as well as standardized measures of word reading and reading comprehension skill. In a series of registered and exploratory analyses, we correlated connectivity coefficients from generalized psychophysiological interactions (gPPI) with behavioral measures and used *z*-scores to test the equality of two correlation coefficients. Results from the preregistered and exploratory analyses indicated weak evidence that functional connectivity of dIFG to pSTG during phonological processing is positively correlated with better word reading skill, but no evidence that connectivity in the vIFG-pMTG pathway during semantic processing is related to better reading comprehension skill. Moreover, there was no evidence to support the differentiation between the dorsal pathway's relation to word reading and the ventral pathway's relation to reading comprehension skills. Our finding suggesting the importance of phonological processing to word reading is in line with prior behavioral and neurodevelopmental models.

KEYWORDS

phonology, semantics, word reading, reading comprehension, fMRI

Introduction

Reading is facilitated by three main interconnected systems: orthography, phonology, and semantics involving the occipitotemporal, temporoparietal, and inferior frontal cortex (Harm and Seidenberg, 2004; Sandak et al., 2012). Prior work, predominantly based on languages with alphabetic scripts, has established that the functional architecture of this network is associated with different reading skills throughout development (see Pugh et al., 2010; Landi et al., 2013 for review). Relations between network engagement and reading ability are commonly used to characterize differences in individuals with reading difficulties (e.g., Hoeft et al., 2006; van der Mark et al., 2011; Norton et al., 2015) and in skilled readers (e.g., Turkeltaub et al., 2003; Jobard et al., 2011; Welcome and Joanisse, 2012; Aboud et al., 2016; Ryherd et al., 2018). However, most neurobiological theories and extant computational models examine reading outcomes at the single-word level (Seidenberg, 2012; Pugh et al., 2013). In the current study, we examine how engagement of the neural pathways for phonological and semantic processing are related to individual differences in word reading versus reading comprehension skills in children ages 8–15 years old.

In alphabetic languages, successful word reading skills are strongly associated with phonological awareness abilities whereas reading comprehension skills are strongly associated with semantic knowledge (e.g., Wagner and Torgesen, 1987; Melby-Lervåg et al., 2012; Lervåg et al., 2018; Hjetland et al., 2019). Phonological skills are particularly relevant during the early stages of reading acquisition when children heavily rely on phoneme awareness and letter knowledge to decode words (Hjetland et al., 2019). They may also be involved in learning to read in non-alphabetic orthographies such as Chinese (e.g., McBride-Chang and Suk-Han Ho, 2005). However, this study is restricted to considering the role of phonological skills in learning to read in English. Patterns of neurodevelopment also suggest that successful reading is initially supported by brain connectivity for phonological decoding with a decreased reliance on this strategy as reading becomes more automated (Harm and Seidenberg, 2004; Shaywitz et al., 2004; Martin et al., 2015; Younger et al., 2017). Together, word decoding and language comprehension skills accounts for a large percent of variance in concurrent reading comprehension skills and its growth over time (Gough and Tunmer, 1986; Lervåg et al., 2018; Hjetland et al., 2019). Individuals with reading comprehension deficits, despite adequate phonemic decoding skill, can have difficulty with word-level semantic processing (e.g., Nation and Snowling, 1999; Landi and Perfetti, 2007; Cutting et al., 2013; Henderson et al., 2013; Spencer et al., 2014) and with higher-level word to text integration (Oakhill and Cain, 2012; Silva and Cain, 2015). Thus, assessing how the neural pathways for phonological and semantic processing relate to reading beyond single words may inform targeted remediation strategies and contribute to understanding long-term literacy outcomes (Landi and Ryherd, 2017).

Multiple overlapping and distinct brain regions support phonological and semantic processing (e.g., Fiebach et al., 2002; Vigneau et al., 2006; Mathur et al., 2020; Hodgson et al., 2021). Research suggests that the fronto-temporal network for reading is evident in children as young as 4 years old (Mathur et al., 2020; Jasińska et al., 2021). In a recent meta-analysis by Hodgson et al. (2021) comparing the two, phonological processing primarily involved a large cluster in the frontal lobe including the precentral gyrus and inferior frontal gyrus (IFG) pars opercularis and left posterior superior temporal gyrus (STG), as well as the superior parietal lobe. These hubs make-up the *dorsal* pathway of the reading circuitry. By contrast, the *ventral* pathway is associated with semantic processing and involves clusters in the left IFG pars triangularis and orbitalis and left posterior middle temporal gyrus (MTG), as well as the left anterior temporal lobe and angular gyrus (Hodgson et al., 2021). Prior cross-modal work suggests that these regions are generally engaged during phonological and semantic processing irrespective of the visual or auditory modality (e.g., Booth et al., 2002; Landi et al., 2010; Williams et al., 2015; Oron et al., 2016).

Prior work suggests a functional separation of the dorsal versus ventral left IFG for phonological and semantic processing (e.g., Jobard et al., 2003; Coltheart, 2005; Vigneau et al., 2006; Mathur et al., 2020; Hodgson et al., 2021; Wang et al., 2021b). Phonological processing of speech sounds involves perceptual processing in the STG and articulatory processing in the dorsal IFG (Hickok and Poeppel, 2004, 2007). Specifically, the dorsal IFG accesses phonological representations through connections via the arcuate fasciculus with STG (Saur et al., 2010; Boets et al., 2013) and is specialized for phonological processing during language production (Vigneau et al., 2006; Klaus and Hartwigsen, 2019). Specialization of the dorsal IFG for phonological processing is also evident when using a visual word rhyming task (e.g., Mathur et al., 2020). The MTG is engaged in lexical-semantic processes while the ventral IFG supports controlled processes such as meaning judgments or plausibility categorization (Thompson-Schill et al., 1997, 1999; Badre et al., 2005; Binder et al., 2009; Friederici and Gierhan, 2013). The ventral IFG accesses stored semantic knowledge in the temporal cortex through connections via the uncinate fasciculus (Thompson-Schill et al., 1997; Lau et al., 2008). It appears that these interconnected regions involved in phonological and semantic processing become more specialized with increased language experience (e.g., Weiss-Croft and Baldeweg, 2015; Skeide and Friederici, 2016; Perrone-Bertolotti et al., 2017).

Across studies using auditory and visual rhyming paradigms, phonological specialization in the left STG is evident in children by age five (e.g., Weiss et al., 2018; Mathur et al., 2020; Yamasaki et al., 2021), whereas specialization in the dorsal IFG is thought to develop later around age seven (e.g., Wang et al., 2021b; Yamasaki et al., 2021). Notably, the engagement of this *dorsal* fronto-temporal pathway during phonological processing is predictive of children's word reading skills throughout reading development (Wang et al., 2013). During an auditory word rhyming task, there was significant activation in the posterior left STG in children 6-years-old (Wang et al., 2020) and in the left IFG pars opercularis and posterior STG in children 7.5-years-old (Wang et al., 2021a). In the younger children, phonological processing in the left STG was a significant predictor of word reading skills 1.5 years later, even after controlling for initial levels of reading (Wang et al., 2020). In the older children, stronger functional connectivity from the dorsal IFG to STG during phonological processing predicted better word reading skills 1.5 years later (Wang et al., 2021a). These findings suggest that, by early elementary school, dorsal IFG and STG are specialized for phonological processing and that effective access of phonological representations via this pathway scaffolds children's word reading development (Wang et al., 2021a).

Semantic specialization in the left MTG is also evident in children by age five (e.g., Mathur et al., 2020; Wang et al., 2021b), whereas specialization in the ventral IFG seems to develop later around age seven (e.g., Wang et al., 2021b). Across studies with skilled adult and younger readers, engagement of the *ventral* fronto-temporal pathway during semantic processing is related to discourse-level reading skills (e.g., Lee et al., 2016; Yu et al., 2018; Jasińska et al., 2021). During a word reading task in adults, reading comprehension skill was significantly related to brain activation in the left IFG pars triangularis (Malins et al., 2016) and in the left MTG (Welcome and Joanisse, 2012). In adolescents ages 12–18 years, skilled comprehenders showed greater activation in left IFG pars triangularis and bilateral MTG during a discourse comprehension task (e.g., Landi et al., 2010; Ryherd et al., 2018). Together, this literature suggests that the ventral IFG and MTG are reliably engaged during both word- and discourse-level semantic tasks, and those with poor comprehension skills often struggle with accessing the lexical-semantic representations via this pathway during reading (e.g., Cutting et al., 2013).

In a recent study using a word rhyming and a word meaning task, better readers showed greater engagement of the dorsal IFG (pars opercularis) during phonological processing ($r = 0.40$) and a trend for greater engagement of the ventral IFG (pars triangularis) during semantic processing ($r = 0.30$; Brozdowski and Booth, 2021, *preprint*). However, reading skill was only assessed at the single-word level. The current study builds on the prior literature suggesting that phonological ability is strongly associated with better word reading skills, whereas semantic knowledge is strongly related to better reading comprehension skills. This study is the first to directly compare brain-behavior correlations to test how the engagement of the dorsal and ventral pathways may differentially relate to word- and passage- level reading skills. Specifically, we examine if the engagement of the dorsal pathway (i.e., dIFG to pSTG) during phonological processing is related to word reading skill, whereas the engagement of the ventral pathway (i.e., vIFG to pMTG) during semantic processing is related to reading comprehension skill in children ages 8–15 years old.

Based on the prior literature, we hypothesized the following: (1) the correlation between connectivity in the dorsal pathway (dIFG-pSTG) and word reading skills will be stronger than the correlation between connectivity in the ventral pathway (vIFG- pMTG) and word reading skills, (2) the correlation between connectivity in the ventral pathway (vIFG-pMTG) and reading comprehension skills will be stronger than the correlation between connectivity in the dorsal pathway (dIFG-pSTG) and reading comprehension skills, (3) the correlation between word reading skills and connectivity in the dorsal pathway (dIFG-pSTG) will be stronger than the correlation between reading comprehension skills and connectivity in the dorsal pathway (dIFG-pSTG), and (4) the correlation between reading comprehension skills and connectivity in the ventral pathway (vIFG-pMTG) will be stronger than the correlation between word reading skills and connectivity in the ventral pathway (vIFG-pMTG).

Materials and methods

This study was conducted used the Cross-Sectional Multidomain Lexical Processing dataset available on (Lytle et al., 2020). The research questions, hypotheses, and analytical plan were preregistered through the Open Science Framework after data cleaning but prior to beginning the data analyses.[1]

Participants

The dataset contains a sample of 91 native English-speaking children with normal hearing and normal or corrected-to-normal vision, with no neurological or psychiatric disorders, and not taking medications impacting the central nervous system. Data from participants who met the following inclusionary criteria were analyzed for the current study: (1) primarily right-handed assessed using five actions of writing, drawing, picking-up, opening, and throwing; score ≥ 3 indicates right-handedness (N = 4 excluded); (2) a standard score of at least 70 on the performance IQ subscale of the Wechsler Abbreviated Scale of Intelligence (WASI-II; Wechsler, 1999;

---

[1] https://osf.io/re7au

N = 2 excluded); and (3) having complete behavioral and imaging data with limited movement and acceptable task performance in the scanner (N = 38 excluded, see details below). One additional participant was excluded for errors found during data pre-processing. Forty-six participants are included in the final analyses (M_{age} = 11.7, SD = 2.2, 25 females, see **Table 1**). A list of included participant IDs is reported in the **Supplementary Table 1**. The excluded participants had comparable non-verbal IQ (t = 0.92, p = 0.32), word reading (t = 1.23, p = 0.32), and reading comprehension (t = 1.52, p = 0.13) scores to participants that were included in the final analysis.

Participants were recruited from the greater Chicago area. In total, 13% of the participants identified as Hispanic or Latinx. In total, 67% of participants identified as White, 11% as Black or African American, 9% as "other", 9% as multiracial, and 4% as Asian. Caregivers and children completed informed consent and assent forms before participation. All study procedures were approved by the Institutional Review Board at Northwestern University and Evanston Northwestern Healthcare Research Institute.

Procedure

Participants completed behavioral and fMRI tasks over two or more visits. First, participants completed standardized behavioral assessments followed by a practice imaging session in a mock scanner within a week of their fMRI session. This allowed participants to become familiar with the in-scanner tasks as well as the scanning environment. Practice tasks did not include any stimuli used in the experimental tasks. Lastly, participants completed the fMRI sessions.

Behavioral assessments of reading

We used raw scores from the *Word Identification* and *Passage Comprehension* subtests of the Woodcock-Johnson III Tests of Achievement (WJ-III; Woodcock et al., 2001) to assess word reading and reading comprehension skills, respectively. Six participants from the original dataset were excluded for not having complete reading data. The *Word Identification* subtest involves reading a series of words aloud, arranged from low- to high- difficulty, with a total of 76 items. Standard scores on the word reading task for the selected participants ranged from 82 to 130 [M(SD) = 105 (11)]. The Passage Comprehension subtest involves reading short sentences and identifying a missing key word that made sense in the context of the passage, with a total of 47 items. Standard scores on the reading comprehension task for the selected participants ranged from 76 to 133 [M(SD) = 103 (11)].

Functional magnetic resonance imaging lexical judgment tasks

Participants completed a rhyming task and a meaning task in the scanner. For both tasks, two words were visually presented in sequential order and contained three condition types: lexical, perceptual control, and fixation.

In all trials, the first stimulus was presented for 800 ms followed by an intertrial interval of 200 ms and the second stimulus for 800 ms. The second stimulus was followed by a red fixation cross lasting 2,600 ms indicating that participants should respond. Participants could respond as soon as the second stimulus was presented up until the start of the next trial. The second stimulus was offset right or left 1/2 a letter/symbol from the first to prevent judgments based on visual persistence. Stimuli were presented in the same order for all participants, optimized for event-related design using OptSeq (Greve, 2002). Word characteristics are provided in the stimuli directory of the OpenNeuro dataset (Lytle et al., 2020).

In the *Rhyming Task*, participants read two words and judged if the pair of words rhymed. Word pairs were grouped into four lexical conditions with 24 pairs in each condition: orthographically similar and phonologically similar (O+P+, *gate-hate*), orthographically different and phonologically similar (O–P+, *has-jazz*), orthographically similar and phonologically different (O+P–, *pint-mint*), and orthographically different and phonologically different (O–P–, *press-list*). Trial order was optimized and divided into two 108 trial runs collected in 240 volumes. All but five participants completed both runs in the same scanning session.

In the *Meaning Task*, participants read two words and judged if the pair of words were related in meaning. Word pairs were grouped into three conditions based on free association values (Nelson et al., 1998) with 24 pairs in each condition: strongly related (*found-lost*), weakly related (*dish-plate*), and unrelated (*tank-snap*). The average strength of association between word pairs in the strongly related condition was 0.60 (range = 0.36–0.77), 0.30 in the weakly related condition (range = 0.14–0.60), and 0 in the unrelated condition. Six word pairs from the weakly related condition overlapped in association values (>0.36) with the strongly related condition. Trial order was optimized and divided into one run with 91 trials and a second run with 89 trials. Due to the difference in length, run 1 for was collected in 203 volumes and run 2 was collected in 198 volumes. All but three participants completed both runs on the same scanning session.

In addition to the lexical trials, both tasks contained perceptual control and fixation trials. Participants were presented with a pair of symbols and were asked if the pair matched or not. Perceptual control trials were not modeled as conditions of interest in the present study. The fixation condition controlled for motor responses. In these trials,

TABLE 1 Participant demographics, standard scores on reading assessments, and behavioral performance during the two functional magnetic resonance imaging (fMRI) tasks.

	N = 46 (25 females)	
	M (SD)	Range
Age in years	11.7 (2.2)	8.7–15.5
WASI-II non-verbal IQ	106 (15)	77–144
WJ-III letter word ID	105 (12)	82–130
WJ-III passage comp.	103 (11)	76–133

	Accuracy (%)		Response time (ms)	
	M (SD)	Range	M (SD)	Range
Rhyming task				
O+P+	90 (11)	58–100	1,391 (347)	716–2,177
O–P+	78 (19)	21–100	1,462 (339)	815–2,392
O+P–	64 (25)	4–100	1,563 (359)	935–2,398
O–P–	93 (11)	44–100	1,390 (375)	705–2,293
Fixation	96 (7)	68–100	1,354 (357)	677–2,044
Meaning task				
Strongly related	92 (12)	54–100	1,344 (338)	588–2,158
Weakly related	89 (13)	46–100	1,331 (317)	656–2,122
Unrelated	90 (14)	42–100	1,466 (375)	764–2,337
Fixation	97 (7)	62–100	1,367 (326)	680–1,980

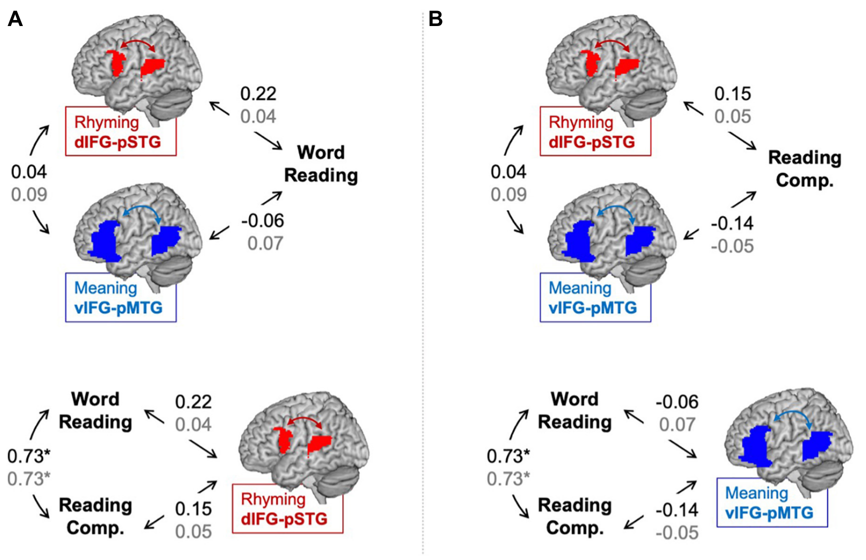


FIGURE 1

Partial correlations (controlling for age) between connectivity and *reading assessments* in the lexical > fixation (black) and the target-conditions > fixation (gray) contrasts. Panel **(A)** tests the hypotheses that the dorsal pathway for phonology is associated with word reading skill more than the ventral pathway for semantics and that word reading skill is associated with the dorsal pathway for phonology more than reading comprehension skill. Panel **(B)** tests the hypotheses that the ventral pathway for semantics is associated with reading comprehension skill more than the dorsal pathway for phonology and that reading comprehension skill is associated with the ventral pathway for semantics more than word reading skill. Statistically significant correlation denoted by an asterisk. $^*p < 0.001$.

participants were presented with a black fixation during the first and second stimulus phases and a red fixation during the response phase. Participants were asked to press the button under their index finger when the black cross turned red. The number of trials for the perceptual and fixation trials were as follows: 24 matching perceptual trials, 24 non-matching perceptual trials, and 72 fixation control trials.

Only those with complete data for both runs of the two tasks were included in the analysis (N = 20 excluded for missingness). Additionally, those who scored within an acceptable accuracy range and had no response bias were included in the analysis. Acceptable accuracy was defined as at least 50% accuracy in the O+P+, strongly related, and fixation conditions (N = 5 excluded). The lack of response bias was defined by no greater than a 50% difference in accuracy between the O+P+ and O–P– conditions for the rhyming task and the strongly related and unrelated conditions for the semantic task (N = 1 excluded).

Functional magnetic resonance imaging data acquisition

Magnetic resonance data were acquired using a 1.5 T General Electric (GE) Signa Excite scanner at Evanston Hospital, using a quadrature birdcage head coil. Participants were placed supine in the scanner and their head position was secured using a vacuum pillow (Bionix, Toledo, OH, USA). A response box was placed in the participant's right hand to allow them to respond to the tasks. Task stimuli were projected onto a screen, which the participants viewed through a mirror attached to the inside of the head coil. Structural T1-weighted SPGR images were collected using the following parameters: TR 1/4 33.333 ms, TE 1/4 8 ms, matrix size 1/4 256 × 256, bandwidth 1/4 114.922 Hz/Px, slice thickness 1/4 1.2 mm, number of slices 1/4 124, voxel size 1/4 1 mm isotropic, flip angle 1/4 30°. Blood oxygen level dependent signal (BOLD) was acquired using a T2-weighted susceptibility weighted single-shot echo planar imaging (EPI) and the following parameters: TR 1/4 2,000 ms, TE 1/4 25 ms, matrix size 1/4 64 × 64, bandwidth 1/4 7812.5 Hz/Px, slice thickness 1/4 5 mm, number of slices 1/4 24, voxel size 1/4 3.75 mm × 3.75 mm × 5 mm, flip angle 1/4 90°. Slices were acquired interleaved from bottom to top, odd first.

Functional magnetic resonance imaging data analysis

Preprocessing

Functional magnetic resonance imaging data was analyzed using SPM12.[2] Images were spatially realigned to the mean functional volume to correct for head movements, co-registered to the corresponding skull stripped T1 anatomical image and normalized to the Montreal Neurological Institute (MNI) space standard, with voxel size 2 $\text{mm}^3 \times 2\ \text{mm}^3 \times 2\ \text{mm}^3$. Functional images were then spatially smoothed with a 6-mm full-width at half-maximum isotropic Gaussian kernel. We used ArtRepair (Mazaika et al., 2007) to detect outlier volumes with more than 1.5 mm volume-to-volume movement, or with more than 4% deviation from the mean global signal. Outlier volumes were repaired by interpolating between the nearest non-outlier volumes. Interpolated volumes were then de-weighted when calculating first-level models on repaired images (Mazaika et al., 2007). No more than 10% of the volumes from each run and no more than six consecutive volumes for any individual were interpolated in this way. Six participants were excluded from analysis for excessive movement.

First-level analysis

First-level statistical analyses were performed on individual participants' data using the general linear model (GLM) as implemented in SPM12. The following regressors were entered into the GLM for the two runs: six motion regressors of head movement, two perceptual control conditions of no interest, one fixation condition, and four rhyme (O+P+, O+P–, O–P+, and O–P–) and three meaning (strongly related, weakly related, and unrelated) lexical conditions, for each run. The contrast of lexical > fixation was defined to produce individual level activation maps, which include the four rhyme or three meaning conditions.

Regions of interest masks

Based on the prior literature, four anatomical masks were used to isolate the ROIs for each task: left dorsal inferior frontal gyrus (dIFG; pars opercularis) and left posterior superior temporal gyrus (pSTG) for phonology and the left ventral inferior frontal gyrus (vIFG; pars triangularis and pars orbitalis) and left posterior middle temporal gyrus (pMTG) for semantics (Hodgson et al., 2021). The regions were identified using the anatomical automatic labeling (AAL) atlas template from WFU PickAtlas toolbox[3] and the MarsBar toolbox (Brett et al., 2002). The pSTG was defined as the posterior half of the left STG with $y < -24$ and the pMTG was defined as the posterior half of the left MTG with $y = -33$.

General psychophysiological interaction analysis

For each task, the top 100 voxels showing maximal activation (regardless of significance) for each participant for the lexical > fixation contrast in the dIFG (for the rhyming task) and vIFG (for the meaning task) was used as the seed region. We chose the top 100 voxels at the subject-level to define

---

[2] http://www.fil.ion.ucl.ac.uk/spm

[3] http://www.nitrc.org/projects/wfu_pickatlas

the seed region to focus on individual differences rather than a group-based cluster. This approach of using individualized ROIs is thought to be more sensitive at capturing the experimental manipulation and detecting differences between conditions and groups (Fedorenko et al., 2010; Tong et al., 2016). Specifically, the method of using the top 100 voxels regardless of significance, has been used by several previous studies to examine brain-behavioral correlations (Suárez-Pellicioni and Booth, 2018; Suárez-Pellicioni et al., 2019; Younger et al., 2019; Wang et al., 2020), including capturing individual differences in phonological and semantic processing using comparable paradigms of language and reading (e.g., Wang et al., 2021a; Yamasaki et al., 2021; Wang, 2022). Next, the following regressors were entered into a GLM in the individual level analysis for the two runs: the timeseries from the seed region, the experimental parameter regressors (seven for the rhyming task and six for the meaning task), the PPI regressors of the interaction (seven for the rhyming task and six for the meaning task), and the six motion regressors of head movement, for each run. The contrast of lexical > fixation was defined to produce individual level functional connectivity maps. Following, the average gPPI beta values were extracted from the top 100 voxels with the strongest connectivity in the pSTG (for rhyming) and pMTG (for meaning) anatomical mask. The dIFG-pSTG and vIFG-pMTG path coefficients for each participant were entered into the correlation analyses with reading scores. See **Supplementary material** for activation maps showing overlap across participants for the seed and target regions (**Supplementary Figure 1**) and the whole brain activation maps for the pre-registered (**Supplementary Figure 2**) and exploratory (**Supplementary Figures 3–5**) contrasts.

Brain and behavior analysis

Beta values from general psychophysiological interaction analysis (gPPI and raw scores from the reading assessments were entered into a partial correlation analysis using Pearson's r with age a covariate (ppcor package in R; Kim, 2015). Each brain-behavior partial correlation was independently calculated prior to computing the comparisons of correlation coefficients. To test each hypothesis, we used an interactive calculator to compute the z-score between two correlation coefficients (Lee and Preacher, 2013).[4] This calculator tests for the difference between two correlation coefficients obtained from the same dataset with the two correlations sharing one variable in common. Each test of equality between correlation coefficients was evaluated using a 1-tailed $p < 0.05$ threshold, given that we expected the correlations to be in a specific direction. For example, we expected that the correlation between connectivity in the dorsal pathway and word reading skills would be stronger, in the positive direction, than the correlation between connectivity in the ventral pathway and word reading skills.

Results

Preregistered analyses

Partial correlations between connectivity during the lexical > fixation contrasts and reading skills are shown in **Figure 1**. No correlation is defined as $r < 0.2$ and weak correlation is defined as $0.2 < r < 0.4$ (Dancey and Reidy, 2017). There was a weak correlation between word reading skill and dIFG-pSTG connectivity during the rhyming task ($r = 0.22$, $p = 0.14$), but no correlation between word reading skill and vIFG-pMTG connectivity during the meaning task ($r = -0.06$, $p = 0.68$). There was also no correlation between reading comprehension skill and dIFG-pSTG connectivity ($r = 0.15$, $p = 0.34$) or between reading comprehension skill and vIFG-pMTG connectivity ($r = -0.14$, $p = 0.36$); however, none of these correlations were significant. There was a strong correlation between the two reading skill measures ($r = 0.73$, $p < 0.001$), but no correlation between connectivity in the two pathways ($r = 0.04$, $p = 0.81$).

Results comparing correlation coefficients for each hypothesis are reported in **Table 2** with weak evidence is defined as z-score > 1.0. There was weak evidence to support the hypotheses that word reading skill was associated with the dorsal pathway for phonological processing more than the ventral pathway for semantic processing (z-score = 1.36, $p = 0.09$). Contrary to our hypothesis, there was weak evidence to show that reading comprehension was associated with the dorsal pathway for phonological processing more than the ventral pathway for semantic processing (z-score = −1.37, $p = 0.08$). There was no evidence to support the hypotheses that the dorsal pathway was associated with word reading more than reading comprehension skill (z-score = 0.69, $p = 0.24$) and that the ventral pathway was associated with reading comprehension more than word reading skill (z-score = 0.72, $p = 0.24$).

Exploratory analyses 1 – Change in measurement of brain activation

In the first exploratory analysis, we changed the contrast of interest to increase specificity in the measurement of phonological and semantic processing in the brain. The rhyming task contains four lexical conditions of which two require a "yes" response (O+P+, O−P+) and two require a "no" response (O+P−, O−P−). The semantic task has three lexical conditions with two "yes" response conditions (strongly and weakly related) and one "no" response condition (unrelated). This mismatch in response types across the tasks may place different demands on language

---

[4] http://quantpsy.org/corrtest/corrtest2.htm

TABLE 2 Results comparing correlation coefficients from the preregistered analysis (pre reg) using the lexical > fixation contrast (lex > fix) and exploratory analyses (explor 1 and explor 2) using the target-conditions > fixation (target > fix) contrasts.

Analysis	Contrast	Variable j	Variables k, h	r_{jk}	r_{jh}	r_{kh}	*z*-score	1-tail *p*
pre reg	lex > fix	Word reading	dIFG-pSTG vIFG-pMTG	0.22	−0.06	0.04	1.36	0.09
pre reg	lex > fix	Reading comp.	dIFG-pSTG vIFG-pMTG	0.15	−0.14	0.04	−1.37	0.08
pre reg	lex > fix	dIFG-pSTG	Word reading Reading comp.	0.22	0.15	0.73	0.69	0.24
pre reg	lex > fix	vIFG-pMTG	Word reading Reading comp.	−0.06	−0.14	0.73	0.72	0.24
explor 1	target > fix	Word reading	dIFG-pSTG vIFG-pMTG	0.04	0.07	0.09	−0.15	0.44
explor 1	target > fix	Reading comp.	dIFG-pSTG vIFG-pMTG	0.05	−0.05	0.09	0.49	0.31
explor 1	target > fix	dIFG-pSTG	Word reading Reading comp.	0.04	0.05	0.73	−0.09	0.46
explor 1	target > fix	vIFG-pMTG	Word reading Reading comp.	0.07	−0.05	0.73	1.07	0.14
explor 2	lex > fix	Rhyme accuracy	dIFG-pSTG vIFG-pMTG	0.19	0.01	0.18	0.93	0.17
explor 2	lex > fix	Mean accuracy	dIFG-pSTG vIFG-pMTG	0.04	0.04	0.18	0.00	0.50
explor 2	lex > fix	dIFG-pSTG	Rhyme accuracy Mean accuracy	0.19	0.04	0.71	1.31	0.09
explor 2	lex > fix	vIFG-pMTG	Rhyme accuracy Mean accuracy	0.01	0.04	0.71	−0.26	0.40
explor 2	target > fix	Rhyme accuracy	dIFG-pSTG vIFG-pMTG	0.16	0.11	0.18	0.26	0.40
explor 2	targe > fix	Mean accuracy	dIFG-pSTG vIFG-pMTG	0.03	0.03	0.18	0.00	0.50
explor 2	target > fix	dIFG-pSTG	Rhyme accuracy Mean accuracy	0.16	0.03	0.71	1.13	0.13
explor 2	target > fix	vIFG-pMTG	Rhyme accuracy Mean accuracy	0.11	0.03	0.71	0.69	0.24

and cognitive processes and could lead to variations in brain activations and the localization of these effects. Additionally, participants' task accuracy across the lexical conditions was higher for the semantic (average ~90%) than the rhyming task (average ~81%). To better equate the two experimental tasks across response type and difficulty, comparable task conditions were chosen for the exploratory analyses – O+P+ and O–P– for rhyming and strongly related and unrelated for meaning. The subsequent analyses using these conditions of interest will be referred to as the "target-conditions," which also align with the task conditions used as part of the study's inclusionary criteria for filtering accuracy and response bias.

In the first exploratory analysis, all first-level analysis parameters remained the same except the contrast of lexical > fixation was changed to [(O+P+ and O–P–) > fixation] for the rhyming task and [(strongly related and unrelated) > fixation] for the semantic task to produce individual level activation maps. The top 100 activated voxels (regardless of significance) in the dIFG and vIFG for these contrasts made up the seed regions for each task. Like the pre-registered analysis, we computed gPPI analysis using the timeseries from the newly defined contrasts to produce individual level functional connectivity maps. We extracted the average gPPI beta values from the top 100 voxels with the strongest connectivity in the pSTG and pMTG anatomical mask and computed brain-behavior analyses.

Partial correlations between connectivity during the target-conditions > fixation contrasts and reading assessments are shown in **Figure 1**. There was no correlation between word reading skill and dIFG-pSTG connectivity during the rhyming task ($r = 0.04$, $p = 0.79$) or between word reading skill and vIFG-pMTG connectivity during the meaning task ($r = 0.07$, $p = 0.64$). Similarly, there was no correlation of reading comprehension skill with dIFG-pSTG connectivity ($r = 0.05$, $p = 0.72$) or vIFG-pMTG connectivity ($r = -0.05$, $p = 0.77$). There was no correlation between connectivity in the two pathways using the new contrasts of interest ($r = 0.09$, $p = 0.56$).

Results comparing correlation coefficients for this exploratory analysis are reported in **Table 2**. There was no evidence to support the hypotheses that word reading skill was associated with the dorsal pathway for phonological processing more than the ventral pathway for semantic processing (z-score = -0.15, $p = 0.44$) or that reading comprehension skill was associated with the ventral pathway more than the dorsal pathway (z-score = 0.49, $p = 0.31$). There was no evidence to support the hypotheses that the dorsal pathway was associated with word reading more than reading comprehension skill (z-score = -0.09, $p = 0.46$). Contrary to our hypothesis, there was weak evidence to show that the ventral pathway for semantics was associated with word reading skill more than reading comprehension skill (z-score = 1.07, $p = 0.14$). However, this comparison was not statistically significant.

Exploratory analyses 2 – Change in behavioral measure of reading skills

In the second exploratory analysis, we changed the behavioral assessments to increase specificity in the measurement of phonological and semantic processing during reading. While reading comprehension does engage vocabulary and activation of semantic knowledge, this measure additionally taps into cognitive processes such as integrating syntax and semantics, making inferences, self-monitoring, etc. (Melby-Lervåg and Lervåg, 2014). To better capture phonological processing as a core component of word reading skill and semantic processing as a central index of comprehension, we replaced the word reading and passage comprehension scores with performance on the in-scanner rhyming and meaning tasks, respectively.

All first-level and gPPI analysis parameters remained the same. Brain-behavior analyses used overall task accuracy as the outcome measure of reading. We used the Spearman rho to evaluate brain-behavior correlations as task accuracy data is non-parametric. We conducted this exploratory analysis using data from both the lexical > fixation and the target-conditions > fixation contrasts to parallel the prior analyses.

Partial correlations between connectivity using the lexical > fixation contrasts and the in-scanner task accuracies are shown in **Figure 2**. There was no correlation between accuracy in the rhyming task and dIFG-pSTG connectivity during the rhyming task ($r = 0.19$, $p = 0.22$). There was no correlation between accuracy on the rhyming task and vIFG-pMTG connectivity during the meaning task ($r = 0.01$, $p = 0.93$). There was no correlation between accuracy on the meaning task and dIFG-pSTG connectivity ($r = 0.04$, $p = 0.79$) and vIFG-pMTG connectivity ($r = 0.04$ $p = 0.82$). There was a significant correlation between the two say in-scanner reading measures ($r = 0.71$, $p < 0.001$), and no correlation between connectivity in the two pathways ($r = 0.18$, $p = 0.22$). Like the preregistered analyses, none of these correlations were significant.

Partial correlations between connectivity using the target-conditions > fixation contrasts and the in-scanner task accuracies are shown in **Figure 2**. There was no correlation between accuracy in the rhyming task and dIFG-pSTG connectivity during the rhyming task ($r = 0.16$, $p = 0.29$) or vIFG-pMTG connectivity during the meaning task ($r = 0.11$, $p = 0.46$). There was no correlation between accuracy on the meaning task and dIFG-pSTG connectivity ($r = 0.03$, $p = 0.84$) or vIFG-pMTG connectivity ($r = 0.03$, $p = 0.86$).

Results comparing correlation coefficients for this exploratory analysis are reported in **Table 2**. Results were similar across the two analyses using the lexical > fixation and target-conditions > fixation contrasts. There was no evidence to support the hypotheses that accuracy on the word rhyming task was associated with the dorsal pathway for phonological processing more than the ventral pathway for semantic processing when using the lexical > fixation contrast (z-score = 0.93, $p = 0.17$) or the target-conditions > fixation contrast (z-score = 0.26, $p = 0.40$). There was no evidence that accuracy on the word meaning task was associated with the ventral pathway more than the dorsal pathway (z-score = 0, $p = 0.50$). There was weak evidence to support the hypotheses that the dorsal pathway for phonology was associated with accuracy on the word rhyming task more than accuracy on the word meaning task, for both the lexical > fixation (z-score = 1.31, $p = 0.09$) and target-conditions > fixation (z-score = 1.13, $p = 0.13$) contrasts. Lastly, there was no evidence that the ventral pathway for semantics was associated with accuracy on the word meaning task more than accuracy on the word rhyming task when using the target-conditions > fixation contrast (z-score = 0.69, $p = 0.24$) or the lexical > fixation contrast (z-score = -0.26, $p = 0.40$).

Discussion

The primary focus of most cognitive neuroscience research in reading has been at the single-word-processing level. Deficits in phonological processing is a key marker of traditionally identified dyslexia, whereas deficits in semantic processing is thought to characterize specific deficits in reading comprehension (e.g., Rueckl and Seidenberg, 2009; Landi et al., 2010; Cutting et al., 2013). The aim of the current study was to examine if engagement of the dorsal pathway during phonological processing is related to word reading skill, whereas engagement of the ventral pathway during semantic processing is related to reading comprehension skill in children ages 8–15 years old. Results from the preregistered and exploratory analyses indicated weak evidence

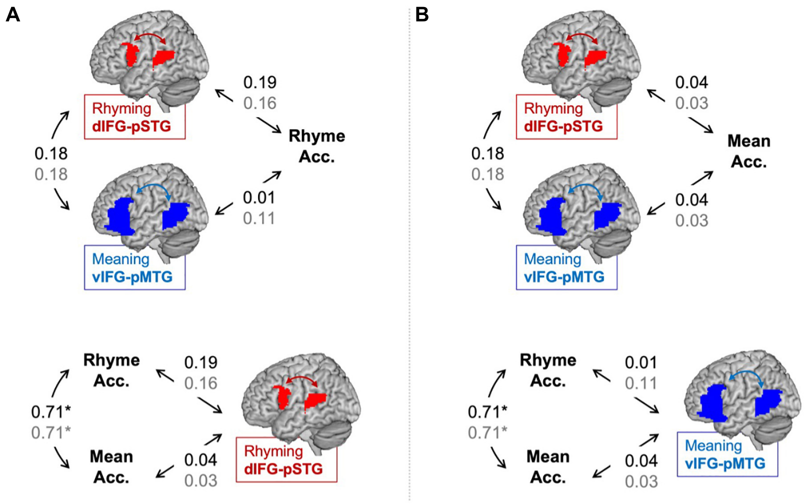


FIGURE 2
Partial correlations (controlling for age) between connectivity and *in-scanner task accuracy* in the lexical > fixation (black) and the target-conditions > fixation (gray) contrasts. Panel **(A)** tests the hypotheses that the dorsal pathway for phonology is associated with word rhyming accuracy more than the ventral pathway for semantics and that word rhyming accuracy is associated with the dorsal pathway for semantics more than word meaning accuracy. Panel **(B)** tests the hypotheses that the ventral pathway for semantics is associated with word meaning accuracy more than the dorsal pathway for phonology and that word meaning accuracy is associated with the ventral pathway for semantics more than word rhyming accuracy. Statistically significant correlation denoted by an asterisk. $^*p < 0.001$.

that dIFG to pSTG functional connectivity during phonological processing is positively correlated with word reading skill (see Table 3). The weak evidence is consistent with prior behavioral and neurodevelopmental models of reading suggesting that phonological awareness is associated with word reading ability (e.g., Melby-Lervåg et al., 2012; Pugh et al., 2013).

The dorsal network's engagement during phonological processing and its relation to word reading skills in English has been extensively examined across the range of reading development, from preliteracy to adolescence (e.g., Wang et al., 2013, 2020; Yu et al., 2018; Jasińska et al., 2021). A few studies have supported the argument that engagement of the ventral network for semantic processing is stronger in children with better reading comprehension skills (e.g., Landi et al., 2010; Cutting et al., 2013; Aboud et al., 2016; Ryherd et al., 2018). Overall, we observed that dorsal functional connectivity between dIFG and pSTG during phonological processing was weakly related to children's word reading skill or accuracy on the visual rhyming task in the scanner. On the other hand, there was no evidence to support the hypothesis that the ventral pathway for semantics from vIFG to pMTG was related to reading comprehension skill or accuracy on the visual meaning task in the scanner. Lastly, there was no evidence to support the hypothesis that differential engagement of the dorsal and ventral pathways is related to word reading versus reading comprehension skill.

In the pre-registered analyses, we related functional connectivity during all lexical conditions with standardized measures of reading. Contrary to the prior literature showing moderate to strong associations between phonological brain systems and word reading skills (e.g., Wang et al., 2013, 2020; Yu et al., 2018; Jasińska et al., 2021; Yamasaki et al., 2021), we only found weak (and statistically unreliable) correlations between connectivity in the dorsal pathway and children's word reading skills. To our surprise, this weak correlation was also evident in our first exploratory analyses which examined connectivity using a task contrast of targeted conditions, (O+P+ and O–P–) > fixation for the rhyming task and (strongly related and unrelated) > fixation for the semantic task. Some methodological parameters may help explain these limited findings.

We chose the IFG as a seed region because models of language and reading suggests a functional separation of the dorsal versus ventral left IFG for phonological and semantic processing, respectively (e.g., Mathur et al., 2020; Hodgson et al., 2021). Specifically, the dorsal IFG (pars opercularis) is thought to access the phonological representations in the STG through connections via the arcuate fasciculus, whereas the ventral IFG (pars triangularis and/or orbitalis) is believed to access stored semantic knowledge in the MTG through connections via the uncinate fasciculus (e.g., Badre et al., 2005; Saur et al., 2010; Friederici and Gierhan, 2013; Hodgson et al., 2021). In 7-year-old children, stronger functional connectivity from dorsal IFG to STG during phonological processing has been shown to predict better word reading skills later in development (Wang et al., 2021a). A recent study using the same word rhyming and meaning tasks as the current study found that better readers showed greater engagement of the dorsal

TABLE 3 Summary of results showing strength of evidence for each hypothesis across the pre-registered and exploratory analyses using the lexical > fixation (lex > fix) and target-conditions > fixation (target > fix) contrasts.

	Preregistered	**Exploratory 1**	**Exploratory 2A**	**Exploratory 2B**
	lex > fix Reading skill	***target > fix Reading skill***	***lex > fix Task accuracy***	***target > fix Task accuracy***
(dIFG-pSTG ⇔ word reading/rhyming acc) > (vIFG-pMTG ⇔ word reading/rhyming acc)	Weak evidence	No evidence	No evidence	No evidence
(vIFG-pMTG ⇔ reading comp/meaning acc) > (dIFG-pSTG ⇔ reading comp/meaning acc)	Weak evidence for the alternate	No evidence	No evidence	No evidence
(word reading/rhyming acc ⇔ dIFG-pSTG) > (reading comp/meaning acc ⇔ dIFG-pSTG)	No evidence	No evidence	Weak evidence	Weak evidence
(reading comp/meaning acc ⇔ vIFG-pMTG) > (word reading/rhyming acc ⇔ vIFG-pMTG)	No evidence	Weak evidence for the alternate	No evidence	No evidence

Weak evidence is defined as $p < 0.25$.

IFG (pars opercularis) during phonological processing and greater engagement of the ventral IFG (pars triangularis) during semantic processing (Brozdowski and Booth, 2021, *preprint*). However, this study focused on single word reading and did not examine connectivity. The current study is the first to directly compare brain-behavior correlations to test how connectivity of the dorsal and ventral pathways may differentially relate to word- and passage- level reading skills.

Meta-analyses of neuroimaging studies across adults and children show high convergence of reading-related activation in the left IFG (Vigneau et al., 2006; Martin et al., 2015). Connectivity between the anterior reading circuit in the IFG and other temporoparietal regions also relate to individual differences in reading skill. For example, the supramarginal and angular gyri in the inferior parietal lobe are thought to be involved in mapping orthographic input to phonological and semantic properties of written words (Welcome and Joanisse, 2012; Lee et al., 2016). Better readers, at the word and sentence level, show greater connectivity between the inferior parietal regions and the left IFG and MTG (e.g., Cutting et al., 2013; Pugh et al., 2013; Aboud et al., 2016; Lee et al., 2016). Yu et al. (2018) observed greater connectivity between left IFG and inferior parietal lobe in 5-year-old children whose phonological abilities increased most over the course of reading development. The strength of these connections predicted later word reading skills at ages 7–8 years old (Yu et al., 2018).

Given that the supramarginal and angular gyri may be involved in both phonological and semantic integration, expanding our posterior mask to include these parietal regions could provide insight into how engagement of the dorsal and ventral pathway differs across tasks (e.g., rhyming versus meaning judgments) in relation to reading skills.

A significant contribution of this study is that we use an individual differences approach to systematically test questions related to the brain bases of reading. First, we used individual functional activation maps to define the seed regions for phonological and semantic processing within each experimental task. We then used individual connectivity maps from the gPPI analyses to examine its relations with children's reading skills and task performance in the scanner. This approach allowed us to capture variability in the engagement of the reading circuit which is apparent even within populations of skilled readers (e.g., Seghier et al., 2004, 2008; Jobard et al., 2011; Welcome and Joanisse, 2012). In the current sample, variability in the spatial distribution of voxels across the frontotemporal regions of interest is shown in the overlap maps in the **Supplementary Figure 1**.

In the second exploratory analyses, we used overall task accuracy on the rhyming and meaning tasks instead of standardized reading assessments as the outcome measure of reading skill. The *Word Identification* subtest of the WJ-III primarily assesses children's oral word decoding but may also engage other processes such as semantics. Similarly, the *Passage Comprehension* subtest involves comprehension skills tapping into vocabulary knowledge, but also involves other processes such as syntax, inferencing, and working memory. We expected accuracy on our experimental word reading tasks might better capture phonological processing, as a core component of word reading skill, and semantic processing, as an essential index of comprehension. Brain-behavior results from these exploratory analyses suggest weak evidence to support the hypothesis that engagement of the dorsal pathway is related to accuracy on the word rhyming task, but no evidence to support the hypothesis that engagement of the ventral pathway is related to accuracy on the word meaning task. Like the previous set of analyses, there was no evidence to support the hypothesis that differential engagement of the dorsal and ventral pathways is related to word rhyming versus word meaning judgments.

Overall, the evidence for our hypotheses is weak and unreliable and therefore needs to be replicated. Thus, we aim to extend these findings to a separate cohort of children and address some of the methodological limitations in the current analyses. Both contrasts we used in the preregistered and exploratory analyses included the fixation condition as the baseline measure. In doing so, the contrasts may have captured general lexical processing effects and/or may not have been sensitive enough to elicit robust engagement of the fronto-temporal regions as related to reading skills. Prior work showing strong associations of brain activation in phonological and semantic hubs with behavioral measures of word- and discourse- level reading skills used non-lexical perceptual stimuli as the baseline subtraction (e.g., Turkeltaub et al., 2003) or contrasted two lexical conditions (e.g., Malins et al., 2016). For example, in participants ages 6–18 years old, phonological awareness ability positively correlated with activation in the left posterior STS (cluster $r = 0.62$) during an implicit reading task that contrasted words with false-font strings (Turkeltaub et al., 2003). Similarly, Malins et al. (2016) tested a “localizer” word reading task to target the orthographic, phonological, and semantic components of reading. Their stimuli type assessing the latter two components were nearly identical to the current study. When contrasting activations pertaining to the semantically related versus unrelated words, engagement of the left IFG (pars triangularis) was related to reading comprehension skills (Malins et al., 2016). In the same study, the authors also observed significant activation in the IFG pars opercularis when contrasting phonologically inconsistent sets of words (O+P–) compared to consistent sets (O+P+), although they did not observe any significant associations between this activation and reading skills. These alternate models may be more sensitive at capturing the phonological and semantic processes that relate to different reading skills.

In conclusion, the present study is the first to directly compare brain-behavior correlations to test how the connectivity of the dorsal and ventral pathways for reading may differentially relate to word- and passage- level reading skills. Our preregistered and exploratory analyses show weak evidence that functional connectivity in the dorsal dIFG-pSTG pathway for phonological processing is positively correlated with word reading skill, but no evidence that connectivity in the ventral vIFG-pMTG pathway during semantic processing is related to reading comprehension skill. Moreover, there was no evidence to support the differentiation between the dorsal pathway’s relation to word reading and the ventral pathway’s relation to reading comprehension skills in children ages 8–15 years old. Our findings need to be replicated with a different sample, and perhaps extended by examining parietal regions implicated in phonological and semantic processing, by using more targeted skill measures of word and passage comprehension and by employing neuroimaging baseline tasks that control more effectively for perceptual processing.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: https://openneuro.org/datasets/ds002236/versions/1.0.0. Additional materials (e.g., scripts used for data analysis) can be found at https://osf.io/4wdcv/.

Ethics statement

The studies involving human participants were reviewed and approved by Institutional Review Board at Northwestern University and Evanston Northwestern Healthcare Research Institute. Written informed consent to participate in this study was provided by the participants’ legal guardian/next of kin.

Author contributions

NW: conceptualization, methodology, formal analysis, and original draft. JB: conceptualization, methodology, review and editing, supervision, and funding acquisition. Both authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Institutes of Health grants R01 DC018171, HD042049, and HD093547.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2022.984328/full#supplementary-material

References

Aboud, K. S., Bailey, S. K., Petrill, S. A., and Cutting, L. E. (2016). Comprehending text versus reading words in young readers with varying reading ability: Distinct patterns of functional connectivity from common processing hubs. *Dev. Sci.* 19, 632–656. doi: 10.1111/desc.12422

Badre, D., Poldrack, R. A., Paré-Blagoev, E. J., Insler, R. Z., and Wagner, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron* 47, 907–918.

Binder, J. R., Desai, R. H., Graves, W. W., and Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb. Cortex* 19, 2767–2796.

Boets, B., Op de Beeck, H. P., Vandermosten, M., Scott, S. K., Gillebert, C. R., Mantini, D., et al. (2013). Intact but less accessible phonetic representations in adults with dyslexia. *Science* 342, 1251–1254. doi: 10.1126/science.1244333

Booth, J. R., Burman, D. D., Meyer, J. R., Gitelman, D. R., Parrish, T. B., and Mesulam, M. M. (2002). Functional anatomy of intra-and cross-modal lexical tasks. *Neuroimage* 16, 7–22. doi: 10.1006/nimg.2002.1081

Brett, M., Anton, J. L., Valabregue, R., and Poline, J. B. (2002). "Region of interest analysis using an SPM toolbox," in *Paper Presented at the 8th International Conference on Functional Mapping of the Human Brain*, Vol. 16, (Sendai), 497.

Brozdowski, C., and Booth, J. R. (2021). Reading skill correlates in frontal cortex during semantic and phonological processing. *PsyArXiv [Preprint]*. doi: 10.31234/osf.io/d3mj7

Coltheart, M. (2005). "Modeling reading: The dual-route approach," in *The Science of Reading: A Handbook*, eds M. J. Snowling & C. Hulme Vol. 6 (Hoboken, NJ: Blackwell Publishing), 23.

Cutting, L. E., Clements-Stephens, A., Pugh, K. R., Burns, S., Cao, A., Pekar, J. J., et al. (2013). Not all reading disabilities are dyslexia: Distinct neurobiology of specific comprehension deficits. *Brain Connect.* 3, 199–211. doi: 10.1089/brain.2012.0116

Dancey, C. P., and Reidy, J. (2017). *Statistics Without Maths for Psychology*. London: Pearson, 632–632.

Fedorenko, E., Hsieh, P. J., Nieto-Castañón, A., Whitfield-Gabrieli, S., and Kanwisher, N. (2010). New method for fMRI investigations of language: Defining ROIs functionally in individual subjects. *J. Neurophysiol.* 104, 1177–1194.

Fiebach, C. J., Friederici, A. D., Müller, K., and Von Cramon, D. Y. (2002). fMRI evidence for dual routes to the mental lexicon in visual word recognition. *J. Cogn. Neurosci.* 14, 11–23. doi: 10.1162/089892902317205285

Friederici, A. D., and Gierhan, S. M. (2013). The language network. *Curr. Opin. Neurobiol.* 23, 250–254.

Gough, P. B., and Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remed. Spec. Educ.* 7, 6–10.

Greve, D. N. (2002). *Optseq2 Computer Software*. Available online at: http://surfer.nmr.mgh.harvard.edu/optseq

Harm, M. W., and Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychol. Rev.* 111:662. doi: 10.1037/0033-295X.111.3.662

Henderson, L., Weighall, A., and Gaskell, G. (2013). Learning new vocabulary during childhood: Effects of semantic training on lexical consolidation and integration. *J. Exp. Child Psychol.* 116, 572–592. doi: 10.1016/j.jecp.2013.07.004

Hickok, G., and Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition* 92, 67–99. doi: 10.1016/j.cognition.2003.10.011

Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402.

Hjetland, H. N., Lervåg, A., Lyster, S. A. H., Hagtvet, B. E., Hulme, C., and Melby-Lervåg, M. (2019). Pathways to reading comprehension: A longitudinal study from 4 to 9 years of age. *J. Educ. Psychol.* 111:751. doi: 10.1542/peds.2004-2390

Hodgson, V. J., Ralph, M. A. L., and Jackson, R. L. (2021). Multiple dimensions underlying the functional organization of the language network. *Neuroimage* 241:118444.

Hoeft, F., Hernandez, A., McMillon, G., Taylor-Hill, H., Martindale, J. L., Meyler, A., et al. (2006). Neural basis of dyslexia: A comparison between dyslexic and nondyslexic children equated for reading ability. *J. Neurosci.* 26, 10700–10708. doi: 10.1523/JNEUROSCI.4931-05.2006

Jasińska, K. K., Shuai, L., Lau, A. N., Frost, S., Landi, N., and Pugh, K. R. (2021). Functional connectivity in the developing language network in 4-year-old children predicts future reading ability. *Dev. Sci.* 24:e13041. doi: 10.1111/desc.13041

Jobard, G., Crivello, F., and Tzourio-Mazoyer, N. (2003). Evaluation of the dual route theory of reading: A metanalysis of 35 neuroimaging studies. *Neuroimage* 20, 693–712. doi: 10.1016/S1053-8119(03)00343-4

Jobard, G., Vigneau, M., Simon, G., and Tzourio-Mazoyer, N. (2011). The weight of skill: Interindividual variability of reading related brain activation patterns in fluent readers. *J. Neurolinguist.* 24, 113–132.

Kim, S. (2015). ppcor: An R package for a fast calculation to semi-partial correlation coefficients. *Commun. Stat. appl. Methods* 22:665. doi: 10.5351/CSAM.2015.22.6.665

Klaus, J., and Hartwigsen, G. (2019). Dissociating semantic and phonological contributions of the left inferior frontal gyrus to language production. *Hum. Brain Mapp.* 40, 3279–3287.

Landi, N., and Perfetti, C. A. (2007). An electrophysiological investigation of semantic and phonological processing in skilled and less-skilled comprehenders. *Brain Lang.* 102, 30–45. doi: 10.1016/j.bandl.2006.11.001

Landi, N., and Ryherd, K. (2017). Understanding specific reading comprehension deficit: A review. *Lang. Linguist. Compass* 11: e12234.

Landi, N., Frost, S. J., Mencl, W. E., Sandak, R., and Pugh, K. R. (2013). Neurobiological bases of reading comprehension: Insights from neuroimaging studies of word-level and text-level processing in skilled and impaired readers. *Read. Writ. Q.* 29, 145–167. doi: 10.1080/10573569.2013.758566

Landi, N., Mencl, W. E., Frost, S. J., Sandak, R., and Pugh, K. R. (2010). An fMRI study of multimodal semantic and phonological processing in reading disabled adolescents. *Ann. Dyslexia* 60, 102–121. doi: 10.1007/s11881-009-0029-6

Lau, E. F., Phillips, C., and Poeppel, D. (2008). A cortical network for semantics:(de) constructing the N400. *Nat. Rev. Neurosci.* 9, 920–933. doi: 10.1038/nrn2532

Lee, I. A., and Preacher, K. J. (2013). *Calculation for the test of the difference between two dependent correlations with one variable in common [Computer software]*. Available online at: http://quantpsy.org

Lee, S. H., Booth, J. R., and Chou, T. L. (2016). Temporo-parietal connectivity uniquely predicts reading change from childhood to adolescence. *Neuroimage* 142, 126–134. doi: 10.1016/j.neuroimage.2016.06.055

Lervåg, A., Hulme, C., and Melby-Lervåg, M. (2018). Unpicking the developmental relationship between oral language skills and reading comprehension: It's simple, but complex. *Child Dev.* 89, 1821–1838. doi: 10.1111/cdev.12861

Lytle, M. N., Bitan, T., and Booth, J. R. (2020). A neuroimaging dataset on orthographic, phonological and semantic word processing in school-aged children. *Data Brief* 28:105091. doi: 10.1016/j.dib.2019.105091

Malins, J. G., Gumkowski, N., Buis, B., Molfese, P., Rueckl, J. G., Frost, S. J., et al. (2016). Dough, tough, cough, rough: A "fast" fMRI localizer of component processes in reading. *Neuropsychologia* 91, 394–406. doi: 10.1016/j.neuropsychologia.2016.08.027

Martin, A., Schurz, M., Kronbichler, M., and Richlan, F. (2015). Reading in the brain of children and adults: A meta-analysis of 40 functional magnetic resonance imaging studies. *Human Brain Mapp.* 36, 1963–1981. doi: 10.1002/hbm.22749

Mathur, A., Schultz, D., and Wang, Y. (2020). Neural bases of phonological and semantic processing in early childhood. *Brain Connect.* 10, 212–223. doi: 10.1089/brain.2019.0728

Mazaika, P., Whitfield-Gabrieli, S., Reiss, A., and Glover, G. (2007). Artifact repair for fMRI data from high motion clinical subjects. *Hum. Brain Mapp.* 47, 70238–1.

McBride-Chang, C., and Suk-Han Ho, C. (2005). Predictors of beginning reading in Chinese and English: A 2-year longitudinal study of Chinese kindergartners. *Sci. Stud. Read.* 9, 117–144.

Melby-Lervåg, M., and Lervåg, A. (2014). Reading comprehension and its underlying components in second-language learners: A meta-analysis of studies comparing first-and second-language learners. *Psychol. Bull.* 140:409. doi: 10.1037/a0033890

Melby-Lervåg, M., Lyster, S. A. H., and Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychol. Bull.* 138:322. doi: 10.1037/a0026744

Nation, K., and Snowling, M. J. (1999). Developmental differences in sensitivity to semantic relations among good and poor comprehenders: Evidence from semantic priming. *Cognition* 70, B1–B13.

Nelson, D. L., McEvoy, C. L., and Schreiber, T. A. (1998). *The University of South Florida word Association, Rhyme, and Word Fragment Norms*. Tampa, FL: University of South Florida.

Norton, E. S., Beach, S. D., and Gabrieli, J. D. (2015). Neurobiology of dyslexia. *Curr. Opin. Neurobiol.* 30, 73–78. doi: 10.1016/j.ajp.2020.102357

Oakhill, J. V., and Cain, K. (2012). The precursors of reading ability in young readers: Evidence from a four-year longitudinal study. *Sci. Stud. Read.* 16, 91–121.

Oron, A., Wolak, T., Zeffiro, T., and Szelag, E. (2016). Cross-modal comparisons of stimulus specificity and commonality in phonological processing. *Brain Lang.* 155, 12–23. doi: 10.1016/j.bandl.2016.02.001

Perrone-Bertolotti, M., Kauffmann, L., Pichat, C., Vidal, J. R., and Baciu, M. (2017). Effective connectivity between ventral occipito-temporal and ventral inferior frontal cortex during lexico-semantic processing. A dynamic causal modeling study. *Front. Hum. Neurosci.* 11:325. doi: 10.3389/fnhum.2017.00325

Pugh, K. R., Frost, S. J., Sandak, R., Landi, N., Moore, D., Della Porta, G., et al. (2010). "Mapping the word reading circuitry in skilled and disabled readers," in *The Neural Basis of Reading*, eds P. L. Cornelissen, P. C. Hansen, M. L. Kringelbach, and K. R. Pugh (New York, NY: Oxford University Press), 281–305. doi: 10.1016/j.bandl.2013.05.001

Pugh, K. R., Landi, N., Preston, J. L., Mencl, W. E., Austin, A. C., Sibley, D., et al. (2013). The relationship between phonological and auditory processing and brain organization in beginning readers. *Brain Lang.* 125, 173–183. doi: 10.1016/j.bandl.2012.04.004

Rueckl, J. G., and Seidenberg, M. S. (2009). "Computational modeling and the neural bases of reading and reading disorders," in *How Children Learn to Read: Current Issues and New Directions in the Integration of Cognition, Neurobiology and Genetics of Reading and Dyslexia Research and Practice*, eds K. Pugh and P. McCardle (London: Psychology press), 101–133.

Ryherd, K., Jasinska, K., Van Dyke, J. A., Hung, Y. H., Baron, E., Mencl, W. E., et al. (2018). Cortical regions supporting reading comprehension skill for single words and discourse. *Brain Lang.* 186, 32–43. doi: 10.1016/j.bandl.2018.08.001

Sandak, R., Frost, S. J., Rueckl, J. G., Landi, N., Mencl, W. E., Katz, L., et al. (2012). "How does the brain read words," in *The Cambridge Handbook of Psycholinguistics*, eds M. J. Spivey, K. McRae, and M. F. Joanisse (Cambridge: Cambridge University Press), 218–235.

Saur, D., Schelter, B., Schnell, S., Kratochvil, D., Küpper, H., Kellmeyer, P., et al. (2010). Combining functional and anatomical connectivity reveals brain networks for auditory language comprehension. *Neuroimage* 49, 3187–3197. doi: 10.1016/j.neuroimage.2009.11.009

Seghier, M. L., Lazeyras, F., Pegna, A. J., Annoni, J. M., Zimine, I., Mayer, E., et al. (2004). Variability of fMRI activation during a phonological and semantic language task in healthy subjects. *Hum. Brain Mapp.* 23, 140–155.

Seghier, M. L., Lee, H. L., Schofield, T., Ellis, C. L., and Price, C. J. (2008). Inter-subject variability in the use of two different neuronal networks for reading aloud familiar words. *Neuroimage* 42, 1226–1236. doi: 10.1016/j.neuroimage.2008.05.029

Seidenberg, M. S. (2012). "Computational models of reading," in *The Cambridge Handbook of Psycholinguistics*, eds M. J. Spivey, K. McRae, and M. F. Joanisse (Cambridge: Cambridge University Press), 186.

Shaywitz, B. A., Shaywitz, S. E., Blachman, B. A., Pugh, K. R., Fulbright, R. K., Skudlarski, P., et al. (2004). Development of left occipitotemporal systems for skilled reading in children after a phonologically-based intervention. *Biol. Psychiatry* 55, 926–933. doi: 10.1016/j.biopsych.2003.12.019

Silva, M., and Cain, K. (2015). The relations between lower and higher level comprehension skills and their role in prediction of early reading comprehension. *J. Educ. Psychol.* 107:321.

Skeide, M. A., and Friederici, A. D. (2016). The ontogeny of the cortical language network. *Nat. Rev. Neurosci.* 17, 323–332.

Spencer, M., Quinn, J. M., and Wagner, R. K. (2014). Specific reading comprehension disability: Major problem, myth, or misnomer? *Learn. Disabil. Res. Pract.* 29, 3–9.

Suárez-Pellicioni, M., and Booth, J. R. (2018). Fluency in symbolic arithmetic refines the approximate number system in parietal cortex. *Hum. Brain Mapp.* 39, 3956–3971. doi: 10.1002/hbm.24223

Suárez-Pellicioni, M., Fuchs, L., and Booth, J. R. (2019). Temporo-frontal activation during phonological processing predicts gains in arithmetic facts in young children. *Dev. Cogn. Neurosci.* 40:100735. doi: 10.1016/j.dcn.2019.100735

Thompson-Schill, S. L., Aguirre, G. K., Desposito, M., and Farah, M. J. (1999). A neural basis for category and modality specificity of semantic knowledge. *Neuropsychologia* 37, 671–676.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., and Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proc. Natl. Acad. Sci. U.S.A.* 94, 14792–14797. doi: 10.1073/pnas.94.26.14792

Tong, Y., Chen, Q., Nichols, T. E., Rasetti, R., Callicott, J. H., Berman, K. F., et al. (2016). Seeking optimal region-of-interest (ROI) single-value summary measures for fMRI studies in imaging genetics. *PLoS One* 11:e0151391. doi: 10.1371/journal.pone.0151391

Turkeltaub, P. E., Gareau, L., Flowers, D. L., Zeffiro, T. A., and Eden, G. F. (2003). Development of neural mechanisms for reading. *Nat. Neurosci.* 6, 767–773.

van der Mark, S., Klaver, P., Bucher, K., Maurer, U., Schulz, E., Brem, S., et al. (2011). The left occipitotemporal system in reading: Disruption of focal fMRI connectivity to left inferior frontal and inferior parietal language areas in children with dyslexia. *Neuroimage* 54, 2426–2436. doi: 10.1016/j.neuroimage.2010.10.002

Vigneau, M., Beaucousin, V., Hervé, P. Y., Duffau, H., Crivello, F., Houde, O., et al. (2006). Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *Neuroimage* 30, 1414–1432. doi: 10.1016/j.neuroimage.2005.11.002

Wagner, R. K., and Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychol. Bull.* 101:192.

Wang, J. (2022). *The Longitudinal Relation Between the Neural Basis of Phonological Awareness and Reading Skill in Developing Children* (Doctoral dissertation). Nashville, Ten: Graduate School of Vanderbilt University.

Wang, J. X., Bartolotti, J., Amaral, L. A., and Booth, J. R. (2013). Changes in task-related functional connectivity across multiple spatial scales are related to reading performance. *PLoS One* 8:e59204. doi: 10.1371/journal.pone.0059204

Wang, J., Joanisse, M. F., and Booth, J. R. (2020). Neural representations of phonology in temporal cortex scaffold longitudinal reading gains in 5-to 7-year-old children. *Neuroimage* 207:116359. doi: 10.1016/j.neuroimage.2019.116359

Wang, J., Pines, J., Joanisse, M., and Booth, J. R. (2021a). Reciprocal relations between reading skill and the neural basis of phonological awareness in 7- to 9-year-old children. *Neuroimage* 236:118083. doi: 10.1016/j.neuroimage.2021.118083

Wang, J., Yamasaki, B. L., Weiss, Y., and Booth, J. R. (2021b). Both frontal and temporal cortex exhibit phonological and semantic specialization during spoken language processing in 7-to 8-year-old children. *Hum. Brain Mapp.* 42, 3534–3546. doi: 10.1002/hbm.25450

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. New York, NY: Harcourt Brace & Company.

Weiss, Y., Cweigenberg, H. G., and Booth, J. R. (2018). Neural specialization of phonological and semantic processing in young children. *Hum. Brain Mapp.* 39, 4334–4348.

Weiss-Croft, L. J., and Baldeweg, T. (2015). Maturation of language networks in children: A systematic review of 22 years of functional MRI. *Neuroimage* 123, 269–281. doi: 10.1016/j.neuroimage.2015.07.046

Welcome, S. E., and Joanisse, M. F. (2012). Individual differences in skilled adult readers reveal dissociable patterns of neural activity associated with component processes of reading. *Brain Lang.* 120, 360–371. doi: 10.1016/j.bandl.2011.12.011

Williams, J. T., Darcy, I., and Newman, S. D. (2015). Modality-independent neural mechanisms for novel phonetic processing. *Brain Res.* 1620, 107–115. doi: 10.1016/j.brainres.2015.05.014

Woodcock, R. W., McGrew, K. S., and Mather, N. (2001). *Woodcock-Johnson III NU Complete*. Rolling Meadows, IL: Riverside Publishing.

Yamasaki, B. L., McGregor, K. K., and Booth, J. R. (2021). Early phonological neural specialization predicts later growth in word reading skills. *Front. Hum. Neurosci.* 15:674119. doi: 10.3389/fnhum.2021.674119

Younger, J. W., Lee, K. W., Demir-Lira, O. E., and Booth, J. R. (2019). Brain lateralization of phonological awareness varies by maternal education. *Dev. Sci.* 22:e12807. doi: 10.1111/desc.12807

Younger, J. W., Tucker-Drob, E., and Booth, J. R. (2017). Longitudinal changes in reading network connectivity related to skill improvement. *Neuroimage* 158, 90–98.

Yu, X., Raney, T., Perdue, M. V., Zuk, J., Ozernov-Palchik, O., Becker, B. L., et al. (2018). Emergence of the neural network underlying phonological processing from the prereading to the emergent reading stage: A longitudinal study. *Hum. Brain Mapp.* 39, 2047–2063. doi: 10.1002/hbm.23985

frontiers | Frontiers in Neuroscience

TYPE Review
PUBLISHED 14 November 2022
DOI 10.3389/fnins.2022.981801

Animal models of developmental dyslexia

Albert M. Galaburda*

Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, United States

OPEN ACCESS

EDITED BY
Charles A. Perfetti,
University of Pittsburgh, United States

REVIEWED BY
Anthony James Krafnick,
Dominican University, United States
Manuel Soriano-Ferrer,
University of Valencia, Spain

*CORRESPONDENCE
Albert M. Galaburda
agalabur@bidmc.harvard.edu

SPECIALTY SECTION
This article was submitted to Neurodevelopment, a section of the journal Frontiers in Neuroscience

RECEIVED 29 June 2022
ACCEPTED 28 October 2022
PUBLISHED 14 November 2022

CITATION
Galaburda AM (2022) Animal models of developmental dyslexia.
Front. Neurosci. 16:981801.
doi: 10.3389/fnins.2022.981801

As some critics have stated, the term "developmental dyslexia" refers to a strictly human disorder, relating to a strictly human capacity – reading – so it cannot be modeled in experimental animals, much less so in lowly rodents. However, two endophenotypes associated with developmental dyslexia are eminently suitable for animal modeling: Cerebral Lateralization, as illustrated by the association between dyslexia and non-righthandedness, and Cerebrocortical Dysfunction, as illustrated by the described abnormal structural anatomy and/or physiology and functional imaging of the dyslexic cerebral cortex. This paper will provide a brief review of these two endophenotypes in human beings with developmental dyslexia and will describe the animal work done in my laboratory and that of others to try to shed light on the etiology of and neural mechanisms underlying developmental dyslexia. Some thought will also be given to future directions of the research.

KEYWORDS

animal models, developmental dyslexia, brain development, endophenotypes, brain asymmetry, ciliopathies, brain anomalies

Introduction

There are many doubts associated with the concept of animal models of human biology and disease, ranging from questioning the degree of molecular, cellular, and higher order homology, to the generalizability and translational potential of animal experimentation to human disease, to ethical considerations regarding animal experimentation, each worthy of serious discussion. These caveats clearly apply to animal models of reading disorder, but I hope to show in this partial review that research findings from animal models of reading disorders still have the potential to shed light on causality, mechanisms, early diagnosis and prevention, and on the design of successful therapies.

In the range of biological levels of representation, from genes and molecules, cell biology and circuits, networks, whole brains and organisms, to cognitive and social systems, non-human and human animals differ least at the first level – genes and molecules – and most at the last – cognitive and social systems. This is illustrated by the example that one can use the same bricks to build schools, supermarkets and post offices, each with very different form and use. One would then be permitted to conclude that animal models are *a priori* most reasonable to pursue for understanding the shared smallest biological units—molecules and cells. Of course, such a conclusion would discourage animal models for understanding reading disabilities, because, although reading acquisition certainly depends on the molecular and cellular integrity of the brain, it cannot happen without the appropriateness and health of higher level structures, such as whole brains and organisms, and social constructs. We understand that social constructs are important because, even as the human brain learns many skills spontaneously, or by imitation, in most cases reading has to be taught, which implies the presence of social structures, including family, teachers and schools, none of which can be modeled in animals. All of this would lead to the conclusion that merely understanding the state of molecules, cells, and circuits is not enough for understanding reading or reading disorders; it is also likely that this situation will not change in the near future, if ever. So, on first inspection, there exist grave restrictions on the utility of animal models for reading and dyslexia. One way out of this conundrum is to focus on preadapted structures and behaviors that are indeed present in animals and humans, which comprise necessary, even if not sufficient, building blocks for the cognitive functions seen only in humans. When these preadapted structures are considered in the genetic context, they are called "endophenotypes" (Gottesman and Gould, 2003).[1]

Endophenotypes studied in animal models have requirements. They must be proven to be reasonable facsimiles of the situation in the human (for a review, see Shanks et al., 2009). This means that at some point it must be shown that the results obtained in the animal are comparable to those that would be obtained in the human were the experiments be able to be performed in the human, and that predictions that come true in the animal model will, within reason, come true for the human. For some questions, this may be a particular challenge in mouse studies, given their phylogenetic distance from the human. For instance, identical or near identical genomic homologies in the mouse and human need not predict for equivalent phenotypes or disease states, as modulation or compensation from other genes or environments in the animal may not be available to the human, or *vice versa*. Thus, it is not uncommon to see that a drug that works in the mouse model fails to work in a human clinical trial (Perrin, 2014). Development and aging have such different time courses in rodents and humans that absence of pathology in the mouse is no guarantee that pathology will not eventually emerge in the human. Also, it may not be possible to mimic all aspects of a disease state in small animals with strikingly different developmental histories, for instance being raised in a mouse cage without social contact, where early experiences can modify the expression of the endophenotype in question (Denenberg, 1981). Here is where the appropriate selection of endophenotypes helps. In that case, a well-chosen endophenotype can shed light on the pathophysiology of the human disorder and can provide ideas for non-invasive testing in the human. In the case of developmental disorders, such as dyslexia, the issue of non-invasive testing becomes particularly important, since often one is dealing with children.

Endophenotypes to model

What makes most sense to model are endophenotypes that *a priori* are more likely to be equivalent in animals and humans. These might include molecular pathways and cellular functions associated with shared dyslexia risk genes, genes that have homologies in both species. Here, even at this low level of representation, care must be taken not to freely generalize from one species to another, since effects of gene manipulation may vary across species and according to the methods used to manipulate gene expression. Based on the known neuroanatomical abnormalities and cognitive deficits in individuals with dyslexia, some preadapted sensory and perceptual behaviors involving the visual and auditory systems, or attention and memory, or laterality, for instance, could comprise suitable behavioral endophenotypes. In our laboratories we chose endophenotypes guided by the original findings in dyslexic autopsy brains—neuronal migration anomalies and anomalous brain asymmetries (Galaburda and Kemper, 1979; Galaburda et al., 1985), which generated additional behavioral research in the animals (see, c.f., Fitch et al., 1994, 1997; Clark et al., 2000a,b).

The first descriptions of structural changes in the brain of dyslexic individuals were described in the 1970s and 1980s on a few dyslexic individuals who had died of unrelated causes (Galaburda and Kemper, 1979; Galaburda et al., 1985, 1994; Humphreys et al., 1990). They ranged in age from the 30s through the 80s, none of them children, and they comprised both men and women. The extent to which the diagnosis was confirmed in life varied, being less secure in the aged individuals. Two types of findings stood out: (1) Subtle disturbances in cortical development, called layer 1 heterotopias, subpial heterotopias, or, simply, ectopias; and (2) abnormal asymmetry of the planum temporale, a region on the upper surface of the temporal lobe associated with language function. In women, the lesions were somewhat later in development and consisted of small, myelinate intracortical scars in the same distribution as the males (Humphreys et al., 1990). In

---

[1] Typically, a measurable, relatively simple biological marker lying between a genomic sequence and a complex behavioral disorder, whereby the marker is likely to predict for the presence of the disorder [coined in the 1970s by U.S. behavioral geneticist and clinical psychologist Irving I. Gottesman (1930–2016) – and British psychiatrist James Shields (1918 – 1978) (Gottesman and Shields, 1972). In animal models an endophenotype must be present in the animal and in the human.

addition, the standard pattern of a left larger planum temporale, seen in two thirds of control human brains (Geschwind and Levitsky, 1968), was not found in any of 7 the dyslexic brains. Attempts were made to confirm the anatomical findings in additional *post-mortem* brains, but this proved to be impossible. Funding to harvest brains in a condition that begins in childhood and normally does not lead to death was simply not forthcoming, and the autopsy project had to be abandoned. Furthermore, the microscopic developmental cortical anomalies, measuring only a few hundred micra in diameter, were not amenable to *in vivo* imaging because the imaging equipment lacked the spatial and contrast resolution needed, although larger structural gray matter heterotopias that are causally related to the smaller anomalies seen in the autopsy brains, have been imaged (Chang et al., 2005). In fact, the difficulty in demonstrating microscopic developmental cortical anomalies in living research participants served as an important stimulus for developing animal models. Another, and perhaps more important reason, was that animal models could be manipulated to test hypotheses about fundamental causes in ways not practically or ethically possible in human studies.

There is a literature about non-right handedness (Abbondanza et al., 2022) and right hemisphere activation (Pugh et al., 2000) for language tasks in dyslexic populations that indicate an aberration in cerebral dominance. Also, a few studies have been published on the issue of structural planum temporale asymmetry as seen in *in vivo* imaging studies (for a review, see Shapleske et al., 1999). In the autopsy studies published by Galaburda and colleagues, the planum temporale was uniformly symmetric, while in the classical study by Geschwind and Levitsky (1968), only 24% of the sample of 100 normal adult human brains showed symmetrical plana. *In vivo* imaging has produced differing results on this topic, which mainly results from slight differences in defining the borders of the planum temporale. One study, that of Altarelli et al. (2014), which defined the planum identically to Geschwind and Levitsky, albeit from MR image reconstructions, rather than from photographs of the upper surface of the temporal lobe in autopsy specimens, found that the asymmetry pattern in the planum of dyslexic brains differed from controls, but only in dyslexic boys, who showed a greater proportion of rightward asymmetrical cases[2]. Also, Heschl's gyrus, which is sometimes duplicated on the right side in control brains, is significantly more often duplicated in dyslexic boys. The difference between the Geschwind and Levitsky and the Altarelli et al. (2014) findings is not understood, although both findings support an anomaly in the manifestation of asymmetry of a language area in the dyslexic brain, at least in boys and men. And, at least in boys, there is a deviation in the distribution of planum asymmetry, which, together with reports of an abnormal distribution of hand dominance in the dyslexic population (Abbondanza et al., 2022), make modeling asymmetry and laterality in animals potentially productive.

Non-human animals display individual paw or claw preference and a directional preference for body rotation and circling behavior. For instance, caged rats will preferentially hug the right or the left wall when exploring the cage (Denenberg, 1981; Glick and Ross, 1981), but show less of a population bias than humans (Glick and Ross, 1981). On the other hand, humans display a strong tendency to right handedness at the population level. This tendency to right-handedness is matched to a large extent by left-hemisphere dominance for speech and language (Knecht et al., 2000). The mechanisms of handedness and language lateralization are not known, but cilia may play a role. Cilia are short, microscopic, hairlike cellular structures that are responsible for the left-right body patterning that results in a left sided-heart and a right sided liver, for instance (Dasgupta and Amack, 2016), but do not easily explain brain laterality (Trulioff et al., 2017). Cilia can be rendered dysfunctional by suppressing the expression of some of the dyslexia-associated genes, and, as they are conserved between humans and animals, they can comprise a useful endophenotype to model in the study of dyslexia.

In addition to asymmetries and developmental cortical malformations, there exist behavioral characteristics displayed by individuals with dyslexia that could also be amenable to animal modeling. Thus, even though the reading disability *per se* cannot be modeled, for obvious reasons, there are some sensory-perceptual traits underlying reading that can be. So, for instance, dyslexic persons have been shown to exhibit phonological deficits as a result, at least in part, of abnormal sound processing at levels lower than the cognitive and cortex (Hornickel and Kraus, 2013; Neef et al., 2017). The idea is that if sounds are not processed properly, which includes the processing of speech relevant sounds, then abnormal phonological representations arise, which represent a barrier to learning to read easily. Phonology is a complex term that includes both speech sound representations (phonetics) and phonological grammar (i.e., rules for combining a limited number of speech sounds to produce unlimited words and meanings). It appears that at least the phonetics part of phonology lends itself to investigation in animals; for instance, it was shown more than fifty years ago that chinchillas can be taught to make speech sound distinctions (Kuhl and Miller, 1971). Furthermore, the phonological grammar appears to be spared in dyslexia (Berent et al., 2012, 2013).

---

[2] Considering the different neuropathology described in men and women with dyslexia, as well as the Altarelli et al. (2014) finding of normal patterns of planum asymmetry in the women and girls in their study, could it be that dyslexia is a different disorder in men and women?

Animal models of dyslexia

Neonatal freezing injury model

Animal studies in this field have exploited the freezing injury rat model, the short hairpin RNA interference (shRNAi) rat model, and the gene deletion (knockout) mouse model. A non-exhaustive review of these studies is presented below for the purpose of illustrating the kinds of discoveries that can be made from using such models.

The neonatal freezing injury model consisted in placing lesions in the developing cortex during the neonatal period, before neuronal migration to the cortex ends. Initially, we used rats prepared by the method established by Dvorak, Feit and Juránková (Dvorak et al., 1978). First, on day one or two after birth, when neuronal migration to the cerebral cortex is still proceeding but reaching the end as the upper layer neurons finish their migration, we apply a freezing probe to the skull of the newborn. Depending on the duration of the probe application, a molecular layer ectopia, a 4-layer microgyrus, or frank porencephaly is produced (Humphreys et al., 1991; Suzuki and Choi, 1991; Rosen et al., 1992, 2000). The coexistence of these very different-looking abnormalities has been recognized in abnormal human brain development for many years (Friede, 1989), so these malformations are considered causally linked. Although molecular layer ectopias were the main finding in the autopsied dyslexic brains, there were also a couple of instances of microgyria, but no instances of porencephaly; porencephaly in the perisylvian language cortex is a more severe lesion and would more likely present with speech and/or language delay and epilepsy as part of the perisylvian syndrome (Kuzniecky et al., 1993). It is not farfetched, then, to hypothesize that microgyria and molecular layer ectopias, by virtue of being a milder pathology, would be associated with more subtle cognitive deficits, *e.g.*, dyslexia, and techniques are available to mimic these pathologies in the developing rat.

Rats with freezing lesions[3] start as normal animals. Starting with a normal animal, anything done to it in the laboratory represents the initial event, thus the cause of what happens subsequently. It is important to stress that the brain reacts to the initial event by a process known as plasticity.[4] However, the reaction to the initial event need not make things better, and anatomical, physiological, and behavioral abnormalities documented later may be the result of this plasticity plus the initial event, rather than the initial event alone, thus making the plasticity potentially a maladaptive rather than an ameliorative phenomenon. This negative effect may be particularly true for very early lesions, in violation of the so-called Kennard Principle (Schneider, 1979; Johnston, 2004; Elliott, 2020). In the case of the rat with the freezing lesion, the injury triggers plasticity in connections and in the cell composition of connectionally related areas (Rosen et al., 1998; Li et al., 2021). In rats with shRNAi and in the knock-out mouse, epigenetic changes are triggered in other genes that are part of the injured gene's network (Che et al., 2014, 2016). The rat freezing injury model has the additional advantage that it has available to it a larger repertoire of behaviors that can be tested in the laboratory, as compared to a substantially reduced repertoire in the mouse. Working with the rat, whether in a freezing injury model or using shRNAi gene knockdown, affects development during late neuronal migration to the neocortex, whereas the gene deletion in the mouse knockout is earlier and precedes neuronal migration. Thus, in addition to the species-specific differences, there is a developmental timing difference that needs to be taken into consideration when interpreting differences in outcomes. That said, even though the freezing injury rat model illustrates the enormous plasticity of the developing brain, there has never been any evidence that dyslexia in humans arises from an episode of intrauterine brain injury, whether traumatic, vascular, infectious, metabolic or other. Instead, there has been growing evidence that variants of certain genes that are expressed in the brain during development contribute significant risk for dyslexia. Therefore, when genetic epidemiological studies began to identify these risk genes, our laboratories retooled to study them in gene-based animal models.

Genetics, dyslexia, and animal research

In the past 20 years, several dyslexia risk genes have been discovered around the world. The first of these genes was *DYX1C1*, followed closely by *DCDC2, KIAA0319, ROBO1, EKN1* (Paracchini et al., 2007).[5] These are called risk, or susceptibility, genes, because they do not predict for a particular phenotype or disorder to arise, but rather for the risk that it will arise. Their discovery involves large scale population studies and statistical associations between the presence of a genetic marker on a chromosome and the presence of a phenotype or disorder. After identifying the marker additional work needs be done to identify the gene and the mutation or variant, and still more work to understand its functions. In many, if not the majority of situations, the variant associated with the condition does not involve the protein coding part of the gene, the exon, but rather a regulatory segment, such as a transcription factor that modulates timing and degree of expression of the exon,

---

[3] This is also the case for rats undergoing shRNAi experiments and for mice with gene deletions.

[4] The term plasticity often conjures up the notion of recovery after injury (Kolb, 2003), but positive outcomes of plasticity do not necessarily follow (Johnston, 2004).

[5] By convention, human genes are italicized in all caps; mouse genes are italicized with only the first letter in cap; proteins for both humans and animals are in all caps without italics.

the so-called epigenetic activity. In many cases, initially it is difficult to see how a particular gene variant or mutation, and its downstream effects, lead to the phenotype of interest, but this discovery is made easier (but not easy!) if the gene in question is expressed in the organ of interest and during the time the science suggests the phenotype originates. In the case of dyslexia, based on what we know about the brain, the gene would at least have to be expressed in the brain during the time of neuronal migration to the cortex, but not necessarily in the developing cortex, since the cortical changes could be secondary to an initial event at other sites. However, it would be surprising, if not embarrassing, to discover that a statistically identified risk gene for dyslexia is only expressed in the liver during senescence!

Genes for dyslexia have effects on human brain development, but it remains a challenge knowing how these effects lead to reading disabilities. The functions of these genes are mainly known from work on cell preparations, rodents, fish, flies and worms, which adds a layer or more of separation from the problem at hand in the human. Furthermore, in general the dyslexia risk genes are broadly expressed in the animal brain and human brain in neurons, so a deficit in a narrow set of cognitive domains, say hearing, vision, language and reading, does not easily follow from such a broad neural distribution, which instead may predict for general intellectual disability, motor and sensory deficits, and/or epilepsy. At the writing of this paper, this conundrum remains an important challenge to the science, but it can be argued that continued work on animal models is likely eventually to provide at least some of the answers (also see below in the discussion section).

DYX1C1

DYX1C1 was the first reported dyslexia susceptibility gene (Taipale et al., 2003). *Currently* termed DNAAF4 (dynein axonemal assembly factor 4[6]), this gene encodes a tetratricopeptide repeat (TPR) domain-containing protein. TPR is a broadly occurring structural motif that helps with protein-protein interactions and the assembly of multiprotein structures and has been linked to several disorders, including primary ciliary dyskinesia (Loges and Omran, 2018), whereby cilia are involved in neuronal migration, particularly interneuron migration, although their role in excitatory neuron migration cannot yet be excluded (Guemez-Gamboa et al., 2014). A chromosomal translocation involving *DYX1C1* confers a susceptibility to developmental dyslexia. Multiple, focal neuronal migration abnormalities primarily in the left perisylvian (language) cortex comprised the most prominent finding in several brains of dyslexic individuals studied at autopsy (Galaburda and Kemper, 1979; Galaburda et al., 1985). So, it was particularly reassuring to discover that suppression of *Dyx1c1* protein translation in the rat by short-hairpin RNA interference (shRNAi) in the late fetal period caused neuronal migration anomalies of cortical projection neurons arising in the ventricular zone (Rosen et al., 2007). Clumps of neurons remained in the ventricular zone, while others over-migrated beyond the layers that would normally accommodate them (Currier et al., 2011). Abnormalities were not restricted to the cerebral cortex. In fact, RNAi transfected rats displayed changes in the medial geniculate nucleus (MGN), with a significant shift to smaller MGN neurons (Szalkowski et al., 2013); autopsied dyslexic brains had shown the same findings in the MGN (Galaburda et al., 1994).

The demonstrated role of *Dyx1c1* varies according to experimental condition. Thus, even though shRNAi interference in rats in late gestation causes cerebro-cortical neuronal migration abnormalities, deletion of exons 2-4 of *Dyx1c1* in mice (*Dyx1c1v* knockout mice), which also renders the gene non-functional, albeit earlier in development, soon after fertilization of the ovum, does not (Chandrasekar et al., 2013; Tarkar et al., 2013*);* also, see below); instead, *Dyx1c1v* knockouts display a phenotype that is reminiscent of human primary ciliary dyskinesia, a disorder characterized by chronic airway disease, laterality defects (*situs inversus*), and male infertility (Lee and Gleeson, 2011; Chandrasekar et al., 2013; Tarkar et al., 2013; Loges and Omran, 2018; Anvarian et al., 2019; Hasenpusch-Theil and Theil, 2021). These knockout mice die soon after birth with hydrocephalus and display situs inversus. Hydrocephalus is an accumulation of cerebrospinal fluid with resultant enlargement of the ventricular system, which implicates dysfunction of the ependymal cell cilia, which are thought to help mobilize the cerebrospinal fluid for resorption (Kumar et al., 2021). In the zebrafish, cilia are present in the Kupffer vesicle (Chandrasekar et al., 2013), which is involved in left-right brain development. However, although cilia are also present in the central nervous system of mammals beyond the empendymal cells, there is no proven relationship between primary cilia dyskinesia and disturbances of cerebral laterality in humans, although in the case of situs inversus without cilia dysfunction, left-handedness has been reported to be increased (Postema et al., 2020). Nevertheless, cilia dysfunction cannot at present clearly explain variations in cerebral asymmetry and increased non-righthandedness among dyslexic individuals. On the other hand, cilia have been implicated in neuronal migration to the cerebral cortex, particularly interneurons migrating tangentially from the ventral germinal zones. Less is known about the radial migration of pyramidal neurons from the ventricular zones, and a portion of patients with the Meckel Gruber Syndrome and Joubert Syndrome, both involving cilia biology, develop heterotopias and other neuronal migration abnormalities (for an excellent review of the role of cilia in neuronal migration, please see Hasenpusch-Theil and Theil, 2021).

---

[6] https://www.ncbi.nlm.nih.gov/gene/161582

One could hypothesize, that cilia dysfunction in dyslexics carrying the *DYX1C1* variant contribute to the neuronal migration defect but also impedes a directional gradient of patterning molecules, which would, in turn, lead to aberrant cerebral asymmetry at the molecular, cellular and perhaps also circuit levels, not yet amenable to demonstration by current *in vivo* tools for human research, let alone in clinical work. In fact, understanding normal and aberrant cerebral asymmetry remains a challenge. Our older studies in rats with experimental cortical microgyria, a type of neuronal migration anomaly described in dyslexia, demonstrated changes in both intra- and interhemispheric connectivity (Rosen et al., 2000), with a theoretical capability of altering patterns of intra and interhemispheric communication, and, thus, lateralization of function. A comparable effect on callosal connections altering lateralization has been suggested for loss of ROBO2 function, another gene implicated in reading disorders in rare families (Hannula-Jouppi et al., 2005). In individuals with dyslexia carrying any of several dyslexia risk genes, the volume of cortical white matter seems to be a predictor of reading comprehension (Darki et al., 2012; Eicher and Gruen, 2013) and alterations in asymmetry of brain activation (Pinel et al., 2012) are seen with the same dyslexia-related polymorphisms. We posit that reorientation of cortical white matter connections in a (seemingly futile) attempt to compensate for the presence of abnormal developmental targets (the malformations), leads to the changes in white matter volumes seen in the imaging studies and in alterations in cerebral lateralization and brain activation during language tasks (but see the glutamatergic hypothesis, below).

DCDC2

A member of the doublecortin superfamily of genes (Reiner et al., 2006), some of which have been linked to abnormal neuronal migration, epilepsy, blindness, and general intellectual disability, *DCDC2* has also been linked to dyslexia (Meng et al., 2005 and others; but see Scerri et al., 2017). This gene serves as a protein-interaction platform (Reiner et al., 2006), where the doublecortin domain binds tubulin and enhances microtubule polymerization. Microtubules are filamentous intracellular structures that are responsible for various kinds of cell movements, including intracellular transport, axon extension and neuronal migration; microtubules are also implicated in the assembly and signaling of primary cilia. Additional functions of *DCDC2* include dendrite morphogenesis, neuronal action potentials, Wnt signaling, sound perception, and excitatory (glutamatergic) synaptic transmission (Massinen et al., 2011; Che et al., 2016[7]).

Elevated glutamate levels were previously found in attention deficit/hyperactivity disorder (Carrey et al., 2007) and autism (Brown et al., 2013) and have more recently been associated with individual differences in reading ability in young readers (Pugh et al., 2014). Our collaborators showed that *Dcdc2* deletion in mice was accompanied by increased excitability and decreased temporal precision in action potential firing in the cortex (Che et al., 2014, 2016). Furthermore, the decreased action potential temporal precision could be fully restored in mutants by treatment with either the NMDA receptor antagonist (2R)-amino-5-phosphonovaleric acid or the NMDAR 2B subunit–specific antagonist Ro 25-6981 (Che et al., 2014). Precise timing of neuronal firing is likely to be essential for representing speech sounds, some of which require a temporal resolution of only a few milliseconds. A deficit in precise firing could explain a tendency for phonological deficits, on the one hand, and, on the other, absence of other perceptual and cognitive deficits that do not depend on precise, rapid neuronal firing. In this way, a ubiquitous neuronal dysfunction could affect one or a few cognitive/perceptual functions, while leaving others intact. This is a testable hypothesis that can help answer the question of why a dysfunction that can affect most neurons can present with a focal behavioral disorder.

As noted previously, under some experimental conditions, neuronal migration anomalies occur when the function of dyslexia risk gene homologs is suppressed *in utero*. An interesting observation was made when rats were transfected with Dcdc2 shRNA, which silences the gene for a few days. Both undermigration and overmigration of cortical neurons were seen, but, whereas over-migration of transfected neurons occurred with transfection late in the intrauterine period, overmigration did not occur with earlier transfection (Adlerr et al., 2013). This difference suggested that compensation could occur in this endophenotype if the gene silencing was early, but not late. This is à propos of reports, and our own results, this time in *Dcdc2* knockout mice, that failed to show migration anomalies. In the knockout, the gene silencing starts much earlier, and the experiments are carried out in mice instead of rats, where species differences may also play a role. It has also been suggested that *Dcdc2* has a role in neuronal migration only when doublecortin is inhibited, whereby deletion of *Dcdc2* increased the severity of the deficits of neuronal migration caused by RNA interference of doublecortin (Wang et al., 2011).

Human carriers of the rs793842 polymorphism of *DCDC2* show a negative correlation between white matter volume and reading comprehension, as well as thickening of the cortex over the left angular and supramarginal gyri (Darki et al., 2014), areas that participate in language and reading. However, excessive glutamatergic activity or hyperexcitability (see above) would be expected to cause increased excitotoxic apoptosis of neurons and oligodendrocytes leading to cortical atrophy; therefore, the cortical thickening remains unexplained, particularly in the parietal lobes, which are particularly vulnerable to

---

[7] https://www.uniprot.org/uniprot/Q5DU00

excitotoxicity. Thus, Alzheimer's disease, a condition associated with excitotoxicity and cell death, shows early atrophic changes in the parietal lobes (Jacobs et al., 2012). The white matter reduction is more easily explained by the special vulnerability of oligodendrocytes to glutamatergic excitotoxicity (see Matute et al., 2007). Increased cortical thickness need not imply better function in a phrenological sense. A thicker cortex can be seen in developmental malformations, such as polymicrogyria (also referred to as "micropolygyria"), in part due to centripetal collapse of 4-layer microgyric cortex and blurring of the cortical-subcortical border and/or decreased developmental neuronal and dendritic pruning. Furthermore, one dyslexia-associated gene variant of *Robo1* causes increased interneuron migration to the cortex (Andrews et al., 2006), which could be another source for the thickening seen, leading to increased intracortical circuits but no increase in longer cortico-cortical pathways. Interneurons' main neurotransmitter is gamma aminobutyric acid (GABA), which is initially excitatory and trophic and switches to inhibitory later in development, the date determined by the degree of GABA activity and blockade (Ganguly et al., 2001). Significant functional changes in the cortex would then be expected by a process that increases migration of GABAergic interneurons to the cortex. Support for the hypothesis of a thicker albeit dysfunctional cortex comes from MRI activation studies showing that posterior left temporoparietal reading related areas (Meda et al., 2008) activate less strongly during reading tasks (Cope et al., 2012; Eicher and Gruen, 2013; D'Mello and Gabrieli, 2018; Richlan, 2020).

KIAA0319

KIAA0319 is a transmembrane protein coded by *KIAA0319*, on chromosome site 6p22.2, with relevant expression in the central nervous system, pituitary, and peripheral nervous system (Franquinho et al., 2017[8]). The gene has been extensively studied in human populations *vis à vis* language, reading and cerebral lateralization [see review by Eberli et al. (2021)]. The gene was linked to dyslexia, and its expression was shown to be reduced in individuals carrying a risk haplotype that included *KIAA0319* (Cope et al., 2005; Paracchini et al., 2006). Expression of the other two genes in the haplotype, the TTRAP gene and portions of THEM2, was not reduced, thus pointing the finger to *KIAA0319* (Paracchini et al., 2006). In rat studies, it has been shown that the protein is involved in neuronal migration during cerebro-cortical development *in utero* (Peschansky et al., 2010; Adlerr et al., 2013; Platt et al., 2013; but see Guidi et al., 2017 in mice). *KIAA0319* may function in a cell autonomous and a non-cell autonomous manner and plays a role in appropriate adhesion between migrating neurons and radial glial fibers during neuronal migration (see text footnote 3). It may also regulate growth and differentiation of dendrites. Thus, negative regulation of axon extension and dendrite development has been demonstrated, as well as effect on auditory responses.[9]

Our group used *in utero* electroporation (Peschansky et al., 2010; Platt et al., 2013) to transfect cells in E15/16 rat neocortical ventricular zone with either shRNA vectors targeting Kiaa0319, with a KIAA0319 expression construct, with a Kiaa0319 shRNA along with KIAA0319 expression construct ("rescue condition"), or with a scrambled version of Kiaa0319 shRNA. Knockdown, but not overexpression, of Kiaa0319 resulted in periventricular heterotopias that contained large numbers of both transfected and non-transfected neurons, the latter considered a non-cell autonomous effect on neuronal migration. Of the Kiaa0319 shRNA– transfected neurons that migrated into the cortical plate, most migrated to their appropriate laminae. In contrast, neurons transfected with the KIAA0319 expression vector attained laminar positions subjacent to their expected positions, indicating that both under- and over-expression of the gene affected neuronal migration. Furthermore, neurons transfected with Kiaa0319 shRNA exhibited apical, but not basal, dendrite hypertrophy. The rescue conditions were successful in inhibiting the migrational and dendritic effects of under- and over-expression, which is a method for excluding off-target effects of the transfection. Off-target effects occur when a short vector contains a sequence that is found not only in the target gene, but also in another unknown gene or genes. In that case there is the danger of interpreting the phenotype as resulting from an effect on the target gene, when in fact it results from effects on some unknown gene sharing the same short sequence. Restitution of the known protein by overexpression would work only on the target gene and is a necessary step for excluding off-target effects. On the other hand, comparable effects were not noted in the mouse undergoing gene deletion (mouse knockouts), which led to controversy (Franquinho et al., 2017; Guidi et al., 2017, Guidi et al., 2018; Martinez-Garay et al., 2017): Does *KIAA0319* have anything to do with neuronal migration? For this writer, finding neuronal migration anomalies is more telling than not finding them, when the research has controlled for off-target effects and other artifacts, unless it can be shown that the process for looking for neuronal migration anomalies itself causes them to appear; this has not been shown to be the case in the rats undergoing shRNAi. On the other hand, one can come up with reasons why anomalies may not arise, especially when the counterexample involves an altogether different species and methodology (see above). Here is a situation where animal studies can shed both light and confusion on the real question, which is whether a genetic variant is responsible for a specific endophenotype in humans. Of additional interest is the fact that suppression of gene expression in the *Dcdc2* knockout mouse still produces abnormal cortical physiology, which illustrates the

---

[8] https://www.proteinatlas.org/ENSG00000137261-KIAA0319

[9] https://www.uniprot.org/uniprot/Q5SZV5

possibility that the neuronal migration anomaly may be only a marker for a more important underlying cortical dysfunction that can exist even in the absence of the marker.

A recent *in vivo* and *post-mortem* study in chimpanzees established a relationship between KIAA0319 variants and gray matter volume in the posterior superior temporal gyrus, as well as neuropil asymmetries in the same region under microscopic examination (Hopkins et al., 2021), suggesting an evolutionary influence by KIAA019 on auditory processing preceding the evolution of language in the primate line. In the absence of linguistic capacities in non-human primates, this effect of KIAA019 supports the notion that dyslexia-related genes are not directed at reading or language *in utero*, but rather to preadapted acoustic endophenotypes that in humans comprise some of the building blocks of language and reading acquisition and efficiency.

In a continuing attempt to make the rodent model as naturalistic as possible, we and others focused on a gene deletion models, known as knock-outs, in the mouse. First, unlike knock-down of gene expression in the rat by shRNAi, deletion of dyslexia risk homologs in the mouse do not result in neuronal migration abnormalities. Instead, deletion of exons 2-4 of *Dyx1c1* in the mouse, which eliminates protein translation, was associated with abnormalities in cilia structure, growth, and function (Chandrasekar et al., 2013; Tarkar et al., 2013). Abnormalities in cilia structure and function were also seen in association with *Dcdc2* dysregulation (Massinen et al., 2011), and a missense mutation in *DCDC2* is known to cause deafness in humans, likely associated with cochlear cilia abnormalities (Grati et al., 2015). Primary ciliopathies are also associated with hearing loss, underscoring the importance of cilia for auditory function. Sonic hedgehog signaling dysregulation causes hearing loss in ciliopathy mouse models (Moon et al., 2020), and Dcdc2 interacts with sonic hedgehog signaling (Massinen et al., 2011). *Kiaa0319* modifications altered axonal growth (Franquinho et al., 2017), and gene overexpression in cortex delayed radial migration, but did not change the pattern of cortical lamination. Similarly, a cell knockout model showed that cilia exhibited increased length and changes in cell migration (Diaz et al., 2022). Finally, *Kiaa0319* knockout animals showed subtle alterations in anxiety-related behavior and in sensorimotor gating (Martinez-Garay et al., 2017).

Other genes

Other genes have been linked to dyslexia. For instance, ROBO1 affects auditory and visual motion processing that predict for reading achievement (Mascheretti et al., 2020) and vocal learning in animals (Wang et al., 2015); the gene has been associated with increased interneuron migration into the cerebral cortex, as well as altered inter and intrahemispheric connectivity (Andrews et al., 2006). Homozygous deletions of *Robo1* in the mouse are also associated with occasional heterotopias [Anthoni et al., 2012; also see review by Gonda et al. (2020)]. Two other genes, *TTRAP* and *THEM2,* are part of the dyslexia risk haplotype that also contains *KIAA0319* on chromosome 6p22.2 and are often included on a list of dyslexia-risk genes. However, the risk haplotype is associated with decreased expression of *KIAA0319*, but not *TTRAP* or *THEM2* (Paracchini et al., 2006). The aromatase gene *CYP19A1* has also been linked to dyslexia (Anthoni et al., 2012), which is interesting, as aromatase determines the conversion of testosterone to estradiol, two sex steroids, and most studies have shown that there is a significant and substantial difference in the prevalence of dyslexia between boys and girls[10]. Neuron specific aromatase has a role in synaptic plasticity and cognitive function in both mouse sexes, and more work is needed to differentiate its effects in males and females (Lu et al., 2019). Expression of *CYP19A1* correlates with expression of dyslexia-risk genes *DYX1C1* and *ROBO1* raising questions as to whether *CYP19A1* acts independently on dyslexia risk. Aromatase has effects on dendritic growth, so an independent role is not excluded, even if not directly proven at present. A study involving Finnish families and an independent study of German families identified a haplotype containing co-regulated genes *C2orf3* and *MRPL19* on chromosome 2p12. The expression of these genes, but not of *FLJ13391* (also in the haplotype) was correlated with the expression of genes *DYX1C1, ROBO1, DCDC2* and *KIAA0319* (Anthoni et al., 2007). No association was found for these genes in a study of Indian families, nor for *ROBO1* or *THEM2* (Venkatesh et al., 2013). Additional animal studies would be useful here to understand the molecular pathways involved and the effects of downregulating the expression of these candidate genes, better to understand possible links to dyslexia.

Genetics and behavior

Most of the studies linking gene and behavior in dyslexia have been performed in humans, often together with *in vivo* functional brain imaging or neurophysiology to link to brain anatomy and/or behavior. Although such studies are good for establishing correlations, and language and reading can be explored directly, it is much more difficult to make statements about first events and causation. So, is what we learn from those studies something about the cause of the dyslexia or a reflection of the reading problem after years of brain plasticity? A partial answer can be obtained by looking for a phenotype in the youngest person possible to study [see, for example, the work

---

[10] This finding implicates a sex hormone effect. Not counting substantial social and environmental influences, there exist sex differences based on the presence or absence of the y chromosome, but this would dictate that the phenotype be present only in boys and men, e.g., ear hair.

of Gaab and colleagues (Raschle et al., 2011)].[11] Animal studies, which permit manipulation of genes or the brain in ways not possible in human beings, are superior for looking at the earliest events and for establishing causation, even in the face of the limitations of animal research discussed in the introduction. In fact, animal studies can help differentiate between causal events and subsequent plasticity changes.

Male rats with bilateral freezing lesions to the cortex, which develop focal microgyria, exhibit difficulties in discriminating two sequential tones that occur 332 msecs or less from each other. Male rats with unilateral induction of microgyria are abnormal at a shorter gap, 249 msecs, compared to control animals with sham interventions (Fitch et al., 1994, 1997; Clark et al., 2000a, b). Female rats exposed to the same treatments failed to show a reduced capacity to distinguish rapidly changing sounds (Clark et al., 2000b), even though quantitative analysis of the anatomical changes in the cortex did not disclose any sex differences. Therefore, female rats appear to be more resistant to the behavioral effects of early brain damage in this specific domain, which in turn may help explain sex differences in the incidence of dyslexia and other neurodevelopmental disorders in humans (Krafnick and Evans, 2019; Romeo et al., 2022). In other words, it is not necessarily the case that females are at a lesser risk of exposure to the causal event, but rather they are more likely to react adaptively compared to the males. The cortical microgyria were not different between the sexes, but plasticity effects differed between male and female rats, with males, but not females, showing a shift toward more small neurons (slow neurons?; Goriounova et al., 2018) in the medial geniculate (auditory) nucleus of the thalamus. This raised the question of whether the thalamus, but not the cortex, is critical for acoustic gap detection (see, *c.f.*, Díaz et al., 2012). Thus, an important benefit of the animal model can be to expand the thinking about the mechanisms involved in dyslexia deficits to subcortical areas, while placing less emphasis on the cerebral cortex. The subcortex is important for skill acquisition (Chen et al., 2021), and there is evidence for involvement of the subcortex, including the brainstem, in dyslexia (Hornickel and Kraus, 2013). In the case where a developmental cognitive disorder implicates both the cerebral cortex and subcortical stations, another benefit of animal models would be to help determine whether the problem begins in the subcortex and spread to the cortex, it starts at multiple sites at the same time, or whether the subcortex represents a secondary change following disruption of cortical development. In the latter case, the plasticity, and not the initial change in the cortex, would be responsible for the deficit. It is possible currently to conditionally delete a gene at a selective location, and at a particular time, to help answer this question. In the case of the freezing lesion induced cortical malformation, unpublished results in rats with freezing lesions showed cell composition changes in the thalamus, but also in the cochlear nucleus in the brainstem, again suggesting that the spread from the induced cortical malformation can reach the earliest stages of auditory representation in the central nervous system. It is much more unlikely that a pathology beginning in the brainstem can developmentally propagate rostrad and lead to neuronal migration abnormalities, although brainstem pathology can certainly lead to functional changes in the cortex. Malformations can, however, arise in the brainstem and cortex at the same time (Barkovich, 2012), but we did not see brainstem malformations in the human cases or in any of the animal models that we have used.

The first behavioral genetic model we tried was in rats, with which our collaborators had extensive experience in studying behavior. The choice of the rat as an experimental model had to do with its more extensive behavioral repertoire than the mouse, and because at that time no knockouts were available in mice. The rats had their dyslexia gene homologs suppressed by transfecting with short hairpin RNA interfering constructs. The first gene we suppressed was *Dyx1c1*. The intervention, which caused focal heterotopias, led to deficits in detecting complex auditory stimuli over time (Threlkeld et al., 2007). Auditory processing deficits were seen in male and female rats (Szalkowski et al., 2013). In addition, those animals that also showed heterotopias in hippocampus had deficits in spatial learning (Threlkeld et al., 2007). Additional subtle, but persistent, working memory deficits were demonstrated in Sprague-Dawley rats suppressed with shRNAi to *Dyx1c1* (Szalkowski et al., 2011). In a subsequent study, *Dyx1c1* suppression in rats, in addition to acoustic processing deficits, impaired visual attention in males, without changes in total cortical volume, hippocampal volume, mid-sagittal callosal volume. On the other hand, there were significant changes in the medial geniculate nucleus, with a switch to greater proportions of smaller neurons (Szalkowski et al., 2013).

As with *Dyx1c1*, *in utero* suppression of *Kiaa0319* in rats produces deficits in speech sound discrimination. The experimental animals needed twice as much training in quiet conditions to perform at control levels and remained impaired at several speech tasks (Centanni et al., 2014a). Training using modified speech sounds was able to normalize speech discrimination and physiology (Centanni et al., 2014a). In a separate experiment, the authors reported that with reduced *Kiaa0319* intracellular recordings from affected neurons showed increased neural excitability and input resistance (Centanni et al., 2014b). shRNAi-mediated knockdown of the homolog of the dyslexia risk gene *DCDC2* in the rat resulted in impaired speech sound discrimination without abnormal responses to sound in the primary auditory cortex (Centanni et al., 2016).

---

[11] "Partial answer" refers to the fact that, although anatomical differences can be imaged before the onset of reading, which indicates that they are not caused by reading differences, the imaging cannot tell how early the differences arise, which is likely to be *in utero* and before the onset of speech itself.

These results contrasted with those found in *Kiaa0319* RNAi, which degrades cortical activity to speech sound (Centanni et al., 2014a). The authors emphasized that different dyslexia risk genes affect the speech processing circuits differently. These deficits could not be confirmed in knockout mice for *Kiaa0319,* but double knockout of *Kiaa0319* and *Kiaa0319l* resulted in deficits in central and peripheral auditory function. Deletion of *Kiaa0319l* alone caused abnormalities in the brainstem acoustic wave (Guidi et al., 2017). This is interesting, because brainstem acoustic responses have been documented to be abnormal in dyslexic individuals (Hornickel and Kraus, 2013), and unpublished findings from our laboratory documented abnormalities in neuronal composition in the human and rodent cochlear nucleus.

Discussion and suggestions for future research

It is clear that animal models offer a limited, albeit important contribution to the understanding of reading disorders, even as such disorders affect only human beings. Endophenotypes such as developmental cortical anomalies and cerebral asymmetries are amenable to modeling even in rodents, as are behavioral endophenotypes involving functional lateralization, sound processing and visual perception. At the cellular level, neuronal hyperexcitability and abnormalities of cilia structure and function occur from dysfunction of dyslexia risk genes in humans and animals. Yet, despite the demonstrated value of animal work, most of the currently funded dyslexia research focuses on human behavior and brain imaging. The value of such research is not in question, but the approaches cannot get directly at the cause of the problem, and therefore cannot link up to powerful available methods for prevention and treatment.

Another limitation of the current human research is its almost exclusive focus on cortical anatomy and physiology and its accompanying behaviors. Thus, although the cerebral cortex is important for language function in adults, and dyslexia in most cases implicates language function, language acquisition requires hearing the sounds of the native language (the congenitally deaf excluded), which begins in infancy (or even *in utero*), and which depends on lower level acoustic processing taking place in the thalamus and brainstem. In the end, if corrupted signals reach the cortex, language can develop abnormally. For speech signals to arrive in the cortex normally, an intact auditory brainstem and thalamus is required, and there is evidence, both from human anatomy and dyslexia animal models, that this may not be the case in dyslexia Tschentscher et al. (2019). That said, *in vivo* imaging the anatomy and function of the human brainstem at the resolution level implicated by the microanatomical studies remains a challenge that relatively few investigators tackle (Tracey and Iannetti, 2006; Beissner et al., 2014; Adil et al., 2021; Lechanoine et al., 2021). Furthermore, although abnormalities in acoustic brainstem physiology in dyslexia has been amply documented (Hornickel and Kraus, 2013; White-Schwoch et al., 2015; Neef et al., 2017), interest in the brainstem's role during early development in the risk for dyslexia has not grown as it should. In the end, even if it is this cortical dysfunction that accounts for the core symptoms in dyslexia, it is important to know how that dysfunction arose and how to prevent it. Here is an area where animal models can be particularly useful.

Genes that provide increased risk for dyslexia are expressed widely in the brain. But, looking at the pattern of expression alone does not provide useful information for figuring out what is going on. So, for instance, if the expression of an anomalous gene leads to increase noise in neural responses to stimuli, it is not likely that this will affect all higher-level functions equally, but rather only those functions that require precise timing, e.g., phonological processing. In other words, hitting neurons that are a part of systems that do not deal with precise timing will not produce noticeable changes in behaviors. The acoustic system is one of the fastest processors in the brain, if not the fastest. It has to be capable of representing stimuli that differ from each other by only a few milliseconds. This is the sort of difference that distinguishes the sound/b/from the sound/p/. Failure to do this may lead to degraded representations of both sounds and thus introduce an additional difficulty for mapping a sound to a letter while attempting to read. Young readers depend much more on this ability in order to read, since adults eventually graduate from letter by letter reading when they are capable of using efficient top-down mechanisms to divine the word without actually having to read it (unless it is a new word or the context is ambiguous and unhelpful). In fact, it is quite clear that those dyslexics who compensate for their earlier reading difficulties do it by relying of top-down, executive processes that avoid having to decode words letter by letter. A corollary would be that dyslexics who cannot compensate as they grow may suffer from executive dysfunction (Brosnan et al., 2016; Smith-Park et al., 2016).

The emphasis on subcortical system concerns the origin of the dyslexia risk in the brain. Developmental plasticity dictates that secondary and further changes will occur in other parts of the brain as a result of the initial event, downstream of the acoustic stimuli, part of a flexible interconnected network. Dysfunction in one node in this network can reroute connections and reframe the network's topography and function. With this type of reorganization under adversity, some compensation for loss of function may emerge, but worsening is a real possibility too. In fact, developmental plasticity did not evolve to reformat a network after a lesion in one or more of its nodes, but rather for learning and growth. When these plasticity mechanisms are summoned to fix a big problem, a pathological event, it should not be expected that they will work well. In fact, more often than not they make matters worse.

Thus, as an infant with a genetic risk for dyslexia grows, eventually the cortex may show cortical reorganization (disorganization?) in its language networks. Imaging and other approaches to demonstrating cortical organization for language will be aberrant, but is that the cause of the reading disorder? Perhaps, it is the immediately proximal cause, but the problem is just as likely not to have started there, but instead at nodes closer to the sensory (in this case acoustic) input. A goal of prevention would be to address the phenomena that are happening earliest in development. And, for as long as it remains out of reach to test and manipulate these early nodes in babies and infants, the use of animal models is crucial for shedding light on those early events.

Nothing has been said in this paper about visual causes of dyslexia. This author believes that visual causes exist, and in fact, he has come in contact with individuals whose dyslexia was visual, without a doubt (see, for instance, Vannuscorps et al., 2021). However, it is likely that visual causes of dyslexia alone are uncommon compared to those of acoustic origin, and it is possible that they affect dyslexic women more often than men (the few cases seen by the author have all been women). Recall also that the limited published neuropathological findings in dyslexic women were different from those of the typical dyslexic man (Humphreys et al., 1990). However, these statements are made in a most tentative manner and are meant mainly to encourage research on sex difference in the causes, brain findings, and cognitive profiles of dyslexia.

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Abbondanza, F., Dale, P. S., Wang, C. A., Hayiou-Thomas, M. E., Toseeb, U., Koomar, T. S., et al. (2022). Non-right handedness is associated with language and reading impairments. *PsyArXiv* [Preprint]. doi: 10.31234/osf.io/74jsd

Adil, S. M., Calabrese, E., Charalambous, L. T., Cook, J. J., Rahimpour, S., Atik, A. F., et al. (2021). A high-resolution interactive atlas of the human brainstem using magnetic resonance imaging. *Neuroimage* 237:118135. doi: 10.1016/j.neuroimage.2021.118135

Adler, W. T., Platt, M. P., Mehlhorn, A. J., Haight, J. L., Currier, T. A., Etchegaray, M. A., et al. (2013). Position of neocortical neurons transfected at different gestational ages with shRNA targeted against candidate dyslexia susceptibility genes. *PLoS One* 8:e65179. doi: 10.1371/journal.pone.0065179

Altarelli, I., Leroy, F., Monzalvo, K., Fluss, J., Billard, C., Dehaene-Lambertz, G., et al. (2014). Planum temporale asymmetry in developmental dyslexia: Revisiting an old question. *Hum. Brain Mapp.* 35, 5717–5735. doi: 10.1002/hbm.22579

Andrews, W., Liapi, A., Plachez, C., Camurri, L., Zhang, J., Mori, S., et al. (2006). Robo1 regulates the development of major axon tracts and interneuron migration in the forebrain. *Development* 133, 2243–2252. doi: 10.1242/dev.02379

Anthoni, H., Sucheston, L. E., Lewis, B. A., Tapia-Páez, I., Fan, X., Zucchelli, M., et al. (2012). The aromatase gene CYP19A1: Several genetic and functional lines of evidence supporting a role in reading, speech and language. *Behav. Genet.* 42, 509–527. doi: 10.1007/s10519-012-9532-3

Anthoni, H., Zucchelli, M., Matsson, H., Müller-Myhsok, B., Fransson, I., Schumacher, J., et al. (2007). A locus on 2p12 containing the co-regulated MRPL19 and C2ORF3 genes is associated with dyslexia. *Hum. Mol. Genet.* 16, 667–677. doi: 10.1093/hmg/ddm009

Anvarian, Z., Mykytyn, K., Mukhopadhyay, S., Pedersen, L. B., and Christensen, S. T. (2019). Cellular signaling by primary cilia in development, organ function and disease. *Nat. Rev. Nephrol.* 15, 199–219.

Barkovich, A. J. (2012). Developmental disorders of the midbrain and hindbrain. *Front. Neuroanat.* 6:7. doi: 10.3389/fnana.2012.00007

Beissner, F., Schumann, A., Brunn, F., Eisenträger, D., and Bär, K. J. (2014). Advances in functional magnetic resonance imaging of the human brainstem. *Neuroimage* 86, 91–98.

Berent, I., Vaknin-Nusbaum, V., Balaban, E., and Galaburda, A. M. (2012). Dyslexia impairs speech recognition but can spare phonological competence. *PLoS One* 7:e44875. doi: 10.1371/journal.pone.0044875

Berent, I., Vaknin-Nusbaum, V., Balaban, E., and Galaburda, A. M. (2013). Phonological generalization in dyslexia: The phonological grammar may not be impaired. *Cogn. Neuropsychol.* 30, 285–310. doi: 10.1017/S0305000909990109

Brosnan, M., Demetre, J., Hamill, S., Robson, K., Shepherd, H., and Cody, G. (2016). Executive function in adults with developmental dyslexia. *Res. Dev. Disabil.* 53, 323–341. doi: 10.1080/02643294.2013.863182

Brown, M. S., Singel, D., Hepburn, S., and Rojas, D. C. (2013). Increased glutamate concentration in the auditory cortex of persons with autism and first-degree relatives: A (1)H-MRS study. *Autism Res.* 6, 1–10. doi: 10.1002/aur.1260

Carrey, N. J., MacMaster, F. P., Gaudet, L., and Schmidt, M. H. (2007). Striatal creatine and glutamate/glutamine in attention-deficit/hyperactivity disorder. *J. Child Adolesc. Psychopharmacol.* 17, 11–17. doi: 10.1089/cap.2006.0008

Centanni, T. M., Booker, A. B., Chen, F., Sloan, A. M., Carraway, R. S., Rennaker, R. L., et al. (2016). Knockdown of dyslexia-gene Dcdc2 interferes with speech sound discrimination in continuous streams. *J. Neurosci.* 36, 4895–4906. doi: 10.1523/JNEUROSCI.4202-15.2016

Centanni, T. M., Chen, F., Booker, A. M., Engineer, C. T., Sloan, A. M., Rennaker, R. L., et al. (2014a). Speech sounds processing deficits and training-induced neural plasticity in rats with dyslexia gene knockdown. *PLoS One* 9:e98439. doi: 10.1371/journal.pone.0098439

Centanni, T. M., Booker, A. B., Sloan, A. M., Chen, F., Maher, B. J., Carraway, R. S., et al. (2014b). Knockdown of the dyslexia-associated gene Kiaa0319 impairs temporal responses to speech stimul in rat primary auditory cortex. *Cereb. Cortex* 24, 1753–1766. doi: 10.1093/cercor/bht028

Chandrasekar, G., Vesterlund, L., Hultenby, K., Tapia-Páez, I., and Kere, J. (2013). The zebrafish orthologue of the dyslexia candidate gene DYX1C1 is essential for cilia growth and function. *PLoS One* 8:e63123. doi: 10.1371/journal.pone.0063123

Chang, B. S., Ly, J., Appignani, B., Bodell, A., Apse, K. A., Ravenscroft, R. S., et al. (2005). Reading impairment in the neuronal migration disorder of periventricular nodular heterotopia. *Neurology* 64, 799–803.

Che, A., Girgenti, M. J., and LoTurco, J. (2014). The dyslexia-associated gene DCDC2 is required for spike-timing precision in mouse neocortex. *Biol. Psychiatry* 76, 387–396. doi: 10.1016/j.biopsych.2013.08.018

Che, A., Truong, D. T., Fitch, R. H., and LoTurco, J. J. (2016). Mutation of the dyslexia-associated gene Dcdc2 enhances glutamatergic synaptic transmission between layer 4 neurons in mouse neocortex. *Cereb. Cortex* 26, 3705–3718. doi: 10.1093/cercor/bhv168

Chen, B., Linke, A., Olson, L., Ibarra, C., Kinnear, M., and Fishman, I. (2021). Resting state functional networks in 1-to-3 typically developing children. *Dev. Cogn. Neurosci.* 51:100991. doi: 10.1016/j.dcn.2021.100991

Clark, M. G., Rosen, G. D., Tallal, P., and Fitch, R. H. (2000a). Impaired processing of complex auditory stimuli in rats with induced cerebrocortical microgyria: An animal model of developmental language disability. *J. Cogn. Neurosci.* 12, 828–839. doi: 10.1162/089892900562435

Clark, M. G., Rosen, G. D., Tallal, P., and Fitch, R. H. (2000b). Impaired two-tone processing of rapid rates in male rates with induced microgyria. *Brain Res.* 871, 94–97. doi: 10.1016/s0006-8993(00)02447-1

Cope, N., Eicher, J. D., Meng, H., Gibson, C. J., Hager, K., Lacadie, C., et al. (2012). Variants in DYX2 locus are associated with altered brain activation in reading-related brain regions in subjects with reading disability. *Neuroimage* 63, 148–156. doi: 10.1016/j.neuroimage.2012.06.037

Cope, N., Harold, D., Hill, G., Moskvina, V., Stevenson, J., Holmans, P., et al. (2005). Strong evidence that KIAA0319 on chromosome 6p is a susceptibility gene for developmental dyslexia. *Am. J. Hum. Genet.* 76, 581–591.

Currier, T. A., Etchegaray, M. A., Haight, J. L., Galaburda, A. M., and Rosen, G. D. (2011). The effects of embryonic knockdown of the candidate dyslexia susceptibility gene homologue Dyx1c1 on the distribution of GABAergic neurons in the cerebral cortex. *Neuroscience* 172, 535–546. doi: 10.1016/j.neuroscience.2010.11.002

Darki, F., Peyrard-Janvid, M., Matsson, H., Kere, J., and Klingberg, T. (2012). Three dyslexia susceptibility genes, DYX1C1, DCDC2, and KIAA0319, affect temporo-parietal white matter structure. *Biol. Psychiatry* 72, 671–676. doi: 10.1016/j.biopsych.2012.05.008

Darki, F., Peyrard-Janvid, M., Matsson, H., Kere, J., and Klingberg, T. (2014). DCDC2 polymorphism is associated with left temporoparietal gray and white matter structures during development. *J. Neurosci.* 34, 14455–14462. doi: 10.1523/JNEUROSCI.1216-14.2014

Dasgupta, A., and Amack, J. D. (2016). Cilia in vertebrate left-right patterning. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 371:20150410.

Denenberg, V. (1981). Hemispheric laterality in animals and the effect of early experience. *Behav. Brain Sci.* 4, 1–21.

Díaz, B., Hintz, F., Kiebel, S. J., and von Kriegstein, K. (2012). Dysfunction of the auditory thalamus in developmental dyslexia. *Proc. Natl. Acad. Sci. U. S. A.* 109, 13841–13846.

Diaz, R., Kronenberg, N. M., Martinelli, A., Liehm, P., Riches, A. C., and Gather, M. C. (2022). KIAA0319 influences cilia length, cell migration and mechanical cell-substrate interaction. *Sci. Rep.* 12:722. doi: 10.1038/s41598-021-04539-3

D'Mello, A. M., and Gabrieli, J. D. E. (2018). Cognitive Neuroscience of Dyslexia. *Lang. Speech. Hear. Serv. Sch.* 49, 798–809.

Dvorak, K., Feit, J., and Juránková, Z. (1978). Experimentally induced focal microgyria and statusverrucosus deformis in rats—pathogenesis and interrelation. Histological and autoradiographical study. *Acta Neuropathol.* 44, 121–129. doi: 10.1007/BF00691477

Eberli, F., Rice, M., and Paracchini, S. (2021). Insights into dyslexia genetics research from the last two decades. *Brain Sci.* 12:27. doi: 10.3390/brainsci12010027

Eicher, J. D., and Gruen, J. R. (2013). Imaging-genetics in dyslexia: Connecting risk genetic variants to brain neuroimaging and ultimately to reading impairments. *Mol. Genet. Metab.* 110, 201–212. doi: 10.1016/j.ymgme.2013.07.001

Elliott, D. (2020). The legacy of the Kennard principle. *J. Undergrad. Neurosci. Educ.* 19, R11–R14.

Fitch, R. H., Brown, C. P., Tallal, P., and Rosen, G. D. (1997). Effects of sex and MK-801 on auditory-processing deficits associated with developmental microgyric lesions in rats. *Behav. Neurosci.* 111, 404–412. doi: 10.1037//0735-7044.111.2.404

Fitch, R. H., Tallal, P., Brown, C. P., Galaburda, A. M., and Rosen, G. D. (1994). Induced microgyria and auditory temporal processing in rats: A model for language impairment? *Cereb. Cortex* 4, 260–270. doi: 10.1093/cercor/4.3.260

Franquinho, F., Nogueira-Rodrigues, J., Duarte, J. M., Esteves, S. S., Carter-Su, C., and Monaco, A. P. (2017). The dyslexia-susceptibility protein KIAA0319 inhibits axon growth through Smad2 signaling. *Cereb. Cortex* 27, 1732–1747. doi: 10.1093/cercor/bhx023

Friede, R. L. (1989). *Developmental Neuropathology*. Berlin: Springer-Verlag.

Galaburda, A. M., and Kemper, T. L. (1979). Cytoarchitectonic abnormalities in developmental dyslexia: A case study. *Ann. Neurol.* 6, 94–100.

Galaburda, A. M., Menard, M. T., and Rosen, G. D. (1994). Evidence for aberrant auditory anatomy in developmental dyslexia. *Proc. Natl. Acad. Sci. U. S. A.* 91, 8010–8013.

Galaburda, A. M., Sherman, G. F., Rosen, G. D., Aboitiz, F., and Geschwind, N. (1985). Developmental dyslexia: Four consecutive patients with cortical anomalies. *Ann. Neurol.* 18, 222–233. doi: 10.1002/ana.410180210

Ganguly, K., Schinder, A. F., Wong, S. T., and Poo, M. (2001). GABA itself promotes the developmental switch of GABAergic responses from excitation to inhibition. *Cell* 105, 521–532.

Geschwind, N., and Levitsky, W. (1968). Human brain: Left-right asymmetries in temporal speech region. *Science* 161, 186–187.

Glick, S. D., and Ross, D. A. (1981). Right-sided population bias and lateralization of activity in normal rats. *Brain Res.* 205, 222–225. doi: 10.1016/0006-8993(81)90737-x

Gonda, Y., Namba, T., and Hanashima, C. (2020). Beyond axon guidance: Roles of slit-robo signaling in neocortical formation. *Front. Cell Dev. Biol.* 8:607415. doi: 10.3389/fcell.2020.607415

Goriounova, N., Heyer, D. B., and Wilbers, R. (2018). Large and fast human pyramidal neurons associate with intelligence. *Elife* 7:41714. doi: 10.7554/eLife.41714

Gottesman, I. I., and Gould, T. D. (2003). The endophenotype concept in psychiatry: Etymology and strategic intentions. *Am. J. Psychiatry* 160, 636–645.

Gottesman, I. I., and Shields, J. (1972). *Schizophrenia and genetics*. New York, NY: Academic Press.

Grati, M., Chakchouk, I., Ma, Q., Bensaid, M., Desmidt, A., Turki, N., et al. (2015). A missense mutation in DCDC2 causes human recessive deafnes DFNB66, likely by interfering with sensory hair cell and supporting cell cilia length regulation. *Hum. Mol. Genet.* 24, 2482–2491. doi: 10.1093/hmg/ddv009

Guemez-Gamboa, A., Coufal, N. G., and Gleeson, J. G. (2014). Primary cilia in the developing and mature brain. *Neuron* 82, 511–521.

Guidi, L. G., Mattley, J., and Martinez-Garay, I. (2017). Knockout mice for dyslexia susceptibility gene homologs KIAA0319 and KIAA0319L have unaffected neuronal migration but display abnormal auditory processing. *Cereb. Cortex* 27, 5831–5845. doi: 10.1093/cercor/bhx269

Guidi, L. G., Velayos-Baeza, A., Martinez-Garay, I., Monaco, A. P., Paracchini, S., Bishop, D. V. M., et al. (2018). The neuronal migration hypothesis of dyslexia: A critical evaluation 30 years on. *Eur. J. Neurosci.* 48, 3212–3233. doi: 10.1111/ejn.14149

Hannula-Jouppi, K., Kaminen-Ahola, N., Taipale, M., Eklund, R., Nopola-Hemmi, J., Kääriäinen, H., et al. (2005). The axon guidance receptor gene ROBO1 is a candidate gene for developmental dyslexia. *PLoS Genet.* 1:e50. doi: 10.1371/journal.pgen.0010050

Hasenpusch-Theil, K., and Theil, T. (2021). The multifaceted roles of primary cilia in the development of the cerebral cortex. *Front. Cell Dev.* 9:630161. doi: 10.3389/fcell.2021.630161

Hopkins, W. D., Staes, N., Mulholland, M. M., Schapiro, S. J., Rosenstein, M., Stimpson, C., et al. (2021). Gray matter variation in the posterior superior temporal gyrus is associated with polymorphisms in the KIAA0319 gene in chipanzees *(Pan troglodytes)*. *eNeuro* doi: 10.1523/ENEURO.0169-21.2021

Hornickel, J., and Kraus, N. (2013). Unstable representation of sound: A biological marker of dyslexia. *J. Neurosci.* 33, 3500–3504. doi: 10.1523/JNEUROSCI.4205-12.2013

Humphreys, P., Kaufmann, W. E., and Galaburda, A. M. (1990). Developmental dyslexia in women: Neuropathological findings in three patients. *Ann. Neurol.* 28, 727–738. doi: 10.1002/ana.410280602

Humphreys, P., Rosen, G. D., Press, D. M., Sherman, G. F., and Galaburda, A. M. (1991). Freezing lesions of the developing rat brain: A model for cerebrocortical microgyria. *J. Neuropathol. Exp. Neurol.* 50, 145–160. doi: 10.1097/00005072-199103000-00006

Jacobs, H. I., Van Boxtel, M. P., Jolles, J., Verhey, F. R., and Uylings, H. B. (2012). Parietal cortex matters in Alzheimer's disease: An overview of structural, functional and metabolic findings. *Neurosci. Biobehav. Rev.* 36, 297–309. doi: 10.1016/j.neubiorev.2011.06.009

Johnston, M. V. (2004). Clinical disorders of brain plasticity. *Brain Dev.* 26, 73–80.

Knecht, S., Dräger, B., Deppe, M., Bobe, L., Lohmann, H., and Flöel, A. (2000). Handedness and hemispheric language dominance in healthy humans. *Brain* 123, 2512–2518.

Kolb, B. (2003). Overview of cortical plasticity and recovery from brain injury. *Phys. Med. Rehabil. Clin. N. Am.* 14, S7–S25.

Krafnick, A. J., and Evans, T. M. (2019). Neurobiological sex differences in developmental dyslexia. *Front. Psychol.* 9:2669. doi: 10.3389/fpsyg.2018.02669

Kuhl, P. K., and Miller, J. D. (1971). Speech perception by the chinchilla. Voiced-voiceless distinction in alveolar plosive consonants. *Science* 190, 69–72. doi: 10.1126/science.1166301

Kumar, V., Umair, Z., Kumar, S., Goutam, R. S., Park, S., and Kim, J. (2021). The regulatory roles of motile cilia in CSF circulation and hydrocephalus. *Fluid Barriers CNS* 18:31. doi: 10.1186/s12987-021-00265-0

Kuzniecky, R., Andermann, F., and Guerrini, R. (1993). Congenital bilateral perisylvian syndrome: Study of 31 patients. The CBPS Multicenter Collaborative Study. *Lancet* 341, 608–612. doi: 10.1016/0140-6736(93)90363-l

Lechanoine, F., Jacquesson, T., Beaujoin, J., Serres, B., Mohammadi, M., Planty-Bonjour, A., et al. (2021). WIKIBrainStem: An online atlas to manually segment the human brainstem at the mesoscopic scale from ultrahigh field MRI. *Neuroimage* 236:118080. doi: 10.1016/j.neuroimage.2021.118080

Lee, J. E., and Gleeson, J. G. (2011). A systems-biology approach to understanding the ciliopathy disorders. *Genome Med.* 3:59.

Li, A. M., Hill, R. A., and Grutzendler, J. (2021). Intravital imaging of neocortical heterotopia reveals aberrant axonal pathfinding and myelination around ectopic neurons. *Cereb. Cortex* 31, 4340–4356.

Loges, N. T., and Omran, H. (2018). "Dynein dysfunction as a cause of primary ciliary dyskinesia and other ciliopathies," in *Dyneins: Structure, biology and disease (second ediction)*, ed. S. M. King (Cambridge, MA: Academic Press), 316–355.

Lu, Y., Sareddy, G. R., Wang, J., Wang, R., Li, Y., Dong, Y., et al. (2019). Neuron-derived estrogen regulates synaptic plasticity and memory. *J. Neurosci.* 39, 2792–2809.

Martinez-Garay, I., Guidi, L. G., Holloway, Z. G., Bailey, M. A., Lyngholm, D., Schneider, T., et al. (2017). Normal radial migration and lamination are maintained in dyslexia-susceptibility candidate gene homolog Kiaa0319 knockout mice. *Brain Struct. Funct.* 222, 1367–1384. doi: 10.1007/s00429-016-1282-1

Mascheretti, S., Riva, V., Feng, B., Trezzi, V., Andreola, C., Giorda, R., et al. (2020). The mediation role of dynamic multisensory processing using molecular genetic data in dyslexia. *Brain Sci.* 10:993. doi: 10.3390/brainsci10120993

Massinen, S., Hokkanen, M. E., Matsson, H., Tammimies, K., Tapia-Páez, I., Dahlström-Heuser, V., et al. (2011). Increased expression of dyslexia candidate gene DCDC2 affects length and signaling of primary cilia in neurons. *PLoS One* 6:e20580. doi: 10.1371/journal.pone.0020580

Matute, C., Alberdi, E., Domercq, M., Sánchez-Gómez, M. V., Pérez-Samartín, A., Rodríguez-Antigüedad, A., et al. (2007). Excitotoxic damage to white matter. *J. Anat.* 210, 693–702.

Meda, S. A., Gelernter, J., Gruen, J. R., Calhoun, V. D., Meng, H., Cope, N. A., et al. (2008). Polymorphism of DCDC2 reveals differences in cortical morphology of healthy individuals – A preliminary voxel based morphometry study. *Brain Imaging Behav.* 2, 21–26. doi: 10.1007/s11682-007-9012-1

Meng, H., Smith, S. D., Hager, K., Held, M., Liu, J., Olson, R. K., et al. (2005). DCDC2 is associated with reading disability and modulates neuronal development in the brain. *Biol. Sci.* 102, 17053–17058. doi: 10.1073/pnas.0508591102

Moon, K., Ma, J., Min, H., Koo, H., Kim, H., Ko, H. W., et al. (2020). Dysregulation of sonic hedgehog signaling causes hearing loss in ciliopathy mouse models. *eLife* 9:56551. doi: 10.7554/eLife.56551

Neef, N. E., Müller, B., Liebig, J., Schaadt, G., Grigutsch, M., Gunter, T. C., et al. (2017). Dyslexia risk gene relates to representation of sound in the auditory brainstem. *Dev. Cogn. Neurosci.* 24, 63–71.

Paracchini, S., Scerri, T., and Monaco, A. P. (2007). The genetic lexicon of dyslexia. *Annu. Rev. Genomics Hum. Genet.* 8, 57–79.

Paracchini, S., Thomas, A., Castro, S., Lai, C., Paramasivam, M., Wang, Y., et al. (2006). The chromosome 6p22 haplotype associated with dyslexia reduces the expression of KIAA0319, a novel gene involved in neuronal migration. *Hum. Mol. Genet.* 15, 1659–1666. doi: 10.1093/hmg/ddl089

Perrin, S. (2014). Preclinical research: Make mouse studies work. *Nature* 507, 423–425.

Peschansky, V. J., Burbridge, T. J., Volz, A. J., Fiondella, C., Wissner-Gross, Z., Galaburda, A. M., et al. (2010). The effect of variation in expression of the candidate dyslexia susceptibility gene homolog Kiaa0319 on neuronal migration and dendritic morphology in the rat. *Cereb. Cortex* 20, 884–897. doi: 10.1093/cercor/bhp154

Pinel, P., Fauchereau, F., Moreno, A., Barbot, A., Lathrop, M., Zelenika, D., et al. (2012). Genetic variants of FOXP2 and KIAA0319/TTRAP/THEM2 locus are associated with altered brain activation in distinct language-related regions. *J. Neurosci.* 32, 817–825. doi: 10.1523/JNEUROSCI.5996-10.2012

Platt, M. P., Adler, W. T., Mehlhorn, A. J., Johnson, G. C., Wright, K. A., Choi, R. T., et al. (2013). Embryonic disruption of the candidate dyslexia susceptibility gene homolog KIAA0319-LIKE results in neuronal migration disorders. *Neuroscience* 248, 585–593. doi: 10.1016/j.neuroscience.2013.06.056

Postema, M. C., Carrion-Castillo, A., Fisher, S. E., Vingerhoets, G., and Francks, C. (2020). The genetics of situs inversus without primary ciliary dyskinesia. *Sci. Rep.* 10:3677.

Pugh, K. R., Frost, S. J., Rothman, D. L., Hoeft, F., Del Tufo, S. N., Mason, G. F., et al. (2014). Glutamate and choline levels predict individual differences in reading ability in emergent readers. *J. Neurosci.* 34, 4082–4089.

Pugh, K. R., Mencl, W. E., Jenner, A. R., Katz, L., Frost, S. J., Lee, J. R., et al. (2000). Functional neuroimaging studies of reading and reading disability (developmental dyslexia). *Ment. Retard. Dev. Disabil. Res. Rev.* 6, 207–213.

Raschle, N., Chang, M., and Gaab, N. (2011). Structural brain alterations associated with dyslexia predate reading onset. *Neuroimage* 57, 742–749. doi: 10.1016/j.neuroimage.2010.09.055

Reiner, O., Coquelle, F. M., Peter, B., Levy, T., Kaplan, A., Sapir, T., et al. (2006). The evolving doublecortin (DCX) superfamily. *BMC Genomics* 7:188. doi: 10.1186/1471-2164-7-188

Richlan, F. (2020). The functional neuroanatomy of developmental dyslexia across languages and writing systems. *Front. Psychol.* 11:155. doi: 10.3389/fpsyg.2020.00155

Romeo, D. M., Venezia, I., Pede, E., and Brogna, C. (2022). Cerebral palsy and sex differences in children. A narrative review of the literature. *J. Neurosci. Res.* [Epub ahead of print]. doi: 10.1002/jnr.25020

Rosen, G. D., Bai, J., Wang, Y., Fiondella, C. G., Threlkeld, S. W., LoTurco, J. J., et al. (2007). Disruption of neuronal migration by targeted RNAi knockdown of Dyx1c1 results in neocortical and hippocampal malformations. *Cereb. Cortex* 17, 2562–2572. doi: 10.1093/cercor/bhl162

Rosen, G. D., Burstein, D., and Galaburda, A. M. (2000). Changes in efferent and afferent connectivity in rats with induced cerebrocortical microgyria. *J. Comp. Neurol.* 418, 423–440.

Rosen, G. D., Jacobs, K. M., and Prince, D. A. (1998). Effects of neonatal freeze lesions on expression of parvalbumin in the rat neocortex. *Cereb. Cortex* 8, 753–761. doi: 10.1093/cercor/8.8.753

Rosen, G. D., Press, D. M., Sherman, G. F., and Galaburda, A. M. (1992). The development of induced cerebrocortical microgyria in the rat. *J. Neuropathol. Exp. Neurol.* 51, 601–611.

Scerri, T., Macpherson, E., and Martinelli, A. (2017). The DCDC2 deletion is not a risk factor for dyslexia. *Transl. Psychiatry* 7:e1182.

Schneider, G. E. (1979). Is it really better to have your brain lesion early? A revision of the "Kennard principle". *Neuropsychologia* 17, 557–583. doi: 10.1016/0028-3932(79)90033-2

Shanks, N., Greek, R., and Greek, J. (2009). Are animal models predictive for humans?. *Philos. Ethics Humanit. Med.* 15:2.

Shapleske, J., Rossell, S. L., Woodruff, P. W., and David, A. S. (1999). The planum temporale; a systematic, quantitative review of it structural, functional and clinical significance. *Brain Res. Rev.* 29, 26–49. doi: 10.1016/s0165-0173(98)00047-2

Smith-Park, J. H., Henry, L. A., Messer, D. J., Edwardsdottir, E., and Zięcik, A. P. (2016). Executive function in adults with developmental dyslexia. *Res. Dev. Disabil.* 53-54, 323–341. doi: 10.1016/j.ridd.2016.03.001

Suzuki, M., and Choi, B. H. (1991). Repair and reconstruction of the cortical plate following closed cryogenic injury to the neonatal rat cerebrum. *Acta Neuropathol.* 82, 93–101. doi: 10.1007/BF00293950

Szalkowski, C. E., Booker, A. B., Truong, D. T., Threlkeld, S. W., Rosen, G. D., and Fitch, R. H. (2013). Knockdown of the candidate dyslexia susceptibility gene homolog Dyx1c1 in rodents: Effects of auditory processing, visual attention, and cortical and thalamic anatomy. *Dev. Neurosci.* 35, 50–68. doi: 10.1159/000348431

Szalkowski, C. E., Hinman, J. R., Threlkeld, S. W., Wang, Y., LePack, A., Rosen, G. D., et al. (2011). Persistent spatial working memory deficits in rats following in utero RNAi of Dyx1c1. *Genes Brain Behav.* 10, 244–252. doi: 10.1111/j.1601-183X.2010.00662.x

Taipale, M., Kaminen, N., Nopola-Hemmi, J., Haltia, T., Myllyluoma, B., Lyytinen, H., et al. (2003). A candidate gene for developmental dyslexia encodes a nuclear tetratricopeptide repeat domain protein dynamically regulated in the brain. *Proc. Natl. Acad. Sci. U. S. A.* 100, 11553–11558.

Tarkar, A., Loges, N. T., Slagle, C. E., Francis, R., Dougherty, G. W., Tamayo, J. V., et al. (2013). DYX1C1 is required for axonemal dynein assembly and ciliary motility. *Nat. Genet.* 45, 995–1003.

Threlkeld, S. W., McClure, M. M., Bai, J., Wang, Y., LoTurco, J. J., Rosen, G. D., et al. (2007). Developmental disruptions and behavioral impairments in rats following in utero RNAi of Dyx1c1. *Brain Res. Bull.* 71, 508–514.

Tracey, I., and Iannetti, D. (2006). Chapter 6 brainstem functional imaging in humans. *Suppl. Clin. Neurophysiol.* 58, 52–67.

Trulioff, A., Ermakov, A., and Malashichev, Y. (2017). Primary cilia as a possible link between left-right asymmetry and neurodevelopmental diseases. *Genes* 8, 48–72. doi: 10.3390/genes8020048

Tschentscher, N., Ruisinger, A., Blank, H., Díaz, B., and von Kriegstein, K. (2019). Reduced structural connectivity between left auditory thalamus and the motion-sensitive planum temporale in developmental dyslexia. *J. Neurosci.* 39, 1720–1732. doi: 10.1523/JNEUROSCI.1435-18.2018

Vannuscorps, G., Galaburda, A. M., and Caramazza, A. (2021). Shape-centered representations of bounded regions of space mediate the perception of objects. *Cogn. Neuropsychol.* 39, 1–50. doi: 10.1080/02643294.2021.1960495

Venkatesh, S. K., Siddaiah, A., Padakannaya, P., and Ramachandra, N. B. (2013). Lack of association between genetic polymorphisms in ROBO1, MRPL19/C2ORF3 and THEM2 with developmental dyslexia. *Gene* 529, 215–219. doi: 10.1016/j.gene.2013.08.017

Wang, R., Chen, C., Hara, E., Rivas, M. V., Roulhac, P. L., Howard, J. T., et al. (2015). Convergent differential regulation of SLIT-ROBO axon guidance genes in the brains of vocal learners. *J. Comp. Neurol.* 523, 892–906. doi: 10.1002/cne.23719

Wang, Y., Yin, X., Rosen, G., Gabel, L., Guadiana, S. M., Sarkisian, M. R., et al. (2011). Dcdc2 knockout mice display exacerbated developmental disruptions following knockdown of doublecortin. *Neuroscience* 190, 398–408. doi: 10.1016/j.neuroscience.2011.06.010

White-Schwoch, T., Woodruff Carr, K., Thompson, E. C., Anderson, S., Nicol, T., Bradlow, A. R., et al. (2015). Auditory processing in noise: A preschool biomarker for literacy. *PLoS Biol.* 13:e102196. doi: 10.1371/journal.pbio.1002196

frontiers | Frontiers in Human Neuroscience

TYPE Original Research
PUBLISHED 15 November 2022
DOI 10.3389/fnhum.2022.922552

Language input in late infancy scaffolds emergent literacy skills and predicts reading related white matter development

Yael Weiss[1,2]*†, Elizabeth Huber[1,2]†, Naja Ferjan Ramírez[1,3], Neva M. Corrigan[1,2], Vasily L. Yarnykh[4] and Patricia K. Kuhl[1,2]

[1]Institute for Learning and Brain Sciences, University of Washington, Seattle, WA, United States, [2]Department of Speech and Hearing Sciences, University of Washington, Seattle, WA, United States, [3]Department of Linguistics, University of Washington, Seattle, WA, United States, [4]Department of Radiology, University of Washington, Seattle, WA, United States

Longitudinal studies provide the unique opportunity to test whether early language provides a scaffolding for the acquisition of the ability to read. This study tests the hypothesis that parental language input during the first 2 years of life predicts emergent literacy skills at 5 years of age, and that white matter development observed early in the 3rd year (at 26 months) may help to account for these effects. We collected naturalistic recordings of parent and child language at 6, 10, 14, 18, and 24 months using the Language ENvironment Analysis system (LENA) in a group of typically developing infants. We then examined the relationship between language measures during infancy and follow-up measures of reading related skills at age 5 years, in the same group of participants (N = 53). A subset of these children also completed diffusion and quantitative MRI scans at age 2 years (N = 20). Within this subgroup, diffusion tractography was used to identify white matter pathways that are considered critical to language and reading development, namely, the arcuate fasciculus (AF), superior and inferior longitudinal fasciculi, and inferior occipital-frontal fasciculus. Quantitative macromolecular proton fraction (MPF) mapping was used to characterize myelin density within these separately defined regions of interest. The longitudinal data were then used to test correlations between early language input and output, white matter measures at age 2 years, and pre-literacy skills at age 5 years. Parental language input, child speech output, and parent-child conversational turns correlated with pre-literacy skills, as well as myelin density estimates within the left arcuate and superior longitudinal fasciculus. Mediation analyses indicated that the left AF accounted for longitudinal relationships between infant home language measures and 5-year letter identification and letter-sound knowledge, suggesting that the left AF myelination at 2 years may serve as a mechanism by which early language experience supports emergent literacy.

KEYWORDS

reading development, language development, brain imaging, parental language input, LENA, white matter myelination, longitudinal, conversational turns

OPEN ACCESS

EDITED BY
Charles A. Perfetti,
University of Pittsburgh, United States

REVIEWED BY
Chun-Hsien Hsu,
National Central University, Taiwan
John S. Hutton,
Cincinnati Children's Hospital Medical Center, United States

*CORRESPONDENCE
Yael Weiss
ylweiss@uw.edu

†These authors share first authorship

SPECIALTY SECTION
This article was submitted to Brain Imaging and Stimulation, a section of the journal Frontiers in Human Neuroscience

RECEIVED 18 April 2022
ACCEPTED 26 October 2022
PUBLISHED 15 November 2022

CITATION
Weiss Y, Huber E, Ferjan Ramírez N, Corrigan NM, Yarnykh VL and Kuhl PK (2022) Language input in late infancy scaffolds emergent literacy skills and predicts reading related white matter development. *Front. Hum. Neurosci.* 16:922552. doi: 10.3389/fnhum.2022.922552

Introduction

Examining whether early language skills predict later emergent literacy skills in preschoolers can shed light on the relationship between language and literacy development and may contribute to developing effective early instruction and interventions and early identification of children at risk of developing reading difficulties. Early language measures coupled with neuroimaging measures in the developing brain can further illuminate the interplay between language input, early language skills, and the emerging neural circuitry for reading. The current study examined the relationship between early language skills and parental language input in infancy (6–24 months), white matter structure at age 2 years, and emergent literacy skills in 5-year-old preschoolers. Additionally, white matter structure at 2 years was tested as a potential mediator of the relationship between infant and preschool behavioral measures.

Predictors of emergent literacy skills

Phonological awareness and letter-sound knowledge have consistently been found as the two best predictors of reading acquisition during the first 2 years in school (Share, 2004; NRP, 2020). These two skills are necessary to develop decoding, which is the ability to activate speech-based information from a string of printed letters (Share, 1995). However, according to the simple view of reading (SVR) framework (Hoover and Gough, 1990), skilled reading reflects adequate reading comprehension. That is, according to SVR, the ability to develop good reading skills depends on the child's language comprehension skills and their ability to decode written words. Consequently, the SVR framework hypothesizes that decoding and language comprehension are two distinct components, and that reading comprehension is the product of these two components. The SVR framework further predicts that reading difficulties can result from inadequate decoding skills, inadequate language comprehension skills, or both (Hoover and Gough, 1990).

Longitudinal and cross-sectional studies have supported the SVR framework by demonstrating that concurrent language comprehension and decoding skills strongly predict reading comprehension skills at school age. These studies also extend the SVR framework by showing that language comprehension and decoding skills in preschool, before the onset of formal literacy instruction, predict reading comprehension skills at school age even several years later (Kendeou et al., 2009; Hjetland et al., 2017, 2019; Chiu, 2018; Dickinson et al., 2019). However, while Kendeou et al. (2009) further support the SVR framework by indicating that language and decoding skills are two distinct components, other studies challenge it by demonstrating strong correlations between decoding and language comprehension in preschoolers, indicating that these are not two distinct components and instead depend on one another (Chiu, 2018; Dickinson et al., 2019; Hjetland et al., 2019). These inconsistencies are suggested to be due to the use of different measures and analytical approaches among different studies (Hjetland et al., 2017). For example, Kendeou et al. (2009) used factor analysis and included receptive vocabulary under the decoding skills component, while other studies used structural equation modeling (SEM) or path and mediation analysis and included vocabulary measures under the language comprehension component (Chiu, 2018; Dickinson et al., 2019; Hjetland et al., 2019). Furthermore, while most of these studies include only standardized measures, others include natural measures from language samples as well (for a comprehensive review see Hjetland et al., 2017). These inconsistencies highlight the need to examine the specific relations between spoken language and decoding skills prior to the onset of literacy instruction.

Language predictors of emergent decoding skills

Despite the clear contribution of language skills to reading comprehension in school-age children, it is not yet clear how early language skills relate to emergent literacy skills. Different studies show inconsistencies related to the specific nature of these relations and suggest that there are additional contributing factors. One comprehensive study explicitly addressed the question regarding the relations between early language skills and emergent and later literacy skills using path analysis (NICHD Early Child Care Research Network, 2005). This study indicates that comprehensive language skills at both 36 and 54 months were directly related to the standardized measures of both phonological awareness and letter and word identification at 54 months. These results indicate both concurrent and longitudinal relations between language and emergent literacy skills in preschoolers.

Other studies have mostly focused on the relations between early vocabulary and emergent literacy skills. For example, Deniz Can et al. (2013) showed that expressive vocabulary (measured by a parental questionnaire) at the age of 2 years predicted emergent literacy skills, including standardized measures of letter-sound correspondence, word recognition, and decoding in kindergarten, but not phonological awareness or letter naming. However, these emergent literacy skills were correlated with concurrent language skills, including standardized measures of vocabulary, syntax, and semantics. Another longitudinal study followed children from 19 months to 16 years (Suggate et al., 2018). They found that vocabulary (measured by a parental questionnaire) at 19 months correlated significantly with emergent literacy skills including standardized measures of letter and word recognition and concept about print, 2–3 years later. In addition, they found that maternal receptive vocabulary (measured by a standardized test) when children were 19-months-old, was related to both early vocabulary and emergent literacy skills in the children. Finally, Silvén et al. (2002) found that both infant's expressive and receptive language skills and mother's interactional sensitivity (both measured with an observational approach) at 12 and 24 months predicted standardized measures of emergent phonological awareness skills at 36 and 48 months (Silvén et al., 2002). These two studies highlight the importance of parental input and the home environment in infancy to the development of emergent literacy skills.

There is evidence of long-term predictive associations between early language skills on later reading skills. Flax et al. (2009) found that expressive and receptive language skills (measured with both a standardized test and a parental questionnaire) in the 2nd and 3rd years of life predicted standardized measures of phonological awareness at age 5-years and word identification and decoding at age 7-years. Psyridou et al. (2018) found that receptive and expressive vocabulary (measured by a parental questionnaire) at 24 and 30 months predicted reading comprehension in 2nd, 3rd, 8th, and 9th grades. Duff et al. (2015) found that receptive and expressive vocabulary (measured by a parental questionnaire) in the 2nd year of life predicted different skills at the age of 4–9 years.

The home language environment

Parental language input is a key component of the early learning environment, and child language development is associated with input quantity, as well as several dimensions of input quality [for a recent review, see Rowe and Snow (2020); Rowe and Weisleder (2020)]. For example, a greater quantity of child-directed speech input has been linked to more advanced expressive vocabulary (Huttenlocher et al., 1991; Hart and Risley, 1995; Hurtado et al., 2008; Shneidman et al., 2013) and stronger lexical processing skills in the 2nd year of life (Weisleder and Fernald, 2013). Linguistic quality of parental speech, including lexical and syntactic diversity, complexity, and narrative content (Huttenlocher et al., 1991, 2002, 2010; Tabors et al., 2001; Pan et al., 2005; Rowe, 2012; Song et al., 2014; Hsu et al., 2017; Uccelli et al., 2019; Leech, 2021) has also been shown to correlate with child language outcomes.

Parental speaking style is another important element of input quality. Adults tend to adopt a style known as "parentese" when they address infants and young children (Fernald, 1985; Fernald and Kuhl, 1987; Grieser and Kuhl, 1988). Parentese is characterized by a slower pace and higher, more variable pitch (Fernald, 1985; Fernald and Kuhl, 1987). Parents have been found to use parentese during activities like book reading (Burnham et al., 2015) as well as during spontaneous child-directed speech. Benefits associated with parentese have been reported for several aspects of language acquisition (Cristià, 2013), including phonemic discrimination (Liu et al., 2003), and vocabulary development (Hartman et al., 2017). For example, Hartman et al. (2017) showed a specific relationship between vowel clarity in maternal parentese at 16 months and children's receptive and expressive vocabulary size at age 2 years (Hartman et al., 2017). In the laboratory, parentese has been shown to facilitate speech segmentation (Thiessen et al., 2005), as well as word recognition and learning in infancy (Singh et al., 2009; Ma et al., 2011). Exposure to parentese in the home environment has been found to correlate with child language output, including speech-like "babbling" in infancy (Ramírez-Esparza et al., 2014; Ferjan Ramírez et al., 2019).

Social interaction is also thought to play an important role in language acquisition, and is considered another important dimension of input quality (Rowe and Snow, 2020; Rowe and Weisleder, 2020). Studies that manipulate social-interactional variables, through intervention or experimental methods, suggest that interaction with responsive adults directly advances infant and child language skills (Kuhl et al., 2003; Goldstein and Schwade, 2008; Ferjan Ramírez et al., 2019, 2020). In correlational studies, conversational turn counts have been associated with immediate as well as longer-term linguistic and cognitive outcomes (Gilkerson et al., 2017, 2018). Moreover, there is evidence that the quantity of parent-child conversational turns correlates not only with child linguistic and cognitive performance but also with measures of brain function and structure (at 4–6 years of age) (Romeo et al., 2018a,b, 2021). In particular, conversational turns counts have been found to correlate with the organization of dorsal white matter pathways associated with expressive and receptive language skills in both children and adults (Romeo et al., 2018b).

The early language environment has also been linked to specific components of child literacy skills. Parent-child book reading interactions at 1–2.5 years have been shown to predict receptive vocabulary, reading comprehension, and internal motivation to read, but not decoding, external motivation to read, or math skills at elementary school (2nd–4th grade) (Demir-Lira et al., 2019). Furthermore, studies that measured the quality of the learning environment, including literacy activities, quality of maternal engagement, and availability of learning materials during the 2nd and 3rd years of life, found that Preschool receptive vocabulary and letter-word identification skills were associated with the quality of the learning environment in the 2nd and 3rd years, as well as the preschool learning environment (Rodriguez and Tamis-Lemonda, 2011; Tamis-LeMonda et al., 2019).

Brain networks supporting language and early literacy development

Brain imaging studies have shown that reading involves two main brain systems in the left hemisphere. The dorsal sub-lexical pathways maps between print and sounds of spoken language and include temporoparietal and frontal regions and the white matter pathways that connect them, namely the arcuate fasciculus (AF) and superior longitudinal fasciculus (SLF). This system is more involved in the beginning stages of reading development. The dorsal pathways are also associated with speech production and auditory-motor integration (Hickok and Poeppel, 2007; Hickok, 2012; Skeide and Friederici, 2016), as well as phonological awareness, vocabulary development, and syntactic processing of speech (Lebel and Beaulieu, 2009; Saygin et al., 2013; Skeide et al., 2016; Su et al., 2018; Reynolds et al., 2019). The ventral lexical system maps between print and meaning and includes occipital and occipitotemporal regions and the inferior fronto-occipital and inferior longitudinal fasciculus (IFOF and ILF), which connect them. This system becomes more automatic with reading experience (Wandell and Yeatman, 2013; Ozernov-Palchik and Gabrieli, 2018; Church et al., 2021).

Longitudinal studies measuring brain function in children from the pre-reading stage throughout reading development have shown that activation in the left posterior superior temporal gyrus and functional connectivity between the left dorsal and ventral pathways, which is related to phonological processing in pre-readers, predict reading skills 1–3 years later at school-age (Yu et al., 2018; Wang et al., 2020; Yamasaki et al., 2021). Similarly, longitudinal studies that measured brain structure in children from the pre-reading stage throughout reading development consistently demonstrate that throughout development the left AF relates to phonological awareness skills that are essential for reading acquisition (Lebel and Beaulieu, 2009; Yeatman et al., 2011; Saygin et al., 2013; Van Der Auwera et al., 2021). Other studies have demonstrated correlations between left AF and cross-modal audio-visual processing in school-age children, and that the direct segment of the AF specifically predicts later reading skills (Gullick and Booth, 2014, 2015). Altogether, these longitudinal studies indicate a strong correlation between the left dorsal white matter and function and reading-related language skills, even before the onset of reading instruction.

Studies with infants (3–12 months) and young children (1–5 years) suggest that white matter development coincides with the emergence of language-related skills. For example, estimates of left AF and SLF fractional anisotropy (FA) obtained shortly after birth have been shown to correlate with receptive and expressive language skills at the age of 2 years (Salvan et al., 2017; Girault et al., 2019a; Sket et al., 2019). Furthermore, magnetic resonance imaging (MRI) myelin water fraction estimates have been shown to correlate with early linguistic and cognitive skills, through 5 years of age (O'Muircheartaigh et al., 2014; Deoni et al., 2016). Similarly, changes in FA from 6 to 24 months have also been found to predict expressive language skills at 24 months (Swanson et al., 2017), suggesting that white matter development during this period influences subsequent language skills. However, it is yet unknown whether the brain structure earlier in life is related to and mediates the correlations between early language, environmental factors, and later reading outcomes.

The current study

In the current study, we examined how infants' spoken language environments and early language skills relate to emergent literacy skills at the age of 5 years. Based the research described above, we hypothesized that parental input and parent-child interactions in the first 2 years of life would predict later pre-reading skills, and that this relationship might be mediated by structural development of the white matter, specifically within pathways related to expressive language skills. Longitudinal data included: (1) Measures of infants' spoken language, parental input, and interactions with caregivers from 6- to 24-months of age, manually coded from home language recordings; (2) Diffusion and quantitative MRI at 26 months; (3) Emergent literacy and related skills at 5-years of age. Correlation and regression analyses were used to assess the relationship between early spoken language skills and parental input, decoding skills prior to the onset of literacy instruction, and the brain mechanisms that support these relations.

Materials and methods

Participants

Seventy-nine families from the Seattle metro area were recruited through the University of Washington (UW) Communications Studies Participants registry with UW Human Subjects Approval that provides subject contact information directly to researchers. These families had previously participated in intervention studies at the Institute for Learning and Brain Sciences (I-LABS) (Ferjan Ramírez et al., 2019, 2020, 2021) and agreed to be re-contacted for future research on their consent forms. All experimental procedures were approved by the UW Institutional Review Board, and all participating families gave informed consent and were compensated monetarily for their time and effort. All families completed audio recordings of their children and environment at five time points in infancy (when infants were 6, 10, 14, 18, and 24 months old). The recordings were employed with the widely used Language ENvironment Analysis system (LENA^{TM} Pro Version 3.4.0, LENA, 2015), which provides audio recordings and measures of different components in children's natural environments (For a more detailed description of the data collection and analysis of the LENA recordings at infancy, see Ferjan Ramírez et al., 2019, 2020, 2021).

As noted above, some families (N = 38 total, and N = 16 in the MRI group) participated in a parental language intervention from child ages 6–18 months. The intervention increased the quantity of parentese speech and parent-child conversational turns observed in home language recordings (Ferjan Ramírez et al., 2019, 2020, 2021), as well as expressive language outcomes from 14 to 30 months of age (Ferjan Ramírez et al., 2019, 2020). The goal of the current study was to evaluate the effects of early language experience, rather than the effects of this specific intervention, which have been described previously. Therefore, data were analyzed across all participants, regardless of their participation in the intervention program. It should be noted, however, that the current sample likely has greater variability and higher rates of parentese speech and parent-child interaction than would be expected in a comparable, non-intervention sample.

All families who participated in the study when their children were infants were invited to return for a follow-up MRI session when children were approximately 26 months old, as well as a follow-up study of pre-reading skills at the age of 5 years. All families who agreed to have their child participate in a follow-up study at the age of 5 years (N = 70) completed an initial phone screening interview to determine whether their children met the following criteria: (1) Pre-K child between the age of 5 years and 5 years and 4 months; (2) Native English is primary in the home (multi-lingual families were included if English was spoken >65% of the time in the home, based on parental report during pre-enrollment screening); (3) Children had no clinically diagnosed congenital, neurological or other physical abnormality. Exclusion criteria included: (1) Any brain injury and medications that impact cognition; (2) Intellectual disability, Autism Spectrum Disorder, mood disorders, and other disorders that impact cognition; (3) Significant and permanent hearing impairments. After the initial screening process, 53 eligible participants were invited to take part in the follow-up study at I-LABS when they were 5-years-old (Weiss et al., 2022). Among the 53 participants who participated in the follow-up study at the age of 5 years, a smaller group also completed brain imaging data collection at 26 months (N = 20, 12 females, M = 27.34 months, SD = 0.73).

Procedure

Measures at 5 years of age

All 53 eligible participants' families completed an online parental questionnaire that included questions regarding children's health and development history, language learning history, family history of dyslexia and reading difficulties, parental education, and socio-economic status (SES). Participants' reading and related skills at the age of 5-years were measured with the following standardized and non-standardized tests that were adapted to online administration by uploading the stimuli to PowerPoint presentations and presented to the participants during online sessions via Zoom (Weiss et al., 2022). The online administration of these tests was a response to the COVID-19 pandemic, when it was not possible to meet with participants in person. Online tests were only used to assess participant's relative performance level, and not as the basis for any kind of formal diagnosis. All participants went through the same procedures, which are also reported in a previous publication (Weiss et al., 2022).

Letter knowledge test

This test is designed to measure Alphabet knowledge and letter sounds. Participants were shown isolated letters on the screen and instructed to name the letters and their corresponding sounds. This test was administered separately for lowercase and uppercase letters. This test resulted in four measures: Uppercase letter names (ULN), uppercase letter sounds (ULS), lowercase letter names (LLN), and lowercase letter sounds (LLS). All 26 letters were presented in random order.

Letter sound scoring was carried out by trained research assistants, under the supervision of the first author. Scorers only accepted isolated pronunciation (not adding any vowel), short vowels, and hard "G," "C," and "X" as correct responses. Video recordings of each session were first scored by the research assistant who administered the session. Two other research team members then provided independent scoring for the videos. Inter-rater inconsistencies were discussed in weekly meetings and resolved as a group.

Woodcock reading mastery tests-third edition

This standardized test (Woodcock, 2011) is designed to assess reading skills in children and adults. We administered the Phonological Awareness (PA) sub-test. Different versions (forms A and B) were counterbalanced between participants.

Expressive vocabulary test-third edition

This standardized test (Williams, 2018) is designed to assess expressive vocabulary test (EVT) and word retrieval based on words in Standard American English in children and adults. Different versions (forms A and B) were counterbalanced between participants.

Language environment and child language measures from 6 to 24 months of age

The Language ENvironment Analysis System (LENA^{TM} Pro Version 3.4.0, LENA, 2015) was used to collect naturalistic first-person recordings from all families over two weekend days when children were 6-, 10-, 14-, 18-, and 24-months-old. Each recording was first pre-processed with the LENA Advanced Data Extractor Tool (ADEX). Custom software was used to identify 50 unique 30-s intervals per day containing the highest daily adult word count (AWC), spaced at least 3 min apart. As described previously (Ferjan Ramírez et al., 2019, 2020), this step identifies intervals with enough language data for analysis and eliminates uninformative periods (e.g., nap times). A total of 100 30-s intervals (50 intervals per day) were obtained for each participant at each age.

Measures of parental language input and child output were then manually coded from the LENA recordings by trained research assistants, following procedures outlined previously (Ramírez-Esparza et al., 2014, 2017a,b; Ferjan Ramírez et al., 2019, 2020, 2021). Coders tabulated the number of intervals containing parentese speech (PT) and/or child speech or speech-like vocalizations (CS), and the total number of parent-child conversational turns (CT) present within each 30-s interval. Ten individuals performed coding for each language variable, and inter-coder reliability was assessed using methods developed by Ramírez-Esparza et al. (2014). Intraclass correlation coefficients (ICC) indicated a high degree of inter-coder agreement (see also Shrout and Fleiss, 1979; Ramírez-Esparza et al., 2017a). ICC estimates for PT, CS, and CT were 0.95, 0.93, and 0.96, respectively.

Although the LENA software package can be used to obtain automated counts of adult words, child vocalizations, and parent-child conversational turns, recent validation studies have shown that these automated estimates are susceptible to error and bias, especially for conversational turn counts in the age range studied here (6–24 months) (Bulgarelli and Bergelson, 2020; Cristià et al., 2020, 2021; Ferjan Ramírez et al., 2021), due to factors like accidental contiguity between speakers (e.g., parent speaking on the phone while the child is babbling to herself, nearby), sibling speech, and noise in the recordings. We, therefore, focused our analysis on manually coded variables. Exact variable definitions, provided in **Table 1**, were based on criteria previously established in the literature (Ramírez-Esparza et al., 2014, 2017a,b; Ferjan Ramírez et al., 2019, 2020, 2021).

Magnetic resonance imaging acquisition at the age 2 years

All data were acquired using a 3.0 T Philips Ingenia MRI system with a 32-channel head coil while children were in natural sleep. High resolution T1-weighted images were acquired using a multi-echo MPRAGE sequence with FOV = 230 × 230 × 180, acquisition voxel size 1.0 mm × 1.0 mm × 1.0 mm, reconstructed voxel size 0.5 mm × 0.5 mm × 0.5 mm, TR/TI/TE1/TE2 = 13/1200/3.7/9.7 ms, shot interval 2,250 ms, and flip angle (FA) = 8°. T1-weighted images were used as a common reference space for later anatomically guided analysis of macromolecular proton fraction (MPF) maps and of diffusion-weighted images, as described below.

For MPF mapping, a fast 3D protocol was implemented according to the single-point synthetic reference method (Yarnykh, 2012, 2016), which included three spoiled gradient-echo sequences with magnetization transfer (MT) (TR = 31 ms, FA = 8°), proton-density (TR = 21 ms, FA = 4°), and T1 (TR = 21 ms, FA = 25°) contrast weightings. Off-resonance saturation in the MT-weighted sequence was applied at the offset frequency 4 kHz with effective FA = 430° and pulse duration 7 ms. All images were obtained in the sagittal plane with dual-echo readout (TE1/TE2 = 4.9 ms/10.0 ms), FOV = 240 × 240 × 200 mm^3, and actual voxel size of 1.25 mm × 1.25 mm × 1.24 mm interpolated to 0.625 mm × 0.625 mm × 0.620 mm. Additionally, actual flip-angle imaging (AFI) B1 maps (Yarnykh, 2007) (TR1/TR2/TE = 60/240/4.8 ms, FA = 60°, voxel size 2.5 mm × 5.0 mm × 5.0 mm) were acquired in the same geometry and reconstructed with 0.625 mm × 0.625 mm × 0.620 mm voxel size.

Diffusion-weighted data were acquired using a single-shot DWI-EPI sequence with FOV = 230 × 230 × 146, acquisition voxel size 1.8 mm × 1.8 mm × 1.9 mm, reconstructed voxel size 1.4 mm × 1.4 mm × 1.9 mm, TR/TE = 11,926/97 ms, FA = 90°. Each diffusion scan included 6 non-diffusion-weighted (b = 0) volumes and 64 diffusion-weighted volumes acquired at either a b-value of 2,000 s/mm^2 (52 non-collinear gradient directions) or a b-value of 800 s/mm^2 (12 additional non-collinear gradient directions). An additional set of 6 non-diffusion-weighted volumes were acquired using the same parameters with a reversed phase encoding direction (posterior–anterior), for use in correcting EPI distortions (Andersson et al., 2003), as described below.

Macromolecular proton fraction mapping

Macromolecular proton fraction maps were reconstructed according to a single-point synthetic reference algorithm (Yarnykh, 2016) with correction of B1 field non-uniformity using custom-written C-language software with previously determined constraints for the non-adjustable two-pool model parameters (Yarnykh, 2012). Software for reconstruction of MPF maps is available at https://www.macromolecularmri.org. Correction of B0 field inhomogeneity was not applied because of a negligible effect of B0-related errors on MPF measurements (Yarnykh et al., 2020). Prior to map reconstruction, individual echo images in each data set were averaged to increase SNR (Helms and Dechent, 2009). Rigid-body registration of the component image volumes was performed using the FLIRT toolbox of the FSL software package (Smith, 2002). Resulting MPF maps were then aligned to each subject's own T1 weighted anatomical image using rigid body registration.

Macromolecular proton fraction is sensitive to changes in myelin content in both gray and white matter (Corrigan et al., 2021), and histological validation studies support a linear relationship between MPF and relative myelin density (Underhill et al., 2011).

TABLE 1 Parent and child language variables measured from 6 to 24 months.

Variable name	Variable definition
Parentese (PT)	Total intervals in which mother, father, or another adult spoke to the infant using parentese speech (high pitch, slow tempo, and exaggerated contours), either alone or in the presence of other adult voices
Child speech and speech-like vocalizations (CS)	Total intervals in which children either repeated or independently produced one or more of the following: fully resonant vowels, consonant–vowel syllables, syllable strings, speech utterances intermixed with non-speech, word-like strings, single words, or word strings
Conversational turns (CT)	Total number of adult utterances directed to child followed within 5 s by a child utterance directed to adult, or vice versa; counted in discrete pairs (child-to-parent = 1 turn, parent-to-child-to-parent = 1 turn, child-to-parent-to-child-to-parent = 2 turns)

Variable definitions used in manual coding of the Language Environment Analysis System (LENA) recordings at each age.

Diffusion magnetic resonance imaging analysis

Diffusion data pre-processing was carried out using the FSL tools (version 6.0.0) for motion and eddy current correction (FSL eddy, Andersson and Sotiropoulos, 2016) and brain extraction (FSL BET, Smith, 2002). Diffusion-weighted volumes first were aligned to an average of the non-diffusion weighted volumes in each scan using rigid body transformation [SPM version 12 (Ashburner and Friston, 1997)]. Volumes were then aligned to the subject's own T1 weighted anatomical image, again using rigid body registration. Diffusion gradients were adjusted to account for rotation applied during motion correction and registration (Leemans and Jones, 2009).

Whole brain tractography was carried out using the MRtrix software package (version 3.0) with the iFOD2 algorithm (Tournier et al., 2010). The resulting whole brain fiber estimates were then entered into the Automated Fiber Quantification software package (AFQ, Yeatman et al., 2012; software available at https://github.com/yeatmanlab/AFQ). Specifically, initial segmentations were made using way-point ROIs in subject native space. The Mori et al. (2006) atlas was then used to refine these segmentations by removing streamlines with 0–3% likelihood of overlapping the desired anatomy.

Tract specific macromolecular proton fraction profiles

Macromolecular proton fraction values were extracted within each fiber tract and summarized at the tract core as a weighted-mean across fiber nodes, which penalizes locations farthest from the spatial center of each tract and thereby minimizes the influence of minor variation in the exact outer boundary of each tract (Yeatman et al., 2012).

Analysis along individual tract locations was carried out by first sampling along 100 evenly spaced locations in native subject space (Yeatman et al., 2012) and then averaging across groups of five nodes to obtain 20 summary locations per subject and tract. Output from AFQ was transformed such that all tract profiles were oriented with nodes increasing (from 0 to 100) right to left, posterior to anterior, and inferior to superior, to simplify presentation of tract profiles. Results were corrected for multiple comparisons along each tract using a permutation-based approach (Nichols and Holmes, 2002), which accounts for spatial similarity within individual white matter tracts (see also Yeatman et al., 2012).

Two participants completed the MPF mapping scans, but not the dMRI scan. For these individuals, MPF profiles were defined as follows: MPF maps were first co-registered to a standard-space template brain, and probabilistic white matter labels were used to define initial candidate tracts (Mori et al., 2006). Tract locations were then visually confirmed in subject native space, relative to the same waypoint ROIs used above, in AFQ (Yeatman et al., 2012). The core of each tract was defined in 3D coordinates using the MATLAB Image Processing Toolbox (using the bwmorph3 and bwconncomp functions), and linearly sampled between termination points at the gray/white matter boundary. All analyses were replicated with and without these participants, to ensure that variation in the definition of fiber bundle "core" values did not change the results.

Data analysis

Behavioral data analysis

First, we wanted to examine whether early parental input and infants' output measures of natural language environment, as recorded with the LENA system and manually coded in infancy are related to later emergent literacy and related skills at the age of 5 years. For the total sample of 53 participants, we calculated the simple bivariate correlations between their early LENA measures at the age of 6, 10, 14, 18, and 24 months, and their emergent literacy and related skills at the age of 5-years.

Second, we examined the correlations between the early LENA measures and 5-years measures for the smaller sub-group of participants that had brain imaging data at 26-months of age ($N = 20$).

In both samples, the average parental education in years was roughly equivalent to a 4-year college degree, with a wide range extending from elementary to postgraduate level degree completion. For both samples, the range of income-to-need ratio (defined as a family's total annual income divided by its corresponding poverty threshold) included families at or below the federal poverty line (ratio < 1) as well as families ranging

TABLE 2 Demographic information of the total and smaller samples.

		All participants (N = 53 total)	MRI participants (N = 20 total)
Gender	Identify as boys	25 (47%)	7 (35%)
	Identify as girls	26 (49%)	11 (55%)
	Other/Prefer not to answer	2 (4%)	2 (10%)
Age at first session	Mean age in months (SD)	60.82 (0.88)	60.54 (0.74)
Socio-economic status	Average years of parental education (SD)	17.43 (2.07)	16.83 (1.64)
	Income-to-need ratio (SD)	6.46 (3.97)	4.76 (3.77)

well into the upper quadrants of wealth (e.g., ratio = 19.62). The summary statistics of the participants' gender, age, and SES in each sample are presented in Table 2.

Brain imaging data analysis

Planned comparisons focused on the following white matter tracts: The left and right AF, SLF, inferior longitudinal fasciculus (ILF), and inferior fronto-occipital fasciculus (IFOF). Two regression models were tested within each tract: (1) 26-month MPF values were regressed on each of the LENA language variables of interest (CS, PT, and CT), and (2) pre-reading measures collected at 5 years (ULN, ULS, LLN, LLS, PA, and EVT) were regressed on 26-month MPF values. Average MPF values were then extracted from regions with significant bivariate effects for both infant LENA measures and the 5-year language measures. These values were entered into a mediation analysis designed to test whether 26-month white matter organization accounts for the longitudinal relationship between infant language measures and 5-year pre-reading skills (see Figure 1).

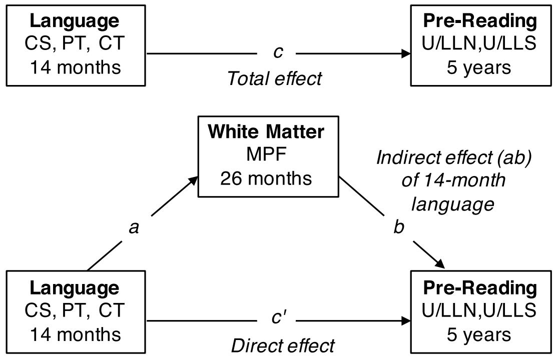


FIGURE 1
Theoretical mediation model. White matter MPF at 26 months was tested as a hypothetical mediator of the relationship between 14-month language measures [child speech (CS), parentese (PT), and conversational turns (CT)] and 5-year pre-reading skills [upper and lowercase letter name (U/LLN), and upper and lowercase letter sound (U/LLS), knowledge].

Results

Behavioral results

For the whole sample of 53 participants, the simple bivariate Pearson correlations between 6- and 24-months and 5-years of age measures revealed some significant results for both parental input and infants' output measures. Parental input, as measured by parentese (PT) when infants were 14-months-old is positively correlated with all skills measured at 5-years of age (i.e., uppercase and lowercase letters and sounds knowledge, phonological awareness, and vocabulary) at the age of 5-years. In addition, PT at 6-months and at 18-months old is positively correlated with lowercase letters knowledge at 5-years. However, there are no significant correlations between conversational turn-taking (CT) at 6–24-months-old and the 5-years-old measures.

Infants' output, as measured by children's production of speech-like vocalizations (CS) at the age of 14-months old is positively correlated with uppercase and lowercase letter-sound knowledge at 5-years. However, CS at 6-months old is negatively correlated with all 5-years measures. In addition, CS at 24-months old is negatively correlated with uppercase and lowercase letters knowledge. The results are summarized in Tables 3–5.

When examining the behavioral results for the sub-group of 20 participants that had also brain imaging data at the age of 26-months, the simple bivariate Pearson correlations between 14-months and 5-years of age measures revealed significant results for both parental input and infants' output measures. As found for the larger sample, PT at 14-months is positively correlated with uppercase and lowercase letters and sounds knowledge at the age of 5-years. However, it is not significantly correlated with phonological awareness, and vocabulary. PT at 6, 10, 18, and 24 months is not significantly correlated with any 5-years measures. Similarly, as found in the larger sample, CS at 14-months is positively correlated with uppercase and lowercase letter-sound knowledge at 5-years. In addition, CS at 24-months old is negatively correlated with uppercase and lowercase letters knowledge, as found in the bigger sample. In contrast to the larger sample, we also found a significant correlation between CT at 14-months-old and ULS knowledge, and between CT at 6-months and ULS and LLS knowledge (but not with

TABLE 3 Correlations between PT (parentese) at 6–24-months and reading-related measures at 5-years of age for the entire sample of 53 participants.

Age	**6 months**		**10 months**		**14 months**		**18 months**		**24 months**	
5-years measures	**Value**	**sig.**	**Value**	**sig.**	**Value**	**sig.**	**Value**	**sig.**	**Value**	**sig.**
ULN	0.023	0.115	0.194	0.185	0.476**‡	0.001	0.257	0.081	0.156	0.512
ULS	0.108	0.463	0.081	0.585	0.428**‡	0.002	0.204	0.169	0.317	0.174
LLN	0.315*	0.029	0.241	0.099	0.541**‡	> 0.001	0.342*	0.018	0.235	0.319
LLS	0.099	0.505	0.069	0.641	0.419**‡	0.003	0.191	0.199	0.213	0.366
PA	0.051	0.731	−0.034	0.817	0.328*	0.024	0.008	0.557	0.354	0.126
EVT	0.179	0.223	0.100	0.499	0.381**‡	0.008	0.225	0.129	0.285	0.223

All results are calculated for the raw scores for each test. *Significance level < 0.05. **Significance level < 0.01. ‡Significant after Bonferroni correction for multiple comparisons ($p \leq 0.008$).

TABLE 4 Correlations between CT (conversational turns) at 6–24-months and reading-related measures at 5-years of age for the entire sample of 53 participants.

Age	**6 months**		**10 months**		**14 months**		**18 months**		**24 months**	
5-years measures	**Value**	**sig.**	**Value**	**sig.**	**Value**	**sig.**	**Value**	**sig.**	**Value**	**sig.**
ULN	−0.018	0.902	−0.068	0.646	0.0175	0.233	0.152	0.302	0.239	0.323
ULS	0.089	0.548	0.055	0.708	0.264	0.070	0.158	0.283	0.389	0.099
LLN	0.028	0.848	0.023	0.877	0.232	0.112	0.192	0.190	0.228	0.348
LLS	0.082	0.581	0.050	0.737	0.269	0.064	0.172	0.242	0.317	0.187
PA	0.104	0.483	−0.037	0.804	0.260	0.074	0.054	0.714	0.188	0.442
EVT	−0.044	0.769	−0.141	0.340	0.235	0.108	0.079	0.595	−0.010	0.966

All results are calculated for the raw scores for each test.

TABLE 5 Correlations between CS (child speech and speech-like vocalizations) at 6–24-months and reading-related measures at 5-years of age for the entire sample of 53 participants.

Age	**6 months**		**10 months**		**14 months**		**18 months**		**24 months**	
5-years measures	**Value**	**sig.**	**Value**	**sig.**	**Value**	**sig.**	**Value**	**sig.**	**Value**	**sig.**
ULN	−0.329*	0.023	−0.101	0.493	0.103	0.487	−0.020	0.896	−0.657**‡	0.002
ULS	−0.436**‡	0.002	0.015	0.919	0.309*	0.032	0.011	0.941	−0.359	0.121
LLN	−0.391**‡	0.006	−0.074	0.618	0.249	0.087	−0.044	0.769	−0.603**‡	0.005
LLS	−0.399**‡	0.005	−0.034	0.817	0.347*	0.016	0.077	0.606	−0.389	0.090
PA	−0.288*	0.047	0.030	0.838	0.198	0.177	−0.048	0.750	−0.181	0.445
EVT	−0.388**‡	0.006	0.032	0.830	0.092	0.535	−0.042	0.780	−0.387	0.092

All results are calculated for the raw scores for each test. *Significance level < 0.05. **Significance level < 0.01. ‡Significant after Bonferroni correction for multiple comparisons ($p \leq 0.008$).

other 5-years-old measures). The results are summarized in Tables 6–8.

Brain imaging results

For the subgroup of 20 participants with MRI data at 26 months, significant bivariate correlations (2-tailed test; all significant results in the positive direction) were observed within the left AF and left SLF for PT and CT at 14 months, within the left AF for CS at 14 months, and within the left AF for LLN and ULS at 5 years (Figure 2). The significant effects within the left AF showed considerable spatial overlap across the 14-month and 5-year measures. Specifically, the Dice coefficients (Dice, 1945) for PT vs. ULS and LLN were 0.29 and 0.50, respectively. Dice coefficients for CT vs. ULS and LLN were 0.25 and 0.44, respectively. As shown in Figure 3, significant bivariate correlations were also observed for LLS in the left and right ILF, for PA in the left IFOF, and for EVT in the left AF, left and right SLF, right ILF, and right IFOF. However, the effects associated with LLS, PS, and EVT did not co-localize with 14-month CS, PT, or CT, and were less anatomically specific than effects observed for ULS and LLN.

TABLE 6 Correlations between PT (parentese) at 6–24-months and reading-related measures at 5-years of age for the smaller sample of 20 participants with brain imaging data at 26-months.

Age	6 months		10 months		14 months		18 months		24 months	
5-years measures	Value	sig.	Value	sig.	Value	sig.	Value	sig.	Value	sig.
ULN	0.086	0.720	−0.023	0.923	0.465*	0.039	0.103	0.666	0.156	0.512
ULS	0.106	0.658	0.213	0.336	0.668**‡	0.001	0.279	0.223	0.317	0.174
LLN	0.194	0.412	0.075	0.753	0.590**‡	0.006	0.290	0.214	0.235	0.319
LLS	−0.049	0.837	0.090	0.705	0.553*	0.011	0.163	0.492	0.213	0.366
PA	−0.128	0.590	−0.084	0.725	0.290	0.215	0.024	0.921	0.354	0.126
EVT	−0.153	0.521	−0.293	0.210	0.228	0.334	−0.119	0.618	0.285	0.223

All results are calculated for the raw scores for each test. *Significance level < 0.05. **Significance level < 0.01. ‡Significant after Bonferroni correction for multiple comparisons ($p \leq 0.008$).

TABLE 7 Correlations between CT (conversational turns) at 6–24-months and reading-related measures at 5-years of age for the smaller sample of 20 participants with brain imaging data at 26-months.

Age	6 months		10 months		14 months		18 months		24 months	
5-years measures	Value	sig.	Value	sig.	Value	sig.	Value	sig.	Value	sig.
ULN	0.131	0.594	−0.073	0.765	0.310	0.196	0.221	0.363	0.239	0.323
ULS	0.483*	0.036	0.228	0.348	0.473*	0.041	0.415	0.078	0.389	0.099
LLN	0.259	0.284	0.011	0.964	0.341	0.153	0.248	0.306	0.228	0.348
LLS	0.476*	0.039	0.192	0.432	0.409	0.082	0.416	0.077	0.317	0.187
PA	0.025	0.920	−0.171	0.484	0.285	0.236	0.078	0.752	0.188	0.442
EVT	−0.116	0.635	−0.286	0.235	0.242	0.319	−0.100	0.684	−0.010	0.966

All results are calculated for the raw scores for each test. *Significance level < 0.05.

TABLE 8 Correlations between CS (child speech and speech-like vocalizations) at 6–24-months and reading-related measures at 5-years of age for the smaller sample of 20 participants with brain imaging data at 26-months.

Age	6 months		10 months		14 months		18 months		24 months	
5-years measures	Value	sig.	Value	sig.	Value	sig.	Value	sig.	Value	sig.
ULN	−0.259	0.269	−0.024	0.920	0.350	0.130	0.024	0.922	−0.657**‡	0.002
ULS	−0.400	0.080	0.443	0.051	0.517*	0.020	0.135	0.569	−0.359	0.121
LLN	−0.381	0.097	0.035	0.883	0.430	0.058	−0.066	0.784	−0.603**‡	0.005
LLS	−0.264	0.261	0.450*	0.046	0.535*	0.015	0.227	0.336	−0.389	0.090
PA	−0.169	0.477	0.112	0.639	0.148	0.533	0.001	0.997	−0.181	0.445
EVT	−0.0141	0.552	0.121	0.610	0.147	0.536	−0.105	0.659	−0.387	0.092

All results are calculated for the raw scores for each test. *Significance level < 0.05. **Significance level < 0.01. ‡Significant after Bonferroni correction for multiple comparisons ($p \leq 0.008$).

To further assess the relationship between 14-month language measures, 26-months MPF, and 5-years pre-reading measures (ULS and LLN), average MPF values were extracted from regions with significant bivariate effects at both 14 months and 5 years (all within the left AF). Mean MPF values were then tested as mediators for each significant correlation between 14-months and 5-years behavioral measures. As shown in Tables 9, 10, the mediation analyses indicated that 26-months MPF values accounted for 19.73% of the total relationship between CS and ULS (indirect/total effect = 0.1973), and 29.65% of the total relationship between CT and ULS (indirect/total effect = 0.2965). Similarly, 26-month MPF values accounted for 14.92% of the total relationship between PT and LLN (indirect/total effect = 0.1492). In other words, the 26-months MPF values accounted for a significant portion of the variance shared between 14-months CT and 5-years ULS, 14-months CS and 5-years ULS, and 14-months PT and 5-years LLN, consistent with a mediation of these effects by the left AF. Note that CS and CT are not included in Table 10, because the 14-months CS correlations in the white matter did not co-localize with 5-years correlations, and neither of the 14-months CT or CS were significant predictors of 5-years LLN in the direct behavioral correlations.

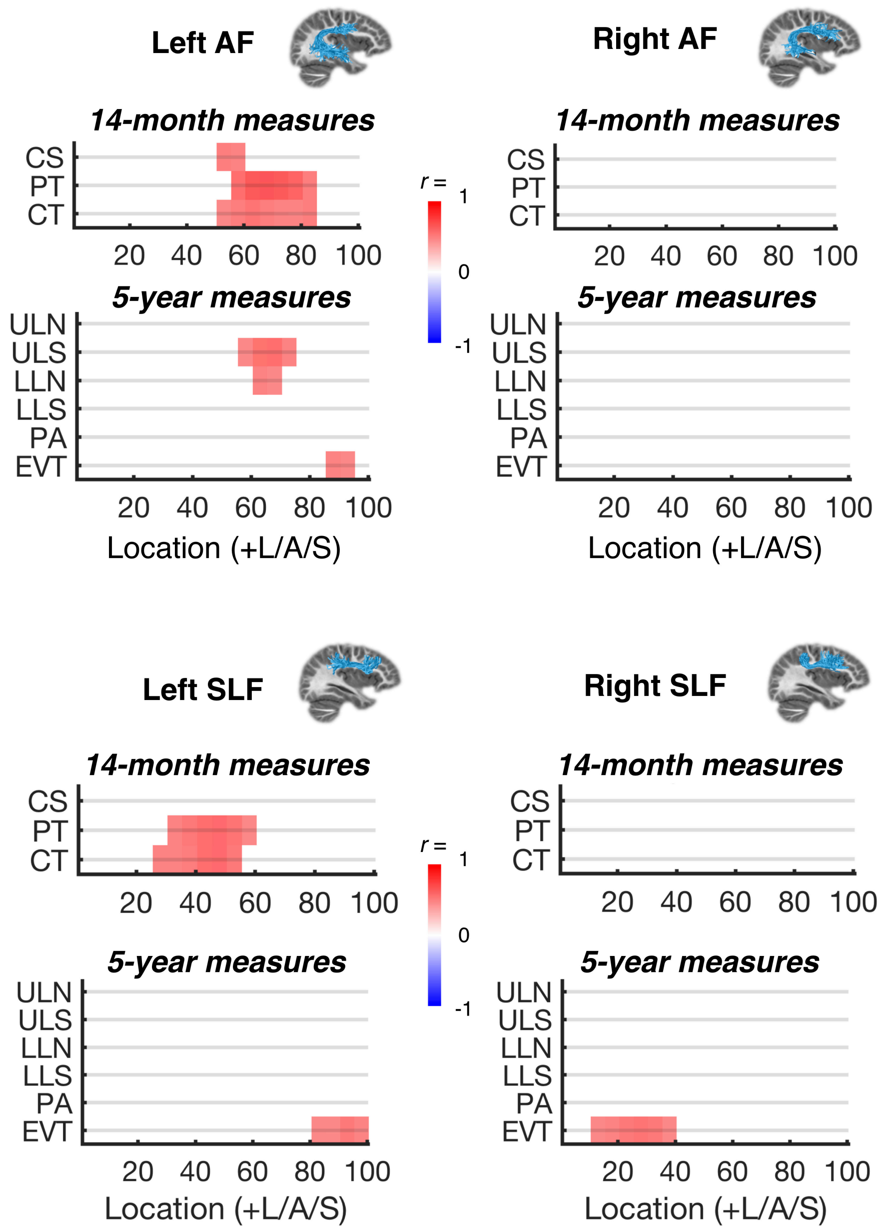


FIGURE 2

Correlations between language measures and MPF estimates within dorsal white matter pathways. Significant bivariate correlations ($p < 0.05$, corrected for multiple comparisons along each tract; see section "Materials and methods") are shown for each sampled white matter location within the left and right arcuate and superior longitudinal fasciculus (AF and SLF) and each behavioral variable: 14-month child speech (CS), parentese (PT), and conversational turns (CT); 5-year uppercase letter naming (ULN), uppercase letter sound knowledge (ULS), lowercase letter naming (LLN), lowercase letter sound knowledge (LLS), phonological awareness (PA), and expressive vocabulary test (EVT). Insets (top and middle right) show example tractography-based reconstructions for each of the white matter regions of interest.

Discussion

The goal of the current study was to examine how parent and child language variables measured in infancy relate to later emergent literacy skills, and whether white matter development mediates these relationships. We first examined correlations between early parental input and child output measures in the natural language environment from 6 to 24 months, and emergent literacy skills at the age of 5 years. We then examined correlations between quantitative MPF estimates of white matter myelination at 26 months and the longitudinal behavioral measures. Finally, we tested whether MPF values at 26 months mediated the relationships between early language measures and later literacy skills. Parental input and infants'

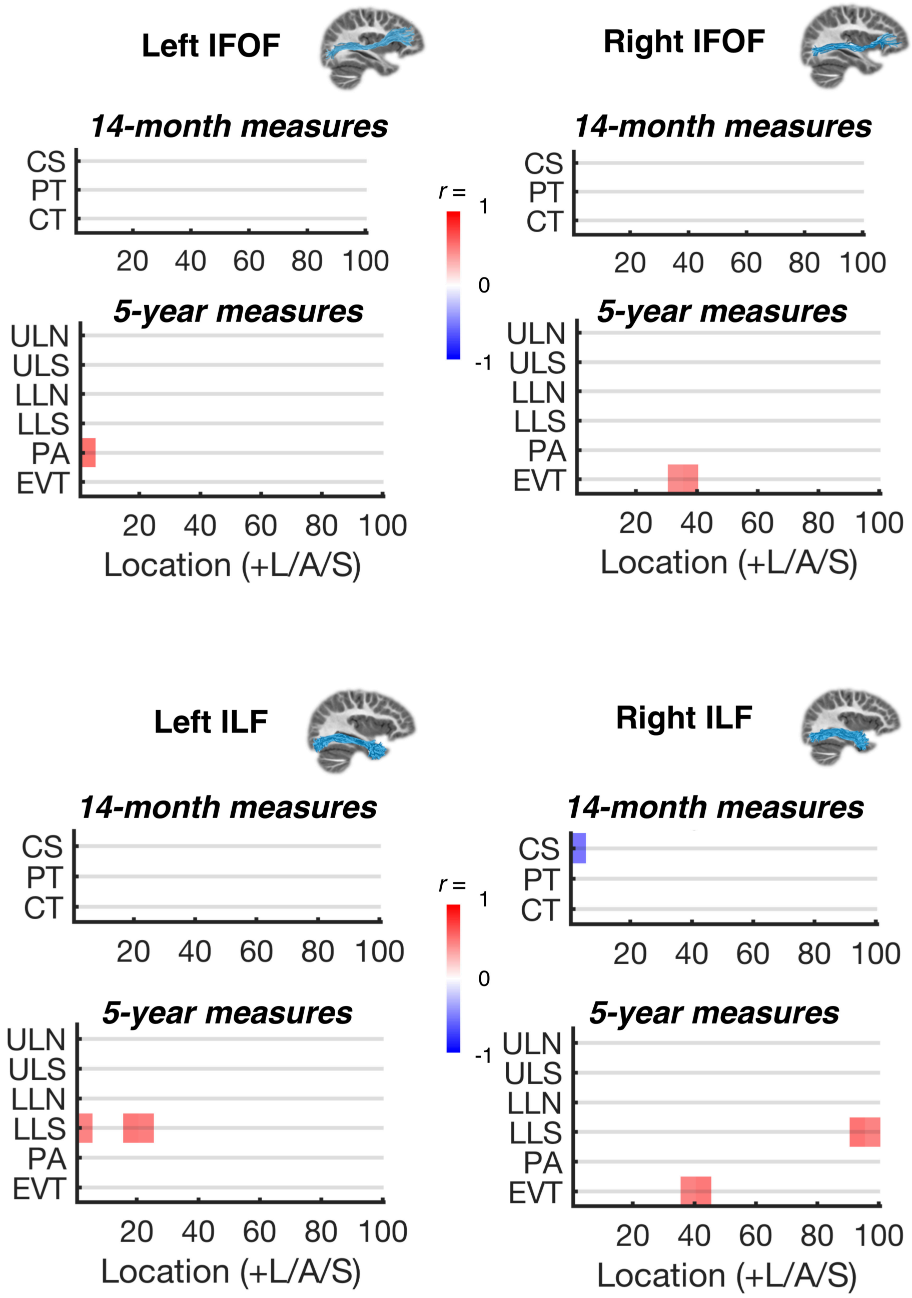


FIGURE 3
Correlations between language measures and MPF estimates within ventral white matter pathways. Significant bivariate correlations ($p < 0.05$, corrected for multiple comparisons along each tract; see section "Materials and methods") are shown for each sampled white matter location within the left and right inferior-fronto-occipital and inferior longitudinal fasciculus (IFOF and ILF) and each behavioral variable: 14-month child speech (CS), parentese (PT), and conversational turns (CT); 5-year uppercase letter naming (ULN), uppercase letter sound knowledge (ULS), lowercase letter naming (LLN), lowercase letter sound knowledge (LLS), phonological awareness (PA), and expressive vocabulary test (EVT). Insets (top and middle right) show example tractography-based reconstructions for each of the white matter regions of interest.

speech and speech-like vocalizations were found to predict emergent literacy skills at 5-years of age. Furthermore, myelin density estimates in the left AF were found to mediate the correlations between the early language measures and later emergent literacy skills. Together, these longitudinal results add to the literature relating to the long-term effect of early language skills and parental input and suggest that parental input and parent-infant interactions support the development of emergent literacy skills partly through myelination of the left arcuate pathway.

TABLE 9 Mediation analysis for uppercase letter sound knowledge (ULS).

	Total effect (c)		Direct effect (c')		Indirect effect (ab)	
14-month measures	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
PT	0.668**	[0.300, 1.037]	0.643*	[0.143, 1.144]	0.025	[−0.002, 0.052]
CS	0.517*	[0.094, 0.941]	0.415	[−0.071, 0.902]	0.102*	[0.078, 0.126]
CT	0.479*	[0.023, 0.936]	0.337	[−0.188, 0.863]	0.142*	[0.110, 0.175]

*Significance level < 0.05. **Significance level < 0.01.

TABLE 10 Mediation analysis for lowercase letter knowledge (LLN).

	Total effect (c)		Direct effect (c')		Indirect effect (ab)	
14-month measures	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
PT	0.590**	[0.190, 0.990]	0.501	[−0.038, 1.041]	0.088*	[0.060, 0.117]

*Significance level < 0.05. **Significance level < 0.01.

Correlations between early language input, child output, and emergent literacy skills

Previous studies have demonstrated that expressive and receptive language skills measured in the first 3 years of life predict literacy skills in kindergarten and elementary school, and emergent literacy skills in preschool (Silvén et al., 2002; NICHD Early Child Care Research Network, 2005; Flax et al., 2009; Deniz Can et al., 2013; Duff et al., 2015; Psyridou et al., 2018; Suggate et al., 2018). These relations have been explained by different models and theories. According to the lexical restructuring hypothesis (Metsala and Walley, 2009), vocabulary growth increases phonological awareness of smaller units, which in turn, drives decoding skills. Alternatively, according to the lexical quality hypothesis (Perfetti, 2007) and the triangle model (Seidenberg and McClelland, 1989), better semantic representations contribute to word identification. Our results extend prior findings by linking language experience and behavior in later infancy to specific pre-reading skills, which provide a foundation for early literacy.

Parentese at 14-months of age predicted emergent literacy skills at 5 years of age. Specifically, all of the 5-year measures, including letter names and sound knowledge, phonological awareness, and expressive vocabulary, in the full sample, and letter name and sound knowledge in the smaller sample of participants who had MRI data. These results are consistent with previous studies reporting that parental input and home language environment in the second year of life are important for literacy skill development (Silvén et al., 2002; Suggate et al., 2018; Tamis-LeMonda et al., 2019), and extend previous studies relating parentese input to expressive language skills (Ramírez-Esparza et al., 2014, 2017a,b; Ferjan Ramírez et al., 2019, 2020). It has been suggested that parentese affects speech development because this unique speaking style increases infants' perception of phonetic categories, and their ability to produce them (Kuhl et al., 1997; Cristià, 2011; Ramírez-Esparza et al., 2017a). The current study is the first to demonstrate a long-term effect of parentese input in infancy on emergent literacy skills at the age of 5-years.

The results of the behavioral analysis further indicate that Infants' speech and speech-like vocalizations at 14-months predict letter-sound knowledge at age 5 years. Recent work has demonstrated that speech-sound production in pre-readers uniquely predicts word later identification in 2nd grade, with additional mediating effects of phonological awareness and letter-knowledge skills (Mues et al., 2021). However, it is not yet clear how different elements of speech-sound production and expressive language relate to emergent literacy skills before the onset of literacy instruction. While the current study suggests a relationship between infant speech and speech-like vocalizations at 14 months and letter-sound knowledge in 5-year-old preschoolers, it should be noted that no clear relationship was found between early child speech output and subsequent 5-year vocabulary. While this is somewhat surprising, it is also important to note that the current measure of child output captures the quantity of child speech, but not lexical diversity or sophistication. Therefore, it does not perfectly correspond to early vocabulary skill. It is also possible that environmental factors at later ages moderate the relationship between these variables over time.

Child speech at 6 and 24 months was negatively correlated with later reading measures, which was unexpected. Importantly, the child speech measure does not differentiate among categories of speech-like vocalization, such as canonical vs. non-canonical babbling. Children with lower 5-year vocabulary skills might therefore produce more, but less sophisticated, vocalizations at 6 and 24 months. Alternatively, child output at these ages might be less reliably measured, especially at 6 months, where children produce fewer vocalizations, overall. Notably, within the MRI sample, negative effects at 6 months were smaller, and non-significant. Future work, using a finer-grained manual coding of the LENA recordings at each age, will be needed to evaluate these possibilities. In contrast, at 14 months, children can be expected to produce canonical babbling (syllables produced with adult-like consonant vowel transitions), as well as a small number of early words. This time point may therefore contain greater individual variability related to expressive language development.

Finally, the results of the behavioral correlations indicate that parent-infant interactions, indexed by conversational turns at 14-months, predict letter-sounds knowledge at the age of 5-years. This effect was strongest for letter-sound knowledge in the subset of participants with MRI data. Importantly, the MRI group primarily included families who previously participated in a parental language intervention (Ferjan Ramírez et al., 2019, 2020), and this group had higher overall conversational turn counts, as compared to the larger sample. Future work is needed to clarify whether these findings generalize to a larger sample with a greater range of conversational exposure, and whether these results are specifically related to the environmental enrichment provided by the intervention.

The current study is the first to demonstrate long-term effects of parent-infant turn-taking on letter-sounds knowledge at the age of 5-years. However, we note that the effect was strongest in the subset of participants with MRI data, and therefore needs further examination in a larger sample. Importantly, the MRI group primarily included families who previously participated in a parental language intervention (Ferjan Ramírez et al., 2019, 2020), and this group had higher overall conversational turn counts, as compared to the larger sample. Future work is needed to clarify whether these findings generalize to a larger sample with a greater range of conversational exposure, and whether these results are specifically related to the environmental enrichment provided by the intervention.

Altogether, the behavioral results from the current study tie together the conclusions from previous studies and illuminate that early parental input, infants' speech production, and parent-child interactions support not only language development, but the development of emergent literacy skills as well.

Brain-behavior correlations

Infant vocalizations, parentese speech input, and parent-child conversational turns at 14-months correlated with estimates of myelin density within left AF 26-months. Parent-child conversational turns, but not infant vocalizations, also correlated with left SLF myelination at 26-months. No significant correlations with early language measures were found for the right hemisphere white matter tracts, or for ventral pathways (IFOF and ILF), reinforcing the idea that the emerging language network is left-hemisphere dominant very early in development. These results extend the findings from previous studies indicating that parent-child conversational turns and language skills in 4–6-year-olds correlate with concurrent structural connectivity in the left AF and SLF (Romeo et al., 2018b), storytelling related activation in the left inferior frontal gyrus (Romeo et al., 2018a), and structural plasticity in the superior marginal gyrus (Romeo et al., 2021).

Further, letter name and sound knowledge at 5-years correlated with estimated myelin density in the left arcuate and the left and right ILF at 26-months. These results extend previous studies indicating that brain structure and function in the left dorsal pathways in 5-year-olds predict later reading skills including letter-word identification (Wang et al., 2020; Yamasaki et al., 2021), phonological awareness (Yu et al., 2018), cross-modal audio-visual processing (Gullick and Booth, 2014, 2015), and word and pseudoword reading (Van Der Auwera et al., 2021). The current study indicates similar correlations between earlier left dorsal white matter structure and 5-year-olds' letter names and sound knowledge, suggesting that these relations already exist in toddlers, long before the onset of reading instruction, and support reading acquisition.

Altogether, the results of the brain-behavior correlations from the current study indicate that the effects of early experience on the left dorsal pathways may have implications for later development of specific emergent literacy skills.

Mediation analysis

Myelination of the left AF, estimated using quantitative MPF mapping at 26-months, was found to mediate the relationship between parent-child conversational turns and child speech at 14 months, and letter sound knowledge at 5 years. Myelin density estimates in the left arcuate were also found to mediate the relationship between parentese at 14 months and letter name knowledge at 5 years. These findings suggest a potential biological mechanism underpinning for the longitudinal relationship between parent-child interactions and later decoding skills.

The left AF and SLF have previously been found to be related to the development of expressive and receptive language skills (O'Muircheartaigh et al., 2014; Deoni et al., 2016; Salvan et al., 2017; Swanson et al., 2017; Romeo et al., 2018a, 2021; Girault et al., 2019b; Sket et al., 2019), and emergent literacy skills (Lebel and Beaulieu, 2009; Yeatman et al., 2011; Saygin et al., 2013; Yu et al., 2018; Wang et al., 2020; Van Der Auwera et al., 2021; Yamasaki et al., 2021). Furthermore, previous studies have shown that the left

AF correlates with parent-child conversational turns in 4–6-year-olds (Romeo et al., 2018a,b, 2021), and that left dorsal activation and plasticity mediate the relations between parent-child conversational turns and comprehensive language skills (Romeo et al., 2018a). Plasticity of the left dorsal structures has also been shown to mediate the relationship between intervention-related changes in parent-child conversational turns, and gains in comprehensive language skills (Romeo et al., 2021). The current results extend these studies and suggest that the left dorsal white matter, and specifically the AF, might serve as a mechanism by which language experience in infancy supports the development of subsequent emergent literacy skills at age of 5 years.

It has been shown that parent-child conversational turns and parentese enhance the infant's speech and speech-like production, which in turn, encourage caregivers to respond and provide contingent feedback, and lead to a positive social feedback loop (Hirsh-Pasek et al., 2015; Gilkerson et al., 2018; Ferjan Ramírez et al., 2020; Romeo et al., 2021). The current data suggest that this process plays a role in the development of left dorsal language pathways, which in turn may facilitate emergent literacy skills. While the differing pattern for results for parentese vs. parent-child interaction hints that distinct developmental processes may be involved, future studies with additional neural measures and measurement time points are needed to clarify the contribution from specific components of early language input and experience.

Limitations and future directions

There are number of limitations in the current study which need to be mentioned. First, the sample is limited to native English speakers and children without known environmental or genetic risk factors, such as lower SES or family history of dyslexia. Further, the sample includes many families who participated a parental language intervention from 6 to 18 months (16 out of the 20 MRI participants, and 38 out of the total 53 participants), which was previously found to increase parentese speech and parent-child conversational turn taking (Ferjan Ramírez et al., 2019, 2020). Finally, the MRI sample size is relatively small ($n = 20$). Future work is therefore needed to clarify whether the current findings generalize to a larger and more demographically diverse sample, with a greater range of conversational exposure, and whether these results are specifically related to the environmental enrichment provided by the intervention.

Future studies should examine how additional variables, such as lexical diversity in parent and child speech, relate to both emergent and longer-term literacy skills, and how measures of brain structure and function might relate to these effects. These goals can be achieved by larger longitudinal studies that follow participants from infancy to school-age using multiple approaches to measure brain structure and function, and behavioral language and literacy skills. Furthermore, as mentioned in the Introduction, according to SVR (Hoover and Gough, 1990), both decoding and language comprehension skills predict the ability to develop good reading skills. Hence, future studies should further investigate how language comprehension skills such as vocabulary, morpho-syntax, and narrative skills in pre-readers relate to parental input and child language early in life.

Conclusion

In the current study we demonstrate for the first time a relationship between parental language input and parent-child interaction during late infancy and later emergent literacy skills in 5-year-olds, with an additional brain measure that suggests a biological mechanism for these effects, namely, developmental myelination of specific components of the left hemisphere's emerging language network, the left dorsal pathways. The key findings are that first, Infants' emergent speech production, together with parental use of parentese speech style and their conversational interactions with their infants directly predict emergent literacy skills, including letter names and sounds knowledge, phonological awareness, and expressive vocabulary in preschool. Second, Infants' emergent speech production, together with parental use of parentese speech style and their conversational interactions with their infants directly relate to myelination of left dorsal pathways (specifically the AF, and SLF) at the age of 26-months. Third, emergent literacy skills, and specifically letter names and sounds knowledge in 5-year-olds directly relate to myelination of the left dorsal pathways (specifically the AF) at the age of 26-months. Fourth, left AF myelination at 26-months may account for the relationship between measures of emergent speech production, parental input, and parent-child interactions in infancy and letter name and sounds knowledge in 5-year-olds.

These findings contribute to our understanding of the brain mechanisms involved in reading development and break new ground by suggesting a potential mechanism by which language experience early in life scaffolds later reading acquisition. Further research is needed to test the mechanism hypothesis.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, following reasonable request.

Ethics statement

The studies involving human participants were reviewed and approved by The University of Washington Human Subjects Division IRB Committee. Written informed consent from the participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

Author contributions

YW, EH, and PK contributed to the conception and design of the study. NF collected and analyzed the LENA data. NC and VY processed the quantitative MRI data to produce MPF maps. EH performed the diffusion MRI analysis. YW and EH organized the database, performed the statistical analysis, and wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

This research was supported by the Petunia Charitable Fund, the Overdeck Family Foundation, the Bezos Family Foundation, the University of Washington Institute for Learning and Brain Sciences Infants to Adolescents Project (Principal Investigator, PK), and the University of Washington's Language Acquisition and Multilingualism Endowment (Principal Investigator, NF).

Acknowledgments

We thank Julia Mizrahi, Bo Woo, Stephanie Purdy, Grace Stewart, Kyli McGillivray, Jake McManus, Jessica Culbreth, and Anna Kunz for assistance in participants recruitment, data collection, and data scoring. We also thank A'Dawnah Pangelinan, Annique Nyman, and Margaux Bouchegnies for manual coding of the recorded data.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Andersson, J. L., Skare, S., and Ashburner, J. J. L. (2003). How to correct susceptibility distortions,in spin-echo echo-planar images: Application to diffusion tensor imaging. *Neuroimage* 20, 870–888.

Andersson, J. L., and Sotiropoulos, S. N. (2016). An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage* 125, 1063–1078. doi: 10.1016/j.neuroimage.2015.10.019

Ashburner, J., and Friston, K. (1997). Multimodal image coregistration and partitioning—a unified framework. *Neuroimage* 6, 209–217. doi: 10.1006/nimg.1997.0290

Bulgarelli, F., and Bergelson, E. (2020). Look who's talking: A comparison of automated and human-generated speaker tags in naturalistic day-long recordings. *Behav. Res. Methods* 52, 641–653. doi: 10.3758/s13428-019-01265-7

Burnham, E. B., Wieland, E. A., Kondaurova, M. V., McAuley, J. D., Bergeson, T. R., and Dilley, L. C. (2015). Phonetic modification of vowel space in storybook speech to infants up to 2 years of age. *J. Speech Lang. Hear. Res.* 58, 241–253. doi: 10.1044/2015_JSLHR-S-13-0205

Chiu, Y. D. (2018). The simple view of reading across development: Prediction of grade 3 reading comprehension from prekindergarten skills. *Remedial Spec. Educ.* 39, 289–303. doi: 10.1177/0741932518762055

Church, J. A., Grigorenko, E. L., and Fletcher, J. M. (2021). The Role of neural and genetic processes in learning to read and specific reading disabilities: Implications for instruction. *Read. Res. Q.* 1–17. doi: 10.1002/rrq.439

Corrigan, N. M., Yarnykh, V. L., Hippe, D. S., Owen, J. P., Huber, E., Zhao, T. C., et al. (2021). Myelin development in cerebral gray and white matter during adolescence and late childhood. *Neuroimage* 227:117678.

Cristià, A. (2011). Fine-grained variation in caregivers' /s/ predicts their infants' /s/ category. *J. Acoust. Soc. Am.* 129, 3271–3280. doi: 10.1121/1.3562562

Cristià, A. (2013). Input to language: The phonetics and perception of infant-directed speech. *Lang. Linguist. Compass* 7, 157–170.

Cristià, A., Bulgarelli, F., and Bergelson, E. (2020). Accuracy of the language environment analysis system segmentation and metrics: A systematic review. *J. Speech Lang. Hear. Res.* 63, 1093–1105. doi: 10.1044/2020_JSLHR-19-00017

Cristià, A., Lavechin, M., Scaff, C., Soderstrom, M., Rowland, C., Räsänen, O., et al. (2021). A thorough evaluation of the language environment analysis (LENA) system. *Behav. Res. Methods* 53, 467–486. doi: 10.3758/s13428-020-01393-5

Demir-Lira, Ö, Applebaum, L., Goldin-Meadow, S., and Levine, S. (2019). Parents' early book reading to children: Relation to children's later language and literacy outcomes controlling for other parent language input. *Dev. Psychol.* 22:e12764. doi: 10.1111/desc.12764

Deniz Can, D., Ginsburg-Block, M., Golinkoff, R. M., and Hirsh-Pasek, K. (2013). A long-term predictive validity study: Can the CDI short form be used to predict language and early literacy skills four years later? *J. Child Lang.* 40, 821–835. doi: 10.1017/S030500091200030X

Deoni, S. C. L. L., O'Muircheartaigh, J., Elison, J. T., Walker, L., Doernberg, E., Waskiewicz, N., et al. (2016). White matter maturation profiles through early childhood predict general cognitive ability. *Brain Struct. Funct.* 221, 1189–1203. doi: 10.1007/s00429-014-0947-x

Dice, L. R. (1945). Measures of the amount of ecologic association between species. *Ecology* 26, 297–302. doi: 10.2307/1932409

Dickinson, D. K., Nesbitt, K. T., and Hofer, K. G. (2019). Effects of language on initial reading: Direct and indirect associations between code and language from preschool to first grade. *Early Child. Res. Q.* 49, 122–137. doi: 10.1016/j.ecresq.2019.04.005

Duff, F. J., Reen, G., Plunkett, K., and Nation, K. (2015). Do infant vocabulary skills predict school-age language and literacy outcomes? *J. Child Psychol. Psychiatry Allied Discip.* 56, 848–856. doi: 10.1111/jcpp.12378

Ferjan Ramírez, N., Hippe, D. S., and Kuhl, P. K. (2021). Comparing automatic and manual measures of parent–infant conversational turns: A word of caution. *Child Dev.* 92, 672–681. doi: 10.1111/cdev.13495

Ferjan Ramírez, N., Lytle, S. R., Fish, M., and Kuhl, P. K. (2019). Parent coaching at 6 and 10 months improves language outcomes at 14 months: A randomized controlled trial. *Dev. Sci.* 22:e12762. doi: 10.1111/desc.12762

Ferjan Ramírez, N., Lytle, S. R., and Kuhl, P. K. (2020). Parent coaching increases conversational turns and advances infant language development. *Proc. Natl. Acad. Sci. U.S.A.* 117, 3484–3491. doi: 10.1073/pnas.1921653117

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behav. Dev.* 8, 181–195.

Fernald, A., and Kuhl, P. K. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behav. Dev.* 10, 279–293.

Flax, J. F., Realpe-Bonilla, T., Roesler, C., Choudhury, N., and Benasich, A. (2009). Using early standardized language measures to predict later language and early reading outcomes in children at high risk for language-learning impairments. *J. Learn. Disabil.* 42, 61–75. doi: 10.1177/0022219408326215

Gilkerson, J., Richards, J. A., Warren, S. F., Montgomery, J. K., Greenwood, C. R., Kimbrough Oller, D., et al. (2017). Mapping the early language environment using all-day recordings and automated analysis. *Am. J. Speech Lang. Pathol.* 26, 248–265. doi: 10.1044/2016_AJSLP-15-0169

Gilkerson, J., Richards, J. A., Warren, S. F., Oller, D. K., Russo, R., and Vohr, B. (2018). Language experience in the second year of life and language outcomes in late childhood. *Pediatrics* 142:e20174276. doi: 10.1542/peds.2017-4276

Girault, J. B., Cornea, E., Goldman, B. D., Knickmeyer, R. C., Styner, M., and Gilmore, J. H. (2019a). White matter microstructural development and cognitive ability in the first 2 years of life. *Hum. Brain Mapp.* 40, 1195–1210.

Girault, J. B., Munsell, B. C., Puechmaille, D., Goldman, B. D., Prieto, J. C., Styner, M., et al. (2019b). White matter connectomes at birth accurately predict cognitive abilities at age 2. *Neuroimage* 192, 145–155. doi: 10.1016/j.neuroimage.2019.02.060

Goldstein, M. H., and Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychol. Sci.* 19, 515–523.

Grieser, D. L., and Kuhl, P. K. (1988). Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Dev. Psychol.* 24:14.

Gullick, M. M., and Booth, J. R. (2014). Individual differences in cross-modal brain activity predict arcuate fasciculus connectivity in developing readers. *J. Cogn. Neurosci.* 26, 1331–1346. doi: 10.1162/jocn_a_00581

Gullick, M. M., and Booth, J. R. (2015). The direct segment of the arcuate fasciculus is predictive of longitudinal reading change. *Dev. Cogn. Neurosci.* 13, 68–74. doi: 10.1016/j.dcn.2015.05.002

Hart, B., and Risley, T. R. (1995). *Meaningful Differences in the Everyday Experience of Young American Children*. Baltimore: Paul H Brookes Publishing.

Hartman, K. M., Ratner, N. B., and Newman, R. S. (2017). Infant-directed speech (IDS) vowel clarity and child language outcomes. *J. Child Lang.* 44, 1140–1162. doi: 10.1017/S0305000916000520

Helms, G., and Dechent, P. (2009). Increased SNR and reduced distortions by averaging multiple gradient echo signals in 3D FLASH imaging of the human brain at 3T. *J. Magn. Reson. Imaging* 29, 198–204. doi: 10.1002/jmri.21629

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nat. Rev. Neurosci.* 13, 135–145.

Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402.

Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Owen, M. T., Golinkoff, R. M., Pace, A., et al. (2015). The contribution of early communication quality to low-income children's language success. *Psychol. Sci.* 26, 1071–1083. doi: 10.1177/0956797615581493

Hjetland, H. N., Brinchmann, E. I., Scherer, R., and Melby-Lervåg, M. (2017). Preschool predictors of later reading comprehension ability: A systematic review. *Campbell Syst. Rev.* 13, 1–155. doi: 10.4073/csr.2017.14

Hjetland, H. N., Lervåg, A., Lyster, S. A. H., Hagtvet, B. E., Hulme, C., and Melby-Lervåg, M. (2019). Pathways to reading comprehension: A longitudinal study from 4 to 9 years of age. *J. Educ. Psychol.* 111, 751–763. doi: 10.1037/edu0000321

Hoover, W. A., and Gough, P. B. (1990). The simple view of reading. *Read. Writ.* 2, 127–160. doi: 10.1007/BF00401799

Hsu, N., Hadley, P. A., and Rispoli, M. (2017). Diversity matters: Parent input predicts toddler verb production. *J. Child Lang.* 44, 63–86. doi: 10.1017/S0305000915000690

Hurtado, N., Marchman, V. A., and Fernald, A. (2008). Does input influence uptake? links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Dev. Sci.* 11:F31–F39. doi: 10.1111/j.1467-7687.2008.00768.x

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., and Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Dev. Psychol.* 27:236. doi: 10.1044/1092-4388(2010/09-0216)

Huttenlocher, J., Vasilyeva, M., Cymerman, E., and Levine, S. (2002). Language input and child syntax. *Cogn. Psychol.* 45, 337–374. doi: 10.1016/S0010-0285(02)00500-5

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., and Hedges, L. V. (2010). Sources of variability in children's language growth. *Cogn. Psychol.* 61, 343–365.

Kendeou, P., van den Broek, P., White, M. J., and Lynch, J. S. (2009). Predicting reading comprehension in early elementary school: The independent contributions of oral language and decoding skills. *J. Educ. Psychol.* 101, 765–778. doi: 10.1037/a0015956

Kuhl, P. K., Andruski, J., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V., et al. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science* 277, 684–6. doi: 10.1126/science.277.5326.684

Kuhl, P. K., Tsao, F.-M., and Liu, H.-M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proc. Natl. Acad. Sci. U.S.A.* 100, 9096–9101. doi: 10.1073/pnas.1532872100

Lebel, C., and Beaulieu, C. (2009). Lateralization of the arcuate fasciculus from childhood to adulthood and its relation to cognitive abilities in children. *Hum. Brain Mapp.* 30, 3563–3573. doi: 10.1002/hbm.20779

Leech, K. A. (2021). An intervention to increase conversational turns between parents and young children. *J. Child Lang.* 48, 399–412. doi: 10.1017/S0305000920000252

Leemans, A., and Jones, D. K. (2009). The B-matrix must be rotated when correcting for subject motion in DTI data. *Magn. Reson. Med.* 61, 1336–1349. doi: 10.1002/mrm.21890

LENA (2015). *LENA^{TM} Pro Version 3.4.0*. Boulder, CO: LENA.

Liu, H., Kuhl, P. K., and Tsao, F. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Dev. Sci.* 6:F1–F10.

Ma, W., Golinkoff, R. M., Houston, D. M., and Hirsh-Pasek, K. (2011). Word learning in infant-and adult-directed speech. *Lang. Learn. Dev.* 7, 185–201.

Metsala, J. L., and Walley, A. C. (2009). Children ' s spoken word recognition and contributions to phonological awareness and nonword repetition: A 1-year follow-up. *Appl. Psycholinguist.* 30, 101–121.

Mori, S., Wakana, S., Nagae-Poetscher, L. M., and Van Zijl, P. C. M. (2006). MRI atlas of human white matter. *Am. J. Neuroradiol.* 27:1384.

Mues, M., Zuk, J., Norton, E., Gabrieli, J., Hogan, T. P., and Gaab, N. (2021). Clarifying the relationship between early speech-sound production abilities and subsequent reading outcomes. *Mapp. Intimacies* [Preprint]. doi: 10.31219/osf.io/zx2k3

NICHD Early Child Care Research Network (2005). Pathways to reading: The role of oral language in the transition to reading. *Dev. Psychol.* 41, 428–442. doi: 10.1037/0012-1649.41.2.428

Nichols, T. E., and Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Hum. Brain Mapp.* 15, 1–25. doi: 10.1002/hbm.1058

NRP. (2020). *Report of the National Reading Panel: Teaching Children to Read.* Midwestern: NRP.

O'Muircheartaigh, J., Dean, D. C., Ginestet, C. E., Walker, L., Waskiewicz, N., Lehman, K., et al. (2014). White matter development and early cognition in babies and toddlers. *Hum. Brain Mapp.* 35, 4475–4487. doi: 10.1002/hbm.22488

Ozernov-Palchik, O., and Gabrieli, J. (2018). Neuroimaging, early identification, and personalized intervention for developmental dyslexia. *Perspect. Lang. Lit.* 44, 15–20.

Pan, B. A., Rowe, M. L., Singer, J. D., and Snow, C. E. (2005). Maternal correlates of growth in toddler vocabulary production in low-income families. *Child Dev.* 76, 763–782. doi: 10.1111/j.1467-8624.2005.00876.x

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Sci. Stud. Read.* 11, 357–383.

Psyridou, M., Eklund, K., Poikkeus, A. M., and Torppa, M. (2018). Reading outcomes of children with delayed early vocabulary: A follow-up from age 2–16. *Res. Dev. Disabil.* 78, 114–124. doi: 10.1016/j.ridd.2018.05.004

Ramírez-Esparza, N., Garcia-Sierra, A., and Kuhl, P. K. (2014). Look who's talking: Speech style and social context in language input to infants are linked to concurrent and future speech development. *Dev. Psychobiol.* 17, 880–891. doi: 10.1111/desc.12172

Ramírez-Esparza, N., García-Sierra, A., and Kuhl, P. K. (2017a). Look who's talking NOW! Parentese speech, social context, and language development across time. *Front. Psychol.* 8:1008. doi: 10.3389/fpsyg.2017.01008

Ramírez-Esparza, N., García-Sierra, A., and Kuhl, P. K. (2017b). The impact of early social interactions on later language development in Spanish–english bilingual infants. *Child Dev.* 88, 1216–1234. doi: 10.1111/cdev.12648

Reynolds, J. E., Grohs, M. N., Dewey, D., and Lebel, C. (2019). Global and regional white matter development in early childhood. *Neuroimage* 196, 49–58.

Rodriguez, E. T., and Tamis-Lemonda, C. S. (2011). Trajectories of the home learning environment across the first 5years: Associations with children's vocabulary and literacy skills at prekindergarten. *Child Dev.* 82, 1058–1075. doi: 10.1111/j.1467-8624.2011.01614.x

Romeo, R. R., Leonard, J. A., Grotzinger, H. M., Robinson, S. T., Takada, M. E., Mackey, A. P., et al. (2021). Neuroplasticity associated with changes in conversational turn-taking following a family-based intervention. *Dev. Cogn. Neurosci.* 49:100967. doi: 10.1016/j.dcn.2021.100967

Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., et al. (2018a). Beyond the 30-million-word gap: Children's conversational exposure is associated with language-related brain function. *Psychol. Sci.* 29, 700–710. doi: 10.1177/0956797617742725

Romeo, R. R., Segaran, J., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., et al. (2018b). Language exposure relates to structural neural connectivity in childhood. *J. Neurosci.* 38, 7870–7877. doi: 10.1523/JNEUROSCI.0484-18.2018

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Dev.* 83, 1762–1774. doi: 10.1111/j.1467-8624.2012.01805.x

Rowe, M. L., and Snow, C. E. (2020). Analyzing input quality along three dimensions: Interactive, linguistic, and conceptual. *J. Child Lang.* 47, 5–21. doi: 10.1017/S0305000919000655

Rowe, M. L., and Weisleder, A. (2020). Language development in context. *Annu. Rev. Dev. Psychol.* 2, 201–223.

Salvan, P., Tournier, J. D., Batalle, D., Falconer, S., Chew, A., Kennea, N., et al. (2017). Language ability in preterm children is associated with arcuate fasciculi microstructure at term. *Hum. Brain Mapp.* 38, 3836–3847. doi: 10.1002/hbm.23632

Saygin, Z. M., Norton, E. S., Osher, D. E., Beach, S. D., Cyr, A. B., Ozernov-Palchik, O., et al. (2013). Tracking the roots of reading ability: White matter volume and integrity correlate with phonological awareness in prereading and early-reading kindergarten children. *J. Neurosci.* 33, 13251–13258. doi: 10.1523/JNEUROSCI.4383-12.2013

Seidenberg, M. S., and McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychol. Rev.* 96:523.

Share, D. L. (1995). Phonological recoding and self-teaching:< i> sine qua non of reading acquisition. *Cognition* 55, 151–218. doi: 10.1016/0010-0277(94)00645-2

Share, D. L. (2004). Knowing letter names and learning letter sounds: A causal connection. *J. Exp. Child Psychol.* 88, 213–233. doi: 10.1016/j.jecp.2004.03.005

Shneidman, L. A., Arroyo, M. E., Levine, S. C., and Goldin-Meadow, S. (2013). What counts as effective input for word learning? *J. Child Lang.* 40, 672–686.

Shrout, P. E., and Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychol. Bull.* 86:420. doi: 10.1037//0033-2909.86.2.420

Silvén, M., Niemi, P., and Voeten, M. J. M. (2002). Do maternal interaction and early language predict phonological awareness in 3- to 4-year-olds? *Cogn. Dev.* 17, 1133–1155. doi: 10.1016/S0885-2014(02)00093-X

Singh, L., Nestor, S., Parikh, C., and Yull, A. (2009). Influences of infant-directed speech on early word recognition. *Infancy* 14, 654–666.

Skeide, M. A., Brauer, J., and Friederici, A. D. (2016). Brain functional and structural predictors of language performance. *Cereb. Cortex* 26, 2127–2139.

Skeide, M. A., and Friederici, A. D. (2016). The ontogeny of the cortical language network. *Nat. Rev. Neurosci.* 17, 323–332. doi: 10.1038/nrn.2016.23

Sket, G. M., Overfeld, J., Styner, M., Gilmore, J. H., Entringer, S., Wadhwa, P. D., et al. (2019). Neonatal white matter maturation is associated with infant language development. *Front. Hum. Neurosci.* 13:434. doi: 10.3389/fnhum.2019.00434

Smith, S. M. (2002). Fast robust automated brain extraction. *Hum. Brain Mapp.* 17, 143–155. doi: 10.1002/hbm.10062

Song, L., Spier, E. T., and Tamis-Lemonda, C. S. (2014). Reciprocal influences between maternal language and children's language and cognitive development in low-income families. *J. Child Lang.* 41, 305–326. doi: 10.1017/S0305000912000700

Su, M., Zhao, J., de Schotten, M. T., Zhou, W., Gong, G., Ramus, F., et al. (2018). Alterations in white matter pathways underlying phonological and morphological processing in Chinese developmental dyslexia. *Dev. Cogn. Neurosci.* 31, 11–19. doi: 10.1016/j.dcn.2018.04.002

Suggate, S., Schaughency, E., McAnally, H., and Reese, E. (2018). From infancy to adolescence: The longitudinal links between vocabulary, early literacy skills, oral narrative, and reading comprehension. *Cogn. Dev.* 47, 82–95. doi: 10.1016/j.cogdev.2018.04.005

Swanson, M. R., Wolff, J. J., Elison, J. T., Gu, H., Hazlett, H. C., Botteron, K., et al. (2017). Splenium development and early spoken language in human infants. *Dev. Sci.* 20:e12360. doi: 10.1111/desc.12360

Tabors, P. O., Roach, K. A., and Snow, C. E. (2001). "Home language and literacy environment: Final results," in *Beginning Literacy with Language: Young Children Learning at Home and School*, eds D. K. Dickinson and P. O. Tabors (Baltimore: Paul H Brookes Publishing), 111–138. doi: 10.1111/1460-6984.12442

Tamis-LeMonda, C. S., Luo, R., McFadden, K. E., Bandel, E. T., and Vallotton, C. (2019). Early home learning environment predicts children's 5th grade academic skills. *Appl. Dev. Sci.* 23, 153–169. doi: 10.1080/10888691.2017.1345634

Thiessen, E. D., Hill, E. A., and Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy* 7, 53–71.

Tournier, J., Calamante, F., and Connelly, A. (2010). "Improved probabilistic streamlines tratography by 2nd order integration over fibre orientation distributions," in *Proceedings of the International Society for Magnetic Resonance in Medicine*, (Stockholm), 1670.

Uccelli, P., Demir-Lira, ÖE., Rowe, M. L., Levine, S., and Goldin-Meadow, S. (2019). Children's early decontextualized talk predicts academic language proficiency in midadolescence. *Child Dev.* 90, 1650–1663. doi: 10.1111/cdev.13034

Underhill, H. R., Rostomily, R. C., Mikheev, A. M., Yuan, C., and Yarnykh, V. L. (2011). Fast bound pool fraction imaging of the *in vivo* rat brain: Association with myelin content and validation in the C6 glioma model. *Neuroimage* 54, 2052–2065. doi: 10.1016/j.neuroimage.2010.10.065

Van Der Auwera, S., Vandermosten, M., Wouters, J., Ghesquière, P., and Vanderauwera, J. (2021). A three-time point longitudinal investigation of the arcuate fasciculus throughout reading acquisition in children developing dyslexia. *NeuroImage* 237:118087. doi: 10.1016/j.neuroimage.2021.118087

Wandell, B. A., and Yeatman, J. D. (2013). Biological development of reading circuits. *Curr. Opin. Neurobiol.* 23, 261–268. doi: 10.1016/j.conb.2012.12.005

Wang, J., Joanisse, M. F., and Booth, J. R. (2020). Neural representations of phonology in temporal cortex scaffold longitudinal reading gains in 5- to 7-year-old children. *NeuroImage* 207:116359. doi: 10.1016/j.neuroimage.2019.116359

Weisleder, A., and Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychol. Sci.* 24, 2143–2152. doi: 10.1177/0956797613488145

Weiss, Y., Yeatman, J. D., Ender, S., Gijbels, L., Loop, H., Mizrahi, J. C., et al. (2022). Can an online reading camp teach 5-Year-old children to read? *Front. Hum. Neurosci.* 16:793213. doi: 10.3389/fnhum.2022.793213

Williams, K. T. (2018). *Expressive Vocabulary Test- Third Edition*. London: Pearson.

Woodcock, R. W. (2011). *Woodcock Reading Mastery Tests- Third Edition*. London: Pearson.

Yamasaki, B. L., McGregor, K. K., and Booth, J. R. (2021). Early phonological neural specialization predicts later growth in word reading skills. *Front. Hum. Neurosci.* 15:674119. doi: 10.3389/fnhum.2021.674119

Yarnykh, V. L. (2007). Actual flip-angle imaging in the pulsed steady state: A method for rapid three-dimensional mapping of the transmitted radiofrequency field. *Magn. Reson. Med.* 57, 192–200. doi: 10.1002/mrm.21120

Yarnykh, V. L. (2012). Fast macromolecular proton fraction mapping from a single off-resonance magnetization transfer measurement. *Magn. Reson. Med.* 68, 166–178. doi: 10.1002/mrm.23224

Yarnykh, V. L. (2016). Time-efficient, high-resolution, whole brain three-dimensional macromolecular proton fraction mapping. *Magn. Reson. Med.* 75, 2100–2106. doi: 10.1002/mrm.25811

Yarnykh, V. L., Kisel, A., and Khodanovich, M. (2020). Scan-rescan repeatability and impact of B0 and B1 field nonuniformity corrections in single-point whole-brain macromolecular proton fraction mapping. *J. Magn. Reson. Imaging* 51, 1789–1798.

Yeatman, J. D., Dougherty, R., Myall, N., Wandell, B., and Feldman, H. (2012). Tract profiles of white matter properties: Automating fiber-tract quantification. *PLoS One* 7:e49790. doi: 10.1371/journal.pone.0049790

Yeatman, J. D., Dougherty, R. F., Rykhlevskaia, E., Sherbondy, A. J., Deutsch, G. K., Wandell, B. A., et al. (2011). Anatomical properties of the arcuate fasciculus predict phonological and reading skills in children. *J. Cogn. Neurosci.* 23, 3304–3317. doi: 10.1162/jocn_a_00061

Yu, X., Raney, T., Perdue, M. V., Zuk, J., Ozernov-Palchik, O., Becker, B. L. L. C., et al. (2018). Emergence of the neural network underlying phonological processing from the prereading to the emergent reading stage: A longitudinal study. *Hum. Brain Mapp.* 39, 2047–2063. doi: 10.1002/hbm.23985

frontiers | Frontiers in Neuroscience

TYPE Review
PUBLISHED 23 November 2022
DOI 10.3389/fnins.2022.1004027

The visual basis of reading and reading difficulties

John Stein*

Department of Physiology, Anatomy and Genetics, University of Oxford, Oxford, United Kingdom

Most of our knowledge about the neural networks mediating reading has derived from studies of developmental dyslexia (DD). For much of the 20th C. this was diagnosed on the basis of finding a discrepancy between children's unexpectedly low reading and spelling scores compared with their normal or high oral and non-verbal reasoning ability. This discrepancy criterion has now been replaced by the claim that the main feature of dyslexia is a phonological deficit, and it is now argued that we should test for this to identify dyslexia. However, grasping the phonological principle is essential for all learning to read; so every poor reader will show a phonological deficit. The phonological theory does not explain why dyslexic people, in particular, fail; so this phonological criterion makes it impossible to distinguish DD from any of the many other causes of reading failure. Currently therefore, there is no agreement about precisely how we should identify it. Yet, if we understood the specific neural pathways that underlie failure to acquire phonological skills specifically in people with dyslexia, we should be able to develop reliable means of identifying it. An important, though not the only, cause in people with dyslexia is impaired development of the brain's rapid visual temporal processing systems; these are required for sequencing the order of the letters in a word accurately. Such temporal, "transient," processing is carried out primarily by a distinct set of "magnocellular" (M-) neurones in the visual system; and the development of these has been found to be impaired in many people with dyslexia. Likewise, auditory sequencing of the sounds in a word is mediated by the auditory temporal processing system whose development is impaired in many dyslexics. Together these two deficits can therefore explain their problems with acquiring the phonological principle. Assessing poor readers' visual and auditory temporal processing skills should enable dyslexia to be reliably distinguished from other causes of reading failure and this will suggest principled ways of helping these children to learn to read, such as sensory training, yellow or blue filters or omega 3 fatty acid supplements. This will enable us to diagnose DD with confidence, and thus to develop educational plans targeted to exploit each individual child's strengths and compensate for his weaknesses.

KEYWORDS

dyslexia, phonology, temporal processing, magnocellular, visual, color filters, omega 3, auditory

EDITED BY
Johannes Ziegler,
Aix-Marseille Université, France

REVIEWED BY
Andrea Facoetti,
University of Padua, Italy
Marie Lallier,
Basque Center on Cognition, Brain and Language, Spain

*CORRESPONDENCE
John Stein
john.stein@dpag.ox.ac.uk

SPECIALTY SECTION
This article was submitted to Neurodevelopment, a section of the journal Frontiers in Neuroscience

RECEIVED 26 July 2022
ACCEPTED 16 September 2022
PUBLISHED 23 November 2022

CITATION
Stein J (2022) The visual basis of reading and reading difficulties. *Front. Neurosci.* 16:1004027.
doi: 10.3389/fnins.2022.1004027

Introduction

Human language evolved around 100,000 years ago, whereas writing was invented much later, only around 6,000 years ago. Also until the last 100 years or so, only a small proportion of humanity ever had to learn to read, and many of these were priests who were celibate, so they did not pass their genes onward. In consequence, whereas language is innate and most children instinctively copy their mothers and thus learn to talk, children usually have to be explicitly taught how to read. But in the UK and USA around 1 in 5 fail to do so, because our education systems are underfunded and reading is the most complicated cognitive skill that everybody is now expected to acquire. Here we will briefly review mainly the visual mechanisms underlying reading and their implications with regard to helping those who are finding the process difficult. But let it be emphasized at the outset reading is multifactorial; in particular not only accurate visual, but also auditory, processing is required.

Until recently most of our knowledge about the neurology of reading derived from observations on people who had failed to do so. The majority of those who fail, do so through low general intelligence and/or for social reasons: a toxic mixture of poor teaching, truancy, lack of parental support in impoverished and disadvantaged families. However, unfortunately the neural mechanisms underlying their failure are not studied much, and most studies are in children with normal or high intelligence and good family support, but who, despite these advantages, for some reason fail to learn to read fluently. These are the children who are said to have "developmental" dyslexia.

The word dyslexia, was derived by Berlin (1884) from the Greek "dys" meaning disordered, and "lexis" meaning words. He applied this word to the rare stroke patients he saw in whom cortical damage had selectively deprived them of the ability to read, but had not affected their speech, comprehension, nor their non-verbal reasoning powers. We now call this condition "acquired dyslexia." Dejerine even found patients who could no longer read, but could still write normally—"dyslexia without dysgraphia." Berlin thought the problem was mainly visual. So he also called it "word blindness." A few years later James Kirk and J. Pringle Morgan postulated that there might be a developmental analog of the brain damage condition, in which, for some reason, children fail to properly develop the brain circuits which mediate reading—"developmental dyslexia (DD)." His now famous patient, 14 year old Percy, was normally intelligent in oral conversation. But despite 9 years of schooling he hadn't even managed to learn to read or spell his own name properly yet (Morgan, 1896).

Morgan also noted that there were many persons in Percy's family with the same sorts of problem. So he speculated that the condition was "congenital." In the UK, Morgan, Kerr and Hinshelwood, and in the USA, Samuel Orton, all thought that it was a subtle visual processing problem, and continued to call it word blindness as well. But the term, developmental dyslexia (DD), soon replaced it because it became clear that sometimes this was not primarily a visual problem. It was assumed that DD was a genetically based failure to properly develop the correct brain connections for reading. And the early pioneers in the subject were confident that future research would uncover what these were. Meantime, the condition could be identified by demonstrating a big difference between a child's normal or high oral and non-verbal intelligence, yet markedly backward reading and spelling, together with a family history of similar problems, suggesting a genetic basis. This "discrepancy" criterion, together with the strong family history, served to identify and diagnose DD successfully for nearly 100 years.

The phonological theory

However, in the second half of the twentieth century the neurological approach to children's reading problems was gradually replaced by a linguistic/psychological one, because reading problems were, by then, mainly encountered by teachers and educational psychologists rather than neurologists. This process was especially influenced by the revolutionary ideas of Noam Chomsky. In particular, he showed how "recursion"—the repeated application of the same rule to its own output—could generate from a small number of speech sounds, "phonemes," a potentially infinite number of different words (Chomsky, 1957). This "phonological principle" underlies all language and writing. Accordingly, it was soon argued that DD was due to failure to acquire this phonological principle, without any visual processing problems (Liberman et al., 1971).

However, this theory does not attempt to explain *why* some children fail to grasp the phonological principle. Indeed it is almost a tautology, since it merely repeats, using different words, that the children have failed to learn to read, since the fundamental skill underlying reading is to learn to translate letters into the sounds (phones) they represent. Any child who cannot read has failed to acquire the phonological principle; hence all such children have a phonological deficit. But importantly there are numerous reasons, other than DD, why a child may fail to grasp this principle and fail to learn to read. These other causes are usually an interacting complex of social, rather than neurological, factors, such as poverty and deprivation, poor teaching, truancy, home chaos, low expectations and lack of family support. However, all the causes, including DD, manifest as failure to learn and absorb the phonological principle. Accordingly, almost all poor readers will fail phonological awareness tests, so low scores on such tests do not distinguish DD from any of the other causes. Hence, if we attempt to use these tests alone to identify DD, all poor readers will be diagnosed as dyslexic, and the diagnosis would become pointless because it would not point to any specific cause or to any treatment that might help them (Elliott, 2020).

The current situation

Indeed, this is what seems to be happening nowadays to a worrying extent. Almost everyone whose reading or other educational achievements are behind that expected for their chronological age may be called "dyslexic." Currently, practitioners are dissuaded from looking for discrepancies between reading and oral language competence to diagnose it, but instead they are instructed to emphasize phonological tests. As we have seen, since almost all poor readers fail these tests, they don't distinguish DD from other causes of reading failure, hence they cannot be used to identify it.

In reality, however, assessors actually do use discrepancy criteria, but covertly—for example by requiring a child to have at least a "normal" IQ for their age, yet recording more than 1.49 sds below average in their reading, phonological or temporal processing scores, in order to be classified as dyslexic. However, full scale IQ tests are highly unsuitable for measuring oral and non-verbal ability, since they involve a lot of reading (Thomson, 1982).

Furthermore, and in practice more importantly, measuring IQ and administering a variety of other psychometric tests takes up a great deal of time and expense, which has meant that only the relatively wealthy middle classes can afford to have their children tested privately. Many UK Education Authorities have now abandoned testing for DD at State expense, or even using the term at all. Now also practitioners look for many different possible discrepancies. The consequence is that the diagnosis of DD now covers a wide variety of learning disabilities and doesn't accord at all with the "classical" definition. The combination of the cost and the unclear criteria for identifying dyslexia specifically, has made the diagnosis inconsistent between practitioners and overall a muddle, and this is what has prompted some people to suggest that we abandon the concept altogether (Elliott and Grigorenko, 2014).

Return to discrepancy criterion?

This unhappy situation has arisen as a result of replacing the classical oral intelligence/reading discrepancy with some form of a phonological awareness/chronological age discrepancy definition, because the latter does not distinguish it from any of the other possible causes of reading failure. The solution should surely be to return to the classical reading/oral language discrepancy definition of DD which had served us well for so long, whilst awaiting the development of more objective biomarkers derived from our growing understanding of the physiology of what causes some children to experience such difficulties learning to read, despite having normal oral and non-verbal intelligence. Growth of such understanding was confidently expected by the likes of Morgan, Orton and Hinshelwood to explain how a hereditary disposition can cause learning to read to fail to progress normally in some children; and recent advances in the genetic approach suggest that such biomarkers will indeed be forthcoming in the near future (Mascheretti et al., 2018).

Genetic background

One of the few things that nobody disputes is that DD is strongly hereditary. Twin and family studies all agree that its heritability is around 60% (Olson, 2006). But only for rare single gene disorders is one gene alone, ever wholly responsible for a condition. None of the dozen or so gene variants, nor any of the more than 60 single nucleotide polymorphisms (SNPs) that have been shown to be associated with dyslexia, individually explain more than a tiny proportion of its heritability. Consequently so far, studying the genetic basis of reading difficulties has not added much to our understanding of the neural basis of reading. Nevertheless one of these genes, ROBO1, which helps to guide axons to their correct destinations during brain development, has been found to be associated specifically with visual motion sensitivity (a putative visual temporal processing marker) in dyslexics (Mascheretti et al., 2020). This is perhaps a first step in unraveling how gene variants might endow impaired visual processing to people with dyslexia.

Visual precedes phonological analysis

The great emphasis now placed on children acquiring the phonological principle is especially perverse in view of the fact that it was confirmed more than half a century ago what had been assumed for much longer, that visual processing is the actual starting point for learning to read (Morais et al., 1979). When children, or indeed illiterate adults, are first confronted with a written word they don't automatically see it as a series of letters, but as a single object like a mouse. When you see a mouse, you don't dissect it into a sequence of whiskers, nose, ears, head, body, tail, instead you see it as a single mouse. So the first thing somebody has to do to learn to read, is to learn to visually dissect words into their sequences of letters. Morais et al. (1979) showed that it is not until a child, or even an illiterate adult, has learnt that a written word consists of a sequence of letters, that they can begin to grasp that its spoken form consists of a series of more elementary sounds, "phonemes," which the letters in its written form stand for. So learning to sequence letters visually primes learning to sequence its sounds aurally. Thus visual sequencing always precedes phonological analysis. Often therefore, deficient visual processing is the cause of a dyslexic child's reading problems, but in others deficient auditory processing can be the main cause.

Timing and sequencing

Sequencing letters visually requires *timing* when and where a person sees the first letter, then the next and so on. This timing and rough localization depends on a subsystem of vision—the magnocellular system (from the Latin magnus meaning large) (Nassi and Callaway, 2009). 10% of the ganglion cells in the retina are these magnocells (M- cells) which are much larger than the rest, having receptive fields up to 50x larger than the much more numerous and smaller parvocells (from L. parvus meaning small). Being larger, magnocells have thicker axons and therefore conduct impulses into the brain much more rapidly than the P-cells; the M- cell volley normally reaches the visual cortex c. 10 ms earlier than the P-cells' (Maunsell et al., 1999). The M- cells' large receptive fields mean that they cannot signal fine detail, such as that which distinguishes one letter from another. Thus M- cells could not distinguish between the "d" and the "g" in the word "dog," for instance. But by directing the focus of attention to the right spot, they instruct P- cells to make that distinction (Vidyasagar and Pammer, 2010). Thus M-cells can signal when the eyes and focus of attention first alight on the "d" of "dog," then the "o," then the "g"; and clocking these movements enables representation in short term memory of the order in which the letters were seen. Both acquired and developmental dyslexics are conspicuously bad at correctly sequencing letters correctly.

The main reason for describing magnocellular and parvocellular retinal ganglion cells here, is because there is now overwhelming evidence that many, if not all, developmental dyslexics show impaired development of their visual magnocellular systems (Gori et al., 2016; Stein, 2019). This expresses itself not only in the retina, but in all the brain regions to which the magnocellular system provides significant input. Because magnocellular neurones provide rapid signaling of when and where visual events occur in the external world, their main function is to detect movement, for the visual guidance of attention and of the movements of the eyes and limbs. These functions are mediated by their dominant (90%) input to the "dorsal attentional system." This runs from the primary visual cortex forward via V5 (also known as MT) and the posterior parietal cortex (PPC) to the dorsolateral prefrontal cortex (dPFC) (Laycock et al., 2008). Impaired visuomotor function has been demonstrated in all these areas in dyslexics. Space forbids presenting all this evidence here, but some of the most persuasive will now be discussed.

Selective stimulation of magnocells

Much of this evidence showing that many children with dyslexia have impaired development of their visual magnocellular timing systems, depends on our ability to selectively stimulate these cells, rather than other retinal cells. This is possible because M- cells are most sensitive to low contrast, coarse, brief flashes of yellow light, operationalized as low contrast, low spatial and high temporal frequency stimuli, whereas parvocells are most sensitive to the converse: high contrast, high spatial but low temporal frequency stimuli. Therefore, the most convincing experiments that demonstrate specific abnormalities of the magnocellular system in dyslexics are those that employ low contrast, low spatial and high temporal frequency stimuli, and which compare these with their responses to high contrast, high spatial and low temporal frequency stimuli, to which parvo- cells respond better. These experiments have demonstrated that dyslexics' responses are significantly reduced to the former, but normal or even increased to the latter. This comparison also rules out the possibility that the people with dyslexia may be simply worse at all visual tests due to lack of attention or simply not bothering.

Retina

We will now consider experiments at the different levels of the visual system that demonstrate that in many dyslexics M-cell function is impaired, whereas P- cell function is normal or even enhanced. Lovegrove was the person who really initiated the magnocellular theory of dyslexia. He was the first to demonstrate that dyslexics have reduced sensitivity to gratings of low contrast and low spatial frequency flickered at high temporal frequencies (Lovegrove et al., 1980). Since M- ganglion cells respond best to stimuli with these properties, he suggested that the visual "transient" system in dyslexics was defective; this was the term used in 1980 for the visual magnocellular system. There have been innumerable studies that have confirmed his hypothesis as we shall see. Lovegrove also showed that in some circumstances people with dyslexia were actually more sensitive than ordinary readers to high contrast, high spatial frequency, low temporal frequency stimuli. This suggests that the parvocellular systems in dyslexia may actually be more sensitive, than those of ordinary readers (Lovegrove et al., 1982), a possibility that we will consider in greater detail later on.

The simplest low spatial and high temporal frequency stimulus is just a flickering light, and there have been numerous studies showing that in many dyslexics the frequency at which a flickering light ceases to appear to flicker, but becomes continuous (the so-called critical flicker fusion frequency- CFF), is significantly lower (Brannan and Williams, 1988; Johnston et al., 2017). Indeed the magnitude of this reduction predicts the degree of their reading failure (Talcott et al., 1998). This suggests that this magnocellular deficit may indirectly cause the reading problem, though of course it does not prove this. But this reduction in the

CFF is too small and variable to use diagnostically on an individual basis.

Magnocellular retinal ganglion cells respond "non-linearly"; they are ON/OFF cells, discharging as well to a light switching off as to switching it on, which means that they fire most at the 2nd harmonic of the frequency of a light whose intensity is varied sinusoidally. Thus M- cells give a larger response at the 2nd harmonic, whereas linear P cells do so mainly at the fundamental frequency. Hence typical readers give a larger cortical response at the second harmonic, but dyslexics do so at the fundamental frequency, and this phenomenon has been exploited by comparing the fundamental and 2nd harmonic peaks in steady state visual evoked potentials recorded from dyslexics as a simple test of their magnocellular sensitivity (Stein, 2021).

This non-linear property also explains why black and white stripes, switched on and off 10 times per second or faster, perceptually appear to be twice as fine as they actually are. This is the "spatial frequency doubling illusion" (Rosli et al., 2009). Dyslexics have been shown to need a higher contrast to see these gratings at all (Pammer and Wheatley, 2001) confirming that their M- cells are less sensitive. Again, the increase in contrast they require to see the gratings predicts their degree of reading failure.

Lateral geniculate nucleus

The axons of retinal magnocellular neurons project to the magnocellular layers (1 & 2) of the lateral geniculate nucleus (LGN) in the thalamus. 30 years ago, Livingstone et al. (1991) reported the results of a histological examination of the brain of some 70 year old dyslexics, who had first been seen by Samuel Orton in the 1920s. His patients had bequeathed their brains to the Orton brain bank, now in Harvard. Livingstone and Galaburda found that the LGN M-cells in these brains were significantly smaller compared with those in a control brain. Also they were much more disorganized, encroaching into the koniocellular area which normally clearly separates layers 2 and 3 from each other; in other words they had migrated too far during development. In the same brains Galaburda et al. (1985) had already shown similar excessive migration in the cerebral cortex of the dyslexic brains that had caused an unusually large number of "ectopias" (brain warts) to form on the cortical surface, particularly involving left hemisphere language areas.

Nowadays, higher strength 7 Tesla magnets have increased the spatial resolution of morphological MRI imaging down to less than a millimeter, which is sufficient to resolve the separate layers of the LGN, in life. Accordingly Giraldo-Chica et al. (2015) have now confirmed Livingstone's results in 15 awake volunteer dyslexics. The magnocellular layers of their LGNs were significantly thinner than in typically developing readers, particularly on the left, language, side.

Visual cortical areas

As far as the primary visual (striate) cortex, area V1, situated at the back of the occipital lobe, M- and P- inputs are anatomically separated. But in V1 they interact extensively. Accordingly, the M- "stream" cannot be said to be "pure" M- thereafter (Skottun and Skoyles, 2007). Nevertheless, as well as supplying 90% of the input to the dorsal attentional pathway mentioned earlier, M- cells actually also provide 50% of that to the ventral occipital form and pattern analyzing system, the other main visual pathway passing forward to the infero-temporal cortex. The function of this M-input to the ventral pathway is believed to be to draw attention to what part of visual space needs to be analyzed in detail by the P- system (Vidyasagar and Pammer, 2010).

The visual motion area- V5/MT

Thus the "visual motion area" at the front of the occipital lobe (V5, also known as MT) receives most of its input from retinal magnocells; and most of the neurons there are directionally selective, meaning that they have become specialized for detecting visual motion. In fact, they, like all the cells in the dorsal stream, are genetically related, as indeed are all the "transient" systems in the brain, because they all express similar surface "signature" molecules by which they recognize each other to make preferential connections. Thus M- like cells can be identified anywhere using antibodies specific for that lineage, such as CAT 301 (Hockfield and McKay, 1983); and they are found all over the brain.

The sensitivity of an individual's visual motion system can be tested psychophysically by measuring how many otherwise randomly moving dots on a screen have to move in the same direction, "coherently," in order for an observer to determine in which direction they're going (Britten and Newsome, 1992). The smaller the proportion required, the greater his motion sensitivity, and numerous studies have confirmed that dyslexics, as a group, have lower sensitivity than ordinary readers to these "random dot kinetograms" (RDKs). Furthermore the lower an individual's sensitivity, the worse his reading is (Cornelissen et al., 1995; Franceschini et al., 2012). However, again, individual responses are so variable that this test cannot be used for diagnosis in individuals.

Visual event related potentials

Sensitivity to visual motion can also be indexed by recording Visual event related potentials (VERPs). This was first carried out by Livingstone and Galaburda in 1991 (Livingstone et al., 1991). They measured visual evoked potentials in response to a moving checker board stimulus in 5 dyslexics and found that their latency was longer and amplitude smaller compared with 7 good readers. This result has been confirmed many times, in much larger samples and with more advanced technology (Klistorner et al., 1997; Schulte-Körne and Bruder, 2010; Jednoróg et al., 2011; Stein, 2021) and the result is no longer seriously doubted, although whether the deficit is due to undersampling due to smaller and sparser retinal magnocells, longer integration times or increased local noise, is still debated (Manning et al., 2019).

Functional magnetic resonance imaging

Advances in Functional magnetic resonance imaging (fMRI) have enabled the VERP results to be confirmed with much better spatial resolution using this technology. Thus Eden (Eden et al., 1996) was the first to show reduced activation of V5/MT in dyslexics viewing a moving pattern, compared with controls, an observation that has since been replicated by several other labs. Again, the size of the reduction predicts the subjects' degree of reading difficulty (Demb et al., 1998), but this expensive technique is unlikely to become useful for individual diagnosis.

Eye movement control

Having considered the evidence for impaired magnocellular development in DD at each level of the visual system, we will now discuss how this affects the cortical systems in which the M-system plays an important part in their control. Stable fixation on letters or words being inspected is obviously important for successful reading and this stability depends crucially upon detecting any motion caused by unwanted eye movements which may cause letters to appear to move around. This motion signal is fed back to the ocular motor control system, which negates it by directing the eyes back on to the target. A weak M system therefore leads to less stable visual fixation, which in turn leads to words and letters appearing to move around. This is a problem which many dyslexics experience (Fowler and Stein, 1979; Singleton and Trotter, 2005; Harries et al., 2015). Hence M- impairment impacts on dyslexics' eye movement control very significantly (Eden et al., 1994; Kirkby et al., 2008; Jainta and Kapoula, 2011), and the stability of the fixation of their eyes on letters is significantly reduced (Raymond et al., 1988; Fischer and Hartnegg, 2000).

Vergence

In addition when reading, the two eyes need to converge precisely to focus on letters 30 cms away, and the M- system is crucially involved in the first stage of controlling these convergence eye movements (Mowforth et al., 1981). But the vergence eye movement control system is highly vulnerable to drugs and disease, as we know to our cost if we consume too much alcohol; our eyes cease converging properly and things can seem to go double (diplopia). We and many others have shown that many dyslexics have this kind of unstable vergence control (Fowler and Stein, 1983; Liversedge et al., 2006; Bucci et al., 2007), hence they have a pronounced tendency to experience diplopia when attempting to read.

Furthermore, in individuals the degree of reduction in their visual motion sensitivity predicts the extent of their eye instability problems (Ray et al., 2005). Indeed, motion sensitivity, i.e., M-cell sensitivity, predicts orthographic reading skill, not just in dyslexics, but in everyone (Witton et al., 1998). Thus, because all eye movements depend on M- control, M- insensitivity leads to impaired control of all kinds of eye movement. For instance, when the eyes track a moving target (in "smooth pursuit"), dyslexics tend to fall progressively behind it; so they have to make periodic saccades in order to catch up. Thus such "saccadic intrusions" are much commoner in dyslexics (Adler Grinberg and Stark, 1978; Eden et al., 1994).

In people with dyslexia, the accuracy of saccades is impaired whatever the target, not just when they are reading (Biscaldi et al., 2000). Thus much of the evidence for their impaired eye movement control is derived from recording their responses to targets other than text, i.e., not involving reading at all. This provides yet more evidence suggesting that their poor eye control is a cause of their impaired reading, rather than being just a result of it.

Nevertheless many people still believe that all the eye movement abnormalities found in dyslexics are the result of their difficulties with decoding, rather than their cause (Rayner, 1998). There is no doubt that increased numbers of regressive saccades (returning to words not successfully decoded) and prolonged fixations (due to their longer decoding time) may be partly caused by their decoding problems. But this does not explain why their fixations are so unstable, nor why the eyes diverge inappropriately and cause diplopia, nor why these problems occur when inspecting any visual sequences, not just text.

Visual attention

Because the visual magnocellular system guides attention as well as eye movements, M deficiency leads to slower, less accurate deployment of visual attention (Vidyasagar, 2005). "Serial visual search" is when each in an array of similar targets has to be inspected, one after the other, in order to detect a particular one, whereas "parallel search" is when a feature in one of the objects is so distinctive it just "pops out." Dyslexics are as good or better at parallel search, but much slower at serial search (Vidyasagar and Pammer, 1999; Facoetti et al., 2000;

Iles et al., 2000), and this explains why they are slower and less accurate at any kind of visual sequencing, not just of letters.

Text is an extremely "crowded" visual stimulus. Amidst the distracting surrounding text, the M- system normally directs our attention accurately onto the particular word we are trying to comprehend, but the closer letters and words are to each other, the more difficult it is to concentrate on just this one word. This crowding interferes greatly with the ability to pick it out and accurately read it. This problem is much more pronounced in dyslexics than in good readers. So using more widely spaced print often helps them considerably (Cornelissen et al., 1991; Martelli et al., 2009).

As we have seen the essence of learning to read is learning to associate the shape of a letter with the sound it represents and that the letters translate into those particular sounds. So visual attention must be properly cued to those sounds. But this visual/auditory crossmodal cueing of attention is greatly impaired in those with dyslexia (Gabrieli and Norton, 2012; Harrar et al., 2014).

In order to argue against the possibility that a visual processing deficit may contribute to dyslexia, it is often suggested that people with dyslexia are just bad at all tests, due to a general lack of concentration and motivation; so it is argued that there may be nothing specific about their impaired M-function, but they just don't try hard enough. However, if this were so, they would be equally bad at detecting stimuli designed to stimulate P cells selectively. As we have seen, many studies have compared dyslexics' M- sensitivity to their P- sensitivity to stimuli such as static visual forms, and found it normal. Indeed, as mentioned earlier, they may actually have higher contrast sensitivity for functions mediated by the parvo system, e.g., at high spatial and low temporal frequencies (Lovegrove et al., 1982) and they have better color discrimination in the visual periphery (Dautrich, 1993), possibly because they have more P-cells in the peripheral retina.

The reading connectome

Until recently, the neurological approach to brain function, cognitive skills and disease rested on the classical idea of "localization of function." For each cognitive function, it was argued, there would be a single specific cortical area that is mainly responsible. But as we learn more about how brain structures interact, it has become clear that whilst the primary sensory and motor areas are indeed functionally localized, in the rest of the cortical "association areas," it is more useful to consider how they connect with each other, their "connectomes," than to concentrate on a single area (Dick et al., 2013). Recent advances in functional MRI, in particular diffusion tensor imaging (DTI), allows the functional connectivity between areas to be traced and also to investigate how these change not only over the long term during development (Feng et al., 2022), but also over shorter terms to mediate the acquisition of new cognitive skills, such as reading.

To simplify drastically, the cerebral cortex can be seen as a network of nodes connected to each other by either long or short interconnections, and although these are initially set up by the genes you inherit, they develop continuously throughout life, but particularly in childhood, according to your environment, nutrition, education, experiences, actions, emotions and memories. Those that contribute to successful processing survive; those that do not are eliminated. For reading, the optimal end result is a "small world" network, richly connected to neighboring, but sparsely connected to distant nodes (Bullmore and Sporns, 2012). Thus in a recent UK study, reading and maths networks were compared in a large number of children covering the full range of abilities in these two domains. There was no evidence that the poorer readers had missing connections or nodes compared with the good readers, but the strength of some crucial connections were much weaker in the poorer readers compared with the better ones (Bathelt et al., 2017).

Dyslexia seems to be equally common in the very different Chinese character script (Peng et al., 2017). Given the very visual nature of the characters it was not surprising that network analysis showed that fluent adult Chinese readers develop significantly stronger visually based connections than do children who have not yet learnt to read Chinese fluently (Zhou et al., 2021).

Cause or effect? Reading age matches

One problem with many of the results that we have discussed so far, is that reduced M cell sensitivity could conceivably be a result rather than a cause of failing to learn to read, because the children will have had less practice at the required visual skills (Goswami, 2015), and therefore might have failed to develop them because of this lack of exposure. One way of avoiding this problem is to compare dyslexics' M sensitivity with that of younger, typically developing children, matched for reading age with the dyslexics. This is known as a "reading age match" design. Under these conditions, these younger children will have had no greater experience of reading than the older dyslexics have had, and therefore their M- sensitivity should be no better than that of the dyslexics if visual experience were the crucial factor. However, as we have seen, many studies have shown clearly that younger readers who have had the same amount of reading experience as people with dyslexia, have already developed much better visual M- function (Gori et al., 2016).

Nevertheless it is clear that impaired visual M- function alone is neither a sufficient, nor a necessary cause of dyslexia, and therefore cannot be advanced as its sole cause. But it probably makes an important contribution in most dyslexic people. As we shall see, however, auditory and probably other neural temporal processing problems may also be relevant.

Cohort studies

Another powerful way to demonstrate that timing and sequencing deficits may be a cause of children's reading problems, rather than merely their consequence, is to look for M- deficits in children before they even begin to learn to read. The best way to do this is to select children at genetic risk of dyslexia on the basis that close relatives in the family have been diagnosed as dyslexic, so they have a much greater risk of becoming dyslexic themselves. Better still, make this a "cohort" study, where the children at genetic risk of dyslexia are studied both before they begin to learn to read and then again at intervals after they have or have not managed to do so. Two large studies of this sort have been carried out, one in Finland (Hamalainen et al., 2008) and the other in Holland (van der Leij, 2013). They have both shown that neurological differences manifest themselves in infants, and even in newborn babies, in those who are going to go on to become dyslexic later on. In the Dutch study of children at family risk, habituation to a simple visual, non-alphabetic, stimulus measured by ERPs at the age of 5, before learning to read had commenced, predicted whether the child would be identified as dyslexic by the age of 8 (Regtvoort et al., 2006). Thus these cohort studies have established that impairments in temporal processing precede a child's failing to learn to read and therefore strongly support the hypothesis that they contribute causally to any subsequent failure.

Intervention studies

However, the most convincing way of demonstrating that impaired visual magnocellular performance causes visual difficulties with reading is to show that improving it by training, or in other ways, helps children to overcome those difficulties and to make significant progress with their learning to read. Ideally, we would show this by means of randomized control trials (RCTs). Unfortunately given the opposition to the whole concept (leading to lack of funding) only a few of these have been carried out. Even though all have come to positive conclusions, none of them have been large enough to overcome the opposition. Space forbids much detail here, but we can consider some of this evidence.

Lawton has spent a lifetime developing a system based on the known properties of the M-system for training poor readers with an initially weak M-system (Lawton, 2016). She uses a low contrast, low spatial frequency, moving grating presented against a static background grating of higher contrast and spatial frequency. Subjects have to report the direction of motion whilst the program iteratively reduces the contrast of the moving grating as their threshold for correct responses improves. This training procedure greatly helps the children to improve their M- sensitivity. Thus she can simultaneously measure the children's current motion sensitivity, how this improves with her training and whether this enhances their reading progress. And this has shown clearly that this training helps the majority of dyslexic children to improve their M- sensitivity and thereby make considerable progress with their reading, thus offering powerful evidence that their improved M- function was what enabled their improved reading (Lawton and Stein, 2022).

Action video games

Another way of training M- function, not involving reading, that has been used by many labs is to ask the children to play action videogames. These incorporate the active tracking of moving targets with the eyes and limbs (Bavalier and Davidson, 2013), obviously engaging the M- system. These games have proved to be extremely effective in improving children's M- function. Accordingly, their ability to focus visual attention and concentrate on the task in hand increases. Indeed, not only is the direction of visual attention improved, but also that of auditory attention (Mancarella et al., 2022) consequently these changes are followed by impressive improvements in their reading (Peters et al., 2019).

Color filters

Perhaps the most controversial way of trying to improving magnocellular function is the use of color filters. These have mainly been used on the basis of other theories of why children can have visual reading problems. Here we will discuss their relevance to the magnocellular theory.

Yellow filters

M-cells do not contribute to color vision, but they receive mainly from the retinal cone photoreceptors that are sensitive to Long ("red") and Middle ("green") wavelengths. Hence they are maximally stimulated by the combination of these two wavelengths, which is yellow light. This color is often called "unique yellow" because most observers deem it pure yellow, not tinged with red or green. It is also the dominant wavelength of day light at midday. Band pass filters passing mostly yellow light at 575 nM actually increase the amount of yellow light falling on the retina because with less light entering the eye overall, the pupils dilate. We have shown that viewing text through such filters helps some, but not all, children with visual problems when reading, to overcome them, hence to progress faster, sometimes dramatically. Since we showed that wearing these filters also improved the children's visual motion sensitivity, fixation stability and vergence control we concluded that the yellow filters had improved the children's M- function, hence that this was what enabled their reading progress (Ray et al., 2005).

Blue filters

Another group of dyslexics seem to benefit from viewing text through the opposite color, "unique blue" filters (Clisby et al., 2000). These pass most light at c. 475 nanometers—Oxford blue. Unique blue filters probably work via the brain's internal clock in the hypothalamic suprachiasmatic nucleus (SCN). This clock needs to be synchronized with sunrise which changes throughout the year. At sunrise the first rays of the sun in the morning are blue, and they are detected ganglion cells in the retina that have recently been discovered which respond especially to morning light, because they contain a blue sensitive pigment, melanopsin (Spitschan, 2019). They project to the hypothalamus and specifically activate the M- timing cells via a connection from the SCN to another "blue nucleus," the locus coeruleus (coeruleus, Greek—blue). Blue light thereby increases general arousal and with it children's ability to concentrate, and this is probably how it helps them to improve their reading.

Thus all these treatment interventions that target the magnocellular timing systems can help children with dyslexia to learn to read. Hence they add to the mounting evidence that impaired M- cell development throughout the CNS is an important contributor to dyslexic people's reading problems.

Criticisms, controversies and counter evidence

In summary, after the introduction of the visual transient/magnocellular hypothesis of dyslexia, the great majority of studies asking whether a visual magnocellular impairment can be detected, have found that indeed it can, in at least some dyslexics. In addition, visual attention, visual search and eye movement control are all agreed to be mainly mediated by the M-system. So the many experiments that have demonstrated deficiencies in these systems in dyslexics can be confidently interpreted as supporting the magnocellular hypothesis. Finally, as we have seen, there are now a large number of studies aimed at bolstering M- sensitivity either directly or indirectly, and most of these have demonstrated convincingly that any technique which improves M- cell function can help children to improve their reading.

With such strong evidence, one might well ask why this idea is so often dismissed as disproven, too highly controversial or simply ignored. The main reason is probably the overwhelming dominance of the phonological theory. But as we have seen this theory is more of a tautology, merely repeating that the children find reading difficult, rather than explaining why they fail. Yet many people mistakenly believe that it rules out any visual processing explanation for dyslexia. 50 years ago, however, the work of Morais (Morais et al., 1979) and a succession of later research, has clearly demonstrated the phonological deficit in dyslexics is likely to be often caused by their visual sequencing failures.

Nevertheless, a visual magnocellular weakness cannot be detected in every dyslexic; moreover some children with mildly impaired magnocellular function may still learn to read. Hence a visual M- deficit cannot be said to be the only cause of dyslexia (Skottun and Skoyles, 2004). Yet M- cell sensitivity predicts orthographic skill in both good and bad readers, whether or not they are classed as dyslexic (Talcott et al., 2002). Thus visual M- function appears to contribute significantly to everybody's ability to acquire reading skills. But visual M- weakness seems to confer vulnerability to reading problems, but it is not their sole cause.

Auditory temporal processing

Probably auditory temporal sequencing is also deficient in many people with dyslexia. After learning to sequence the letters in a word visually, the next stage in learning to read is to learn to translate this sequence into the sequence of sounds in the spoken form of the word. This requires the child to learn to time the order in which the sounds occur in the word. This requires precise auditory timing. The auditory timing system has transpired to be remarkably analogous to the visual timing M- system; it is sometimes even called the auditory transient or magnocellular system (Rauschecker, 2018; Meng and Schneider, 2022). Large neurones in the auditory brainstem (cochlear nuclei) and thalamus (the medial geniculate nucleus—MGN) are specialized for the much more precise timing required by audition than for vision. Importantly, like those in the visual LGN, they bind the M- cell specific signature antibody, CAT 301 (Hendry et al., 1988); hence histologically the auditory transient neurones can be considered part of the same neuronal lineage.

This paper is about the visual basis of dyslexia, so we will not consider the auditory analog of the visual transient system in any detail. Suffice it to be said that development of the auditory transient cells is also impaired in many dyslexics, so that anatomical (Galaburda et al., 1994), psychophysical (Hämäläinen et al., 2012) and electrophysiological (Schulte-Körne et al., 1998) studies have all found that dyslexics' auditory temporal processing systems are abnormal. Their responses to frequency or amplitude modulated sounds are reduced compared with good readers, and the degree of this reduction predicts the impairment of their reading (McAnally and Stein, 1996; Goswami, 2014). In cohort studies these abnormalities have been shown to be evident in those at family risk of dyslexia long before learning to read begins, even in new born babies, and they strongly predict the babies' eventual chances of suffering reading problems (Leppänen et al., 2010). Thus, in many ways, auditory temporal processing impairments parallel those seen for the visual M- system in dyslexia, and suggest that both can contribute to causing dyslexia.

Importantly, and probably as a result of their common genetic basis, these visual and auditory temporal processing deficiencies correlate strongly with each other in individuals tested for both (Talcott et al., 2002). Thus the majority of dyslexics appear to have both visual and auditory timing problems and their genetic basis probably represents the basic cause of their visual, phonological and reading difficulties.

Magnocellular timing systems

If we define transient/magnocellular lineages by their expression of the same surface antigens, such as CAT 301 (Hockfield and McKay, 1983), they are not even confined to the visual and auditory systems. They form interconnecting networks, specialized for temporal processing, hence for sequencing, all over the brain. They track changes in light, sound, limb position, etc., for the direction of attention and the control of movement. They are found not only in the visual and auditory systems, but also in the touch and proprioceptive systems (Stoodley et al., 2000), in motor systems, cerebral cortex, hippocampus, cerebellum and brainstem (Hockfield and McKay, 1983). Their large, rapidly conducting, rapidly transmitting, neurons mediate speedy temporal processing. Everywhere, they are also extremely vulnerable. Altered magnocellular development has been reported in fetal prematurity, in fetal alcohol syndrome, DD, dyspraxia, dysphasia, ADHD, ASD, Williams syndrome, and even in schizophrenia and bipolar depression (Whitford et al., 2017).

Docosahexaenoic acid

The high dynamic sensitivity of M- cells is mediated by their membrane ionic channels being able to open and close very rapidly. This requires the membrane to be highly flexible; and this flexibility is crucially dependent on one very important component of the diet, the omega 3 long chain fatty acid, docosahexaenoic acid (DHA) (Haag, 2003; Muskiet et al., 2004). This is normally provided by consuming oily fish. Green vegetables, flax or rape seed and seaweed all contain the shorter chain omega 3, alpha linolenic acid, but humans do not convert this into DHA very efficiently.

Unfortunately consumption of oily fish has decreased greatly over the last 50 years, particularly in disadvantaged households (Maguire and Monsivais, 2015), and so we wondered whether this might be affecting M- cell responses in the children. We therefore carried out a randomized control trial to see whether giving disadvantaged children dietary supplements of DHA could improve their M- cell responses and thereby help them to improve their reading (Richardson and Montgomery, 2005). After 3 months of consuming the supplements, the children's single word reading improved by 9.5 months, whereas those receiving the placebo improved by only 3.3 months—a highly significant difference. Even though not all these children were classified as dyslexic, this effect of improving diet is most likely to be due to the effect of DHA improving the function of magnocellular timing neurons.

Dyslexia strengths

Although many people are convinced that people with dyslexia often have great talents (West, 2009), most studies about it concentrate on their reading problems. But there is now evidence that, alongside their magnocellular weaknesses, people with dyslexia may demonstrate superior parvocellular performance e.g., higher sensitivity than ordinary readers to high spatial but low temporal frequency modulated gratings (Lovegrove et al., 1982), better red and green color discrimination, particularly in the peripheral visual field (Dautrich, 1993). They are also faster and more accurate at identifying "impossible figures"- drawings of objects which cannot exist in reality, such as Escher's stairs and waterfalls (von Károlyi et al., 2003). Male dyslexics have also been shown to be particularly good at identifying shapes in ambiguous figures, remembering and reproducing designs and complex figures and at recreating and navigating in a virtual environment (Brunswick et al., 2010).

90% of the neurones that are born during early brain development fail to survive because they are eliminated in the ruthless competition to make useful connections. Hence the P-cell superiorities seen in dyslexics are not particularly surprising, since a reduced number of M-cells would leave room for more P-cells to survive and make successful connections. Probably therefore, their P- cell connectome is significantly more prolific than that of their ordinary reading peers. This outcome probably explains the very different cognitive style that characterizes dyslexic people. This style is often described as "being able to see the whole picture at once"—a "holistic" approach. Their unusual way of thinking would also account for why so many dyslexics are so remarkably successful in the arts (Wolff and Lundberg, 2002), at practical engineering (West, 1992) and in business (Logan, 2008). If they can survive their schooling which so often fails to recognize their talents, they can often achieve remarkable feats (Brock and Eide, 2012).

Conclusion

The evidence presented here strongly suggests that DD results from impaired development of the magnocellular timing connectome *in utero* and early childhood, particularly for vision and hearing. These cortical systems are responsible for sequencing letters and word sounds for learning to read. The large variety of genetic, histological, electrophysiological, imaging, psychophysical, dietary and achievement data reviewed here, all suggests that if these visual and auditory timing/transient systems fail to develop properly or are damaged, reading becomes difficult or impossible. Precise timing of when vision alights on each letter and the order in which the sounds of a spoken word are heard, is so crucial for reading that proper development of the reading networks is permanently compromised. This coherent body of evidence establishes the neural basis of reading, and strongly suggests, contrary to much current belief, that visual and auditory timing deficits are the main causes of the phonological deficit in people with dyslexia and their reading failure.

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

Funding

This article was kindly funded by a grant from the Shenzen Inst. of Neuroscience.

Acknowledgments

The author thank all my colleagues and students who have contributed to our greater understanding of reading networks and to all our funders over the years: Oxford University, the MRC, the Wellcome, Rodin Remediation, Esmee Fairbairn, BBC Children in Need, Tolkien, Garfield Weston and Dyslexia Research Trusts.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Adler Grinberg, D., and Stark, L. (1978). Eye movements, scanpaths and dyslexia. *Am. J. Optom. Physiol. Opt.* 55, 557–570. doi: 10.1097/00006324-197808000-00005

Bathelt, J., Gathercole, S. E., Butterfield, S., and Astle, D. E. (2017). The role of the structural connectome in literacy and numeracy development in children. *PsyArXiv* [Preprint]. doi: 10.31234/osf.io/jk6yb

Bavalier, D., and Davidson, R. (2013). Games to do you good. *Nature* 494, 425–426. doi: 10.1038/494425a

Berlin, R. (1884). "Über Dyslexie" [About dyslexia]. *Arch. Für. Psychiatr.* 15, 276–278.

Biscaldi, M., Fischer, B., and Hartnegg, K. (2000). Voluntary saccadic control in dyslexia. *Perception* 29, 509–521. doi: 10.1068/p2666a

Brannan, M. C., and Williams, J. (1988). The effects of age and reading ability on flicker threshold. *Clin. Vis. Sci.* 3, 137–142.

Britten, K., and Newsome, W. (1992). Comparison of neuronal and psychophysical motion detection. *J. Neurosci.* 12, 4745–4765. doi: 10.1523/JNEUROSCI.12-12-04745.1992

Brock, L., and Eide, F. F. E. (2012). *The Dyslexic Advantage: Unlocking the Hidden Potential of the Dyslexic Brain*. New York, NY: Hudson St Press.

Brunswick, N., Martin, G. N., and Marzano, L. (2010). Visuospatial superiority in developmental dyslexia: Myth or reality? *Learn. Individ. Differ.* 20, 421–426. doi: 10.1016/j.lindif.2010.04.007

Bucci, M. P., Bremond-Gignac, D., and Kapoula, Z. (2007). Poor binocular coordination of saccades in dyslexic children. *Graefes Arch. Clin. Exp. Ophthalmol.* 246, 417–428. doi: 10.1007/s00417-007-0723-1

Bullmore, E., and Sporns, O. (2012). The economy of brain network organization. *Nat. Rev. Neurosci.* 13, 336–349. doi: 10.1038/nrn3214

Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton, 113. doi: 10.1515/9783112316009

Clisby, C., Fowler, M., Hebb, G., Walters, J., Southcott, P., and Stein, J. (2000). Outcome of treatment of visual problems of children with reading difficulties. *Prof. Assoc. Teach. Spec Situations Bull.* 13, 9–14.

Cornelissen, P., Bradley, L., Fowler, S., and Stein, J. (1991). What children see affects how they spell. *Dev. Med. Child Neurol.* 33, 755–762. doi: 10.1111/j.1469-8749.1991.tb14959.x

Cornelissen, P., Richardson, A., Mason, A., Fowler, S., and Stein, J. (1995). Contrast sensitivity and coherent motion detection measured at photopic luminance levels in dyslexics and controls [see comments]. *Vision Res.* 35, 1483–1494. doi: 10.1016/0042-6989(95)98728-R

Dautrich, B. R. (1993). Visual perceptual differences in the dyslexic reader: Evidence of greater visual peripheral sensitivity to color and letter stimuli. *Percept. Mot. Skills.* 73, 755–764. doi: 10.2466/pms.1993.76.3.755

Demb, J. B., Boynton, G. M., Best, M., and Heeger, D. J. (1998). Psychophysical evidence for a magnocellular pathway deficit in dyslexia. *Vis. Res.* 38, 1555–1559. doi: 10.1016/S0042-6989(98)00075-3

Dick, A. S., Bernal, B., and Tremblay, P. (2013). The Language Connectome: New Pathways. *New Concepts. Neurosci.* 20, 453–467. doi: 10.1177/1073858413513502

Eden, G. F., Stein, J. F., Wood, H. M., and Wood, F. B. (1994). Differences in eye movements and reading problems in dyslexic and normal children. *Vision Res.* 34, 1345–1358. doi: 10.1016/0042-6989(94)90209-7

Eden, G. F., VanMeter, J. W., Rumsey, J. M., Maisog, J. M., Woods, R. P., and Zeffiro, T. A. (1996). Abnormal processing of visual motion in dyslexia revealed by functional brain imaging. *Nature* 382, 66–69. doi: 10.1038/382066a0

Elliott, J. G. (2020). It's time to be scientific about dyslexia. *Read. Res. Q.* 55:S61–S75. doi: 10.1002/rrq.333

Elliott, J. G., and Grigorenko, E. L. (2014). *The Dyslexia Debate*, Vol. 14. Cambridge: Cambridge University Press, 296. doi: 10.1017/CBO9781139017824

Facoetti, A., Paganoni, P., Turatto, M., Marzola, V., and Mascetti, G. G. (2000). Visual-spatial attention in developmental dyslexia. *Cortex* 36, 109–123. doi: 10.1016/S0010-9452(08)70840-2

Feng, X., Monzalvo, K., Dehaene, S., and Dehaene-Lambertz, G. (2022). Evolution of reading and face circuits during the first three years of reading acquisition. *Neuroimage* 259:119394. doi: 10.1016/j.neuroimage.2022.119394

Fischer, B., and Hartnegg, K. (2000). Stability of gaze control in dyslexia. *Strabismus* 8, 119–122. doi: 10.1076/0927-3972(200006)821-2FT119

Fowler, S., and Stein, J. F. (1979). New evidence for ambilaterality in visual dyslexia. *Neurosci. Letts.* 3:214.

Fowler, S., and Stein, J. F. (1983). Ocular motor dominance - an aetiological factor in some orthoptic problems. *BritOrthoptic. J.* 40, 43–45.

Franceschini, S., Gori, S., Ruffino, M., Pedrolli, K., and Facoetti, A. (2012). A causal link between visual spatial attention and reading acquisition. *Curr. Biol.* 22, 814–819. doi: 10.1016/j.cub.2012.03.013

Gabrieli, J., and Norton, E. (2012). Reading abilities: importance of visual-spatial attention. *Curr Biol.* 22:R298–R299. doi: 10.1016/j.cub.2012.03.041

Galaburda, A. M., Menard, M. T., and Rosen, G. D. (1994). Evidence for aberrant auditory anatomy in developmental dyslexia. *Proc. Natl. Acad. Sci. U.S.A.* 91, 8010–8013. doi: 10.1073/pnas.91.17.8010

Galaburda, A. M., Sherman, G. F., Rosen, G. D., Aboitiz, F., and Geschwind, N. (1985). Developmental dyslexia?: four consecutive patients with cortical anomahes. *Ann. Neurol.* 18, 222–233. doi: 10.1002/ana.410180210

Giraldo-Chica, M., Hegarty, J. P., and Schneider, K. A. (2015). Morphological differences in the lateral geniculate nucleus associated with dyslexia. *NeuroImage Clin.* 7, 830–836. doi: 10.1016/j.nicl.2015.03.011

Gori, S., Seitz, A. R., Ronconi, L., Franceschini, S., and Facoetti, A. (2016). Multiple causal links between magnocellular-dorsal pathway deficit and developmental dyslexia. *Cereb. Cortex* 26:bhv206. doi: 10.1093/cercor/bhv206

Goswami, U. (2014). The neural basis of dyslexia may originate in primary auditory cortex. *Brain* 137, 3100–3102. doi: 10.1093/brain/awu296

Goswami, U. (2015). Sensory theories of developmental dyslexia: Three challenges for research. *Nat. Rev. Neurosci.* 16, 43–54. doi: 10.1038/nrn3836

Haag, M. (2003). Essential fatty acids and the brain. *Can. J. Psychiatry* 48, 195–203. doi: 10.1177/070674370304800308

Hamalainen, J. A., Leppanen, P. H. T., Guttorm, T. K., and Lyytinen, H. (2008). Event-related potentials to pitch and rise time change in children with reading disabilities and typically reading children. *Clin. Neurophysiol.* 119, 100–115. doi: 10.1016/j.clinph.2007.09.064

Hämäläinen, J. A., Salminen, H. K., Leppänen, P. H. T., Hamalainen, J. A., Salminen, H. K., and Leppanen, P. H. T. (2012). Basic auditory processing deficits in dyslexia: systematic review of the behavioral and event-related potential/ field evidence. *J. Learn. Disabil.* 46, 413–427. doi: 10.1177/0022219411436213

Harrar, V., Tammam, J., Pérez-Bellido, A., Pitt, A., Stein, J., and Spence, C. (2014). Multisensory integration and attention in developmental dyslexia. *Curr. Biol.* 24, 531–535. doi: 10.1016/j.cub.2014.01.029

Harries, P., Hall, R., Ray, N., Stein, J., Griffiths, P. G., Harries, P., et al. (2015). Using coloured filters to reduce the symptoms of visual stress in children with reading delay. *Scand. J. Occup. Ther.* 22, 328–9. doi: 10.3109/11038128.2015.1033456

Hendry, S., Jones, E., Hockfield, S., and McKay, R. (1988). Neuronal populations stained with the monoclonal antibody Cat-301 in the mammalian cerebral cortex and thalamus. *J. Neurosci.* 8, 518–542. doi: 10.1523/JNEUROSCI.08-02-00518.1988

Hockfield, S., and McKay, R. D. (1983). A surface antigen expressed by a subset of neurons in the vertebrate central nervous system. *Proc. Natl. Acad. Sci. U.S.A.* 80, 5758–5761. doi: 10.1073/pnas.80.18.5758

Iles, J., Walsh, V., and Richardson, A. (2000). Visual search performance in dyslexia. *Dyslexia* 6, 163–177. doi: 10.1002/1099-0909(200007/09)6:3<163::AID-DYS150>3.0.CO;2-U

Jainta, S., and Kapoula, Z. (2011). Dyslexic children are confronted with unstable binocular fixation while reading. *PLoS One* 6:e18694. doi: 10.1371/journal.pone.0018694

Jednoróg, K., Marchewka, A., Tacikowski, P., Heim, S., Grabowska, A., Jednorog, K., et al. (2011). Electrophysiological evidence for the magnocellular-dorsal pathway deficit in dyslexia. *Dev. Sci.* 14, 873–880. doi: 10.1111/j.1467-7687.2011.01037.x

Johnston, R., Pitchford, N. J., Roach, N. W., and Ledgeway, T. (2017). New insights into the role of motion and form vision in neurodevelopmental disorders. *Neurosci. Biobehav. Rev.* 83, 32–45. doi: 10.1016/j.neubiorev.2017.09.031

Kirkby, J. A., Webster, L. A. D., Blythe, H. I., and Liversedge, S. P. (2008). Binocular coordination during reading and non-reading tasks. *Psychol. Bull.* 134, 742–763. doi: 10.1037/a0012979

Klistorner, A., Crewther, D. P., and Crewther, S. G. (1997). Separate magnocellular and parvocellular contributions from temporal analysis of the multifocal VEP. *Vision Res.* 37, 2161–2169. doi: 10.1016/S0042-6989(97)00003-5

Lawton, T. (2016). Improving dorsal stream function in dyslexics by training figure/ground motion discrimination improves attention, reading fluency, and working memory. *Front. Hum. Neurosci.* 10:397. doi: 10.3389/fnhum.2016.00397

Lawton, T., and Stein, J. F. (2022). Visual neural timing problems may interfere with reading, attention, and memory?: looking beyond 20/20 acuity. *Optom. Vis. Perform.* 10, 9–21.

Laycock, R., Crewther, D. P., and Crewther, S. G. (2008). The advantage in being magnocellular: A few more remarks on attention and the magnocellular system. *Neurosci. Biobehav. Rev.* 32, 1409–1415. doi: 10.1016/j.neubiorev.2008.04.008

Leppänen, P. H. T., Hämäläinen, J. A., Salminen, H. K., Eklund, K. M., Guttorm, T. K., Lohvansuu, K., et al. (2010). Newborn brain event-related potentials revealing atypical processing of sound frequency and the subsequent association with later literacy skills in children with familial dyslexia. *Cortex* 46, 1362–1376. doi: 10.1016/j.cortex.2010.06.003

Liberman, I. Y., Shankweiler, D., Orlando, C., Harris, K. S., and Berti, F. B. (1971). Letter confusions and reversals of sequence in the begining reader: implications for orton's theory of developmental dyslexia. *Cortex* 7, 127–142. doi: 10.1016/S0010-9452(71)80009-6

Liversedge, S. P., Rayner, K., White, S. J., Findlay, J. M., and McSorley, E. (2006). Binocular coordination of the eyes during reading. *Curr. Biol.* 16, 1726–1729. doi: 10.1016/j.cub.2006.07.051

Livingstone, M. S., Rosen, G. D., Drislane, F. W., and Galaburda, A. M. (1991). Physiological and anatomical evidence for a magnocellular defect in developmental dyslexia. *Proc. Natl. Acad. Sci. U.S.A.* 88, 7943–7947. doi: 10.1073/pnas.88.18.7943

Logan, J. (2008). Analysis of the Incidence of Dyslexia in Entrepreneurs and its Implications. Cass Business School. London: USASBE, 636–651.

Lovegrove, W., Martin, F., Bowling, A., Blackwood, M., Badcock, D., Paxton, S., et al. (1982). Contrast sensitivity functions and specific reading disability. *Neuropsychologia* 20, 309–315. doi: 10.1016/0028-3932(82)90105-1

Lovegrove, W. J., Heddle, M., and Slaghuis, W. (1980). Reading disability: Spatial frequency specific deficits in visual information store. *Neuropsychologia* 18, 111–115. doi: 10.1016/0028-3932(80)90093-7

Maguire, E. R., and Monsivais, P. (2015). Socio-economic dietary inequalities in UK adults: An updated picture of key food groups and nutrients from national surveillance data. *Br. J. Nutr.* 113, 181–189. doi: 10.1017/S0007114514002621

Mancarella, M., Antzaka, A., Bertoni, S., Facoetti, A., and Lallier, M. (2022). Enhanced disengagement of auditory attention and phonological skills in action video gamers. *Comput. Hum. Behav.* 135:107344. doi: 10.1016/j.chb.2022.107344

Manning, C., Kaneshiro, B., Kohler, P. J., Duta, M., Scerif, G., and Norcia, A. M. (2019). Neural dynamics underlying coherent motion perception in children and adults. *Dev. Cogn. Neurosci.* 38:100670. doi: 10.1016/j.dcn.2019.100670

Martelli, M., Di Filippo, G., Spinelli, D., and Zoccolotti, P. (2009). Crowding, reading, and developmental dyslexia. *J. Vis.* 9, 141–18. doi: 10.1167/9.4.14

Mascheretti, S., Gori, S., Trezzi, V., Ruffino, M., Facoetti, A., Marino, C., et al. (2018). Visual motion and rapid auditory processing are solid endophenotypes of developmental dyslexia. *Genes Brain Behav.* 17, 70–81. doi: 10.1111/gbb.12409

Mascheretti, S., Riva, V., Feng, B., Trezzi, V., Andreola, C., Giorda, R., et al. (2020). The mediation role of dynamic multisensory processing using molecular genetic data in dyslexia. *Brain Sci.* 10:993. doi: 10.3390/brainsci10120993

Maunsell, J. H., Ghose, G. M., Assad, J. A., McAdams, C. J., Boudreau, C. E., and Noerager, B. D. (1999). Visual response latencies of magnocellular and parvocellular neuorns in macaque monkeys. *Vis. Nsc.* 16, 1–14. doi: 10.1017/S0952523899156177

McAnally, K. I., and Stein, J. F. (1996). Auditory temporal coding in dyslexia. *Proc. RS.* 263, 961–965. doi: 10.1098/rspb.1996.0142

Meng, Q., and Schneider, K. A. (2022). A specialized channel for encoding auditory transients in the magnocellular division of the human medial geniculate nucleus. *Neuroreport.* 33, 663–668. doi: 10.1097/WNR.0000000000001830

Morais, J., Cary, L., Alegria, J., Bertelson, P., and Alegria, J. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition* 7, 323–331. doi: 10.1016/0010-0277(79)90020-9

Morgan, W. P. (1896). A case of congenital word blindness. *Br. Med. J.* 2:1378. doi: 10.1136/bmj.2.1871.1378

Mowforth, P., Mayhew, J. E., and Frisby, J. P. (1981). Vergence eye movements made in response to spatial-frequency-filtered random-dot stereograms. *Perception* 10, 299–304. doi: 10.1068/p100299

Muskiet, F. A. J., Fokkema, M. R., Schaafsma, A., Boersma, E. R., and Crawford, M. A. (2004). Is docosahexaenoic acid (DHA) essential? Lessons from DHA status regulation, our ancient diet, epidemiology and randomized controlled trials. *J. Nutr.* 134, 183–186. doi: 10.1093/jn/134.1.183

Nassi, J. J., and Callaway, E. M. (2009). Parallel processing strategies of the primate visual system. *Nat. Rev. Neurosci.* 10, 360–372. doi: 10.1038/nrn2619

Olson, R. K. (2006). Genes, environment, and dyslexia. The 2005 Norman Geschwind Memorial Lecture. *Ann. Dyslexia* 56, 205–238. doi: 10.1007/s11881-006-0010-6

Pammer, K., and Wheatley, C. (2001). Isolating the M(y)-cell response in dyslexia using the spatial frequency doubling illusion. *Vision Res.* 41, 2139–2148. doi: 10.1016/S0042-6989(01)00092-X

Peng, P., Wang, C., Tao, S., and Sun, C. (2017). The deficit profiles of chinese children with reading difficulties: a meta-analysis. *Educ. Psychol. Rev.* 29, 513–564. doi: 10.1007/s10648-016-9366-2

Peters, J. L., De Losa, L., Bavin, E. L., and Crewther, S. G. (2019). Efficacy of dynamic visuo-attentional interventions for reading in dyslexic and neurotypical children: A systematic review. *Neurosci. Biobehav. Rev.* 100, 58–76. doi: 10.1016/j.neubiorev.2019.02.015

Rauschecker, J. P. (2018). Where, when, and how: are they all sensorimotor? towards a unified view of the dorsal pathway in vision and audition. *Cortex* 98, 262–268. doi: 10.1016/j.cortex.2017.10.020

Ray, N. J., Fowler, S., and Stein, J. F. (2005). Yellow filters can improve magnocellular function: Motion sensitivity, convergence, accommodation, and reading. *Ann. N. Y. Acad. Sci.* 1039, 283–293. doi: 10.1196/annals.1325.027

Raymond, J. E., Ogden, N. A., Fagan, J. E., and Kaplan, B. J. (1988). Fixational instability and saccadic eye movements of dyslexic children with subtle cerebellar dysfunction. *Am. J. Optom. Physiol. Opt.* 65, 174–181. doi: 10.1097/00006324-198803000-00006

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychol. Bull.* 124, 372–422. doi: 10.1037/0033-2909.124.3.372

Regtvoort, A. G. F. M., van Leeuwen, T. H., Stoel, R. D., and van der Leij, A. (2006). Efficiency of visual information processing in children at-risk for dyslexia: Habituation of single-trial ERPs. *Brain Lang.* 98, 319–331. doi: 10.1016/j.bandl.2006.06.006

Richardson, A. J., and Montgomery, P. (2005). The oxford-durham study: a randomized, controlled trial of dietary supplementation with fatty acids in children with developmental coordination disorder. *Pediatrics* 115, 1360–1366. doi: 10.1542/peds.2004-2164

Rosli, Y., Bedford, S. M., and Maddess, T. (2009). Low-spatial-frequency channels and the spatial frequency-doubling illusion. *Investig. Ophthalmol. Vis. Sci.* 50, 1956–1963. doi: 10.1167/iovs.08-1810

Schulte-Körne, G., and Bruder, J. (2010). Clinical neurophysiology of visual and auditory processing in dyslexia: A review. *Clin. Neurophysiol.* 121, 1794–1809. doi: 10.1016/j.clinph.2010.04.028

Schulte-Körne, G., Deimel, W., Bartling, J., Remschmidt, H., Schulte-Korne Deimel, W., Bartling, J., et al. (1998). Auditory processing and dyslexia: Evidence for a specific speech processing deficit. *Neuroreport* 9, 337–340. doi: 10.1097/00001756-199801260-00029

Singleton, C., and Trotter, S. (2005). Visual stress in adults with and without dyslexia. *J. Res. Read.* 28, 365–378. doi: 10.1111/j.1467-9817.2005.00275.x

Skottun, B. C., and Skoyles, J. (2007). A few remarks on assessing magnocellular sensitivity in Schizophrenic patients. *Brain* 130:e83. doi: 10.1093/brain/awm153

Skottun, B. C., and Skoyles, J. R. (2004). Some remarks on the use of motion VEPs to assess magnocellular sensitivity. *Clin. Neurophysiol.* 115, 2834–2838. doi: 10.1016/j.clinph.2004.06.021

Spitschan, M. (2019). Melanopsin contributions to non-visual and visual function. *Curr. Opin. Behav. Sci.* 30, 67–72. doi: 10.1016/j.cobeha.2019.06.004

Stein, J. (2019). The current status of the magnocellular theory of developmental dyslexia. *Neuropsychologia* 130, 66–77. doi: 10.1016/j.neuropsychologia.2018.03.022

Stein, J. (2021). Reduced visual magnocellular event-related potentials in developmental dyslexia. *Brain Sci.* 11:48. doi: 10.3390/brainsci11010048

Stoodley, C. J., Talcott, J. B., Carter, E. L., Witton, C., and Stein, J. F. (2000). Selective deficits of vibrotactile sensitivity in dyslexic readers. *Neurosci. Lett.* 295, 13–16. doi: 10.1016/S0304-3940(00)01574-3

Talcott, J., Witton, C., Hebb, G., Stoodley, C., Westwood, E., France, S., et al. (2002). On the relationship between dynamic visual and auditory processing and literacy skills; results from a large primary-school study. *Dyslexia* 8, 204–225. doi: 10.1002/dys.224

Talcott, J. B., Hansen, P. C., Willis-Owen, C., McKinnell, I. W., Richardson, A. J., Stein, P. J. F., et al. (1998). Visual magnocellular impairment in adult developmental dyslexics. *Neuro. Ophthalmol.* 20, 187–201. doi: 10.1076/noph.20.4.187.3931

Thomson, M. E. (1982). The assessment of children with specific reading difficulties (dyslexia) using the British Ability Scales. *Br. J. Psychol.* 73, 461–478. doi: 10.1111/j.2044-8295.1982.tb01828.x

van der Leij, A. (2013). Dyslexia and early intervention: what did we learn from the dutch dyslexia programme? *Dyslexia* 19, 241–255. doi: 10.1002/dys.1466

Vidyasagar, T. R. (2005). Attentional gating in primary visual cortex: A physiological basis for dyslexia. *Perception* 34, 903–911. doi: 10.1068/p5332

Vidyasagar, T. R., and Pammer, K. (1999). Impaired visual search in dyslexia relates to the role of the magnocellular pathway in attention. *Neuroreport* 10, 1283–1287. doi: 10.1097/00001756-199904260-00024

Vidyasagar, T. R., and Pammer, K. (2010). Dyslexia: A deficit in visuo-spatial attention, not in phonological processing. *Trends Cogn. Sci.* 14, 57–63. doi: 10.1016/j.tics.2009.12.003

von Károlyi, C., Winner, E., Gray, W., and Sherman, G. F. (2003). Dyslexia linked to talent: Global visual-spatial ability. *Brain Lang.* 85, 427–431. doi: 10.1016/S0093-934X(03)00052-X

West, T. G. (1992). A future of reversals: dyslexic talents in a world of computer visualization. *Ann. Dyslexia* 42, 124–139. doi: 10.1007/BF02654942

West, T. G. T. (2009). *In The Mind's Eye: Creative Visual Thinkers, Gifted Dyslexics, and The Rise of Visual Technologies.* Buffalo, NY: Prometheus Books.

Whitford, V., O'Driscoll, G. A., and Titone, D. (2017). Reading deficits in schizophrenia and their relationship to developmental dyslexia: A review. *Schizophr. Res.* 193, 11–22. doi: 10.1016/j.schres.2017.06.049

Witton, C., Talcott, J. B., Hansen, P. C., Richardson, A. J., Griffiths, T. D., Rees, A., et al. (1998). Sensitivity to dynamic auditory and visual stimuli predicts nonword reading ability in both dyslexic and normal readers. *Curr. Biol.* 8, 791–797. doi: 10.1016/S0960-9822(98)70320-3

Wolff, U., and Lundberg, I. (2002). The prevalence of dyslexia among art students. *Dyslexia* 8, 34–42. doi: 10.1002/dys.211

Zhou, W., Cui, X., Shi, B., Su, M., and Cao, M. (2021). The development of brain functional connectome during text reading. *Dev. Cogn. Neurosci.* 48:100927. doi: 10.1016/j.dcn.2021.100927

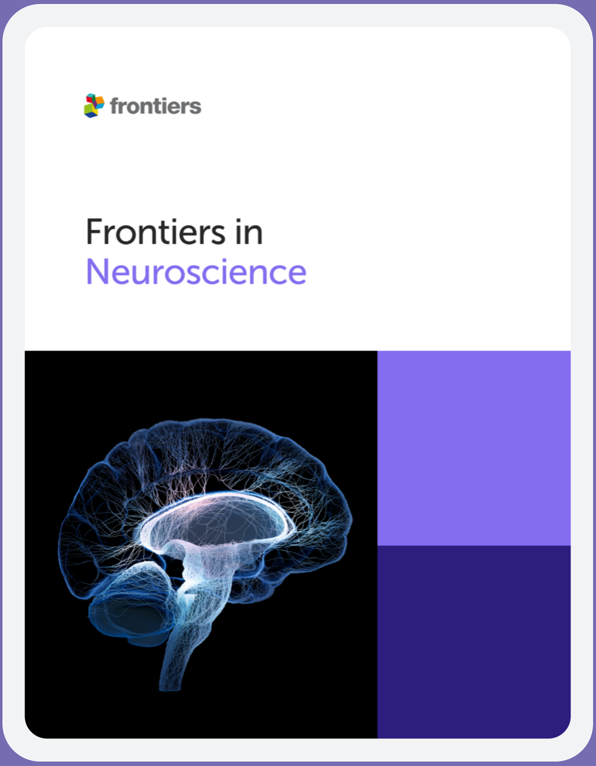
Frontiers in
Neuroscience

Provides a holistic understanding of brain function from genes to behavior

Part of the most cited neuroscience journal series which explores the brain - from the new eras of causation and anatomical neurosciences to neuroeconomics and neuroenergetics.

Discover the latest Research Topics

See more →


frontiers

Frontiers in
Neuroscience

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact