OXFORD

Cerebral Cortex, 2025, **35**, bhaf193
https://doi.org/10.1093/cercor/bhaf193
Advance access publication date 6 October 2025
Original Article

Multimodal investigation of the neurocognitive deficits underlying dyslexia in adulthood

Cristina Cara[1,2], Giulia Zantonello[1,3], Marta Ghio[2], and Marco Tettamanti[4,*]

[1]CIMeC—Center for Mind/Brain Sciences, University of Trento, Corso Bettini 31, 38068 Rovereto (TN), Italy
[2]Institute of Experimental Psychology, Heinrich Heine University, Universitätsstrasse 1, 40225 Düsseldorf, Germany
[3]Department of Educational Psychology, University of Göttingen, Waldweg 26, 37073 Göttingen, Germany
[4]Department of Psychology, University of Milano-Bicocca, Piazza dell'Ateneo Nuovo 1, 20126 Milano, Italy
[*]Corresponding author: Cristina Cara, Department of Psychology, University of Milano-Bicocca, Piazza dell'Ateneo Nuovo 1, I-20126 Milano, Italy.
Email: marcodante.tettamanti@unimib.it

Dyslexia is a neurobiological disorder characterized by reading difficulties, yet its causes remain unclear. Neuroimaging and behavioral studies found anomalous responses in tasks requiring phonological processing, motion perception, and implicit learning, and showed gray and white matter abnormalities in dyslexics compared to controls, indicating that dyslexia is highly heterogeneous and promoting a multifactorial approach. To evaluate whether combining behavioral and multimodal MRI improves sensitivity in identifying dyslexia neurocognitive traits compared to monocomponential approaches, 19 dyslexic and 19 control subjects underwent cognitive assessments, multiple (phonological, visual motion, rhythmic) mismatch-response functional MRI tasks, structural diffusion-weighted imaging (DWI) and T1-weighted imaging. Between group differences in the neurocognitive measures were tested with univariate and multivariate approaches. Results showed that dyslexics performed worse than controls in phonological tasks and presented reduced cerebellar responses to mismatching rhythmic stimuli, as well as structural disorganization in white matter tracts and cortical regions. Most importantly, a machine learning model trained with features from all three MRI modalities discriminated between dyslexics and controls with greater accuracy than single-modality models. The individual classification scores in the multimodal machine learning model correlated with behavioral reading accuracy. These results characterize dyslexia as a composite condition with multiple distinctive cognitive and brain traits.

Keywords: cerebellum; dyslexia; machine learning; mismatch response; multimodal MRI.

Introduction

Developmental dyslexia is a specific learning disability characterized by persistent difficulties in reading efficiently despite normal intelligence and adequate educational opportunities (Peterson and Pennington 2012). Dyslexia represents one of the most common neurobiological disorders, affecting approximately 5%–17% of school-aged children (Shaywitz and Shaywitz 2005). The neurobiology of dyslexia has been widely investigated in the last decades (Paulesu et al. 2014; Norton et al. 2015). Magnetic resonance imaging studies point out gray matter volume (GMV) and cortical thickness (CT) alterations in occipito-temporal, temporo-parietal, and inferior frontal cortices (Ramus et al. 2018), as well as in cerebellar regions (Pernet et al. 2009). Diffusion tensor imaging (DTI) studies further revealed white matter (WM) disorganization in dyslexic subjects affecting association tracts as well as projection and commissural tracts (Vandermosten et al. 2012b; Cui et al. 2016). Functional neuroimaging studies have highlighted the presence of widespread abnormalities in the dyslexic brain, particularly in temporo-parietal, occipito-temporal and inferior frontal cortices (Norton et al. 2015; Peterson and Pennington 2015; Devoto et al. 2022). The involvement of this extensive network is thought to reflect the multifaceted nature of the cognitive deficits underlying dyslexia (Menghini et al. 2010). Indeed, besides the defining symptomatology of the reading disorder, dyslexic subjects may exhibit impairments in manifold cognitive domains. This evidence has given rise to various theories about the etiology of dyslexia, with yet little consensus among them (Carioti et al. 2022). Some of the most influential views on the disorder so far stem from the phonological, magnocellular, and cerebellar theories, which posit that dyslexia is caused by deficits in phonological processing (Ramus 2001; Peterson and Pennington 2015; Snowling et al. 2020), visuo-motion perception (Fischer and Hartnegg 2000; Bucci et al. 2008; Stein 2014; Gori and Facoetti 2015; Gori et al. 2016; Tiadi et al. 2016; Stein 2019), and implicit learning of rhythmic sequences (Vicari et al. 2003; Menghini et al. 2006), respectively. Although the different theoretical accounts have significantly deepened our understanding of dyslexia etiology, there is now growing consensus that none of these theories alone can fully account for the etiology of the disorder. The multiple deficit model (McGrath et al. 2020) has been proposed as a framework for understanding neurodevelopmental disorders, including dyslexia. The model emerged in response to the limitations of single-deficit accounts, providing an explanation that incorporates multiple interacting risk factors, ranging from genetic and neural influences to cognitive and behavioral symptoms, that collectively

Received: December 19, 2024. **Revised:** June 20, 2025. **Accepted:** June 24, 2025

contribute to the heterogeneity observed in dyslexia. In line with this view, dyslexia is increasingly recognized as a multifactorial phenomenon, involving a complex interplay of structural, functional brain, and behavioral impairment that collectively contribute to reading difficulties (Peterson and Pennington 2015). One of the most appropriate methods for examining multidimensionality involves the use of multivariate machine learning (ML) models (Vu et al. 2018; Nenning and Langs 2022). The recent emergence of ML techniques has yielded substantial benefits in clinical and research contexts (Vogt 2018; Vu et al. 2018). ML studies have been conducted to predict dyslexia based on various data sources, including psycho-educational tests, web-based games, eye movement tracking, and neuroimaging measurements (Kaisar 2020; Usman et al. 2021; Nakai et al. 2024). Specifically, GMV and CT have been used to discriminate between dyslexic and control groups (Tamboer et al. 2016; Płoński et al. 2017) and to predict children's future literacy skills (Tamboer et al. 2016; Beyer et al. 2022). Similar results were also obtained using WM diffusion indices (Hoeft et al. 2011; Cui et al. 2016; Langer et al. 2017) and functional magnetic resonance imaging (fMRI) data (Hoeft et al. 2011; Vandermosten et al. 2020; Yu et al. 2020; Zahia et al. 2020). While previous studies have used multimodal neuroimaging approaches to investigate dyslexia (Borghesani et al. 2021; Nemmi et al. 2023), and others have examined the intersection of phonological, magnocellular, and cerebellar processing systems in typical readers (Danelli et al. 2013), to our knowledge, no published studies have integrated structural gray and WM data with fMRI measures within a single predictive model. Moreover, we are not aware of prior studies that have simultaneously tested the three single-domain theoretical accounts of dyslexia (i.e. the phonological, magnocellular, and cerebellar theories) within such an integrated framework. The multimodal approach proposed in the current research allows the independent assessment of each theoretical account at both structural and functional levels, while also exploring the potential interactions between them.

Aim of the study

Within this framework, the current study aimed to investigate the neurocognitive bases of dyslexia in a group of dyslexic adults compared to a group of matched controls by combining behavioral with functional and structural brain imaging measures. We developed an integrated approach to test the phonological, magnocellular, and cerebellar theories of dyslexia at the cognitive and brain levels within the same experimental paradigm. In addition to behavioral and mismatch response (MMR) fMRI tasks testing the participants' sensitivity to the phonological, magnocellular, and cerebellar domains, T1-weighted and diffusion-weighted images were also acquired. Between-group differences for each of the cognitive and neuroimaging measures were tested employing univariate and multivariate approaches. Our prediction was that individuals with dyslexia would perform worse than controls in all the cognitive domains and show reduced MMR in all three fMRI tasks. Significant differences between the two groups were also expected for structural gray and WM indices. Embracing the view of dyslexia as a multifactorial phenomenon, the main aim of the current study was to implement an ML algorithm to evaluate whether the combination of multimodal brain measures could accurately discriminate between subjects with and without dyslexia. We expected the ML model including functional and structural brain measures to outperform both univariate and multivariate models that only include unimodal measures, offering a more comprehensive understanding of the neurocognitive bases of the reading disorder.

Materials and methods

Participants

Nineteen subjects with a diagnosis of dyslexia (10 males, mean age = 23.67 yr, Standard Deviation (SD) = 4.4 yr) and 19 control subjects (7 males, mean age = 23.55 yr, SD = 2.5 yr) participated in the present study. Inclusion criteria for the dyslexic group were: (i) an official diagnosis of dyslexia; (ii) text reading speed and/or accuracy score equal to or below the fifth percentile of the normative data in the MT-16-19 test battery (Cornoldi and Candela 2015); (iii) abstract reasoning skills in the normal range; (iv) no diagnosis of primary developmental language disorder. Inclusion criteria for the control group were: (i) no history or familiarity for developmental learning disabilities or developmental language disorders; (ii) reading speed and accuracy scores equal to or above the 10th percentile of the MT-16-19 normative data (Cornoldi and Candela 2015); (iii) abstract reasoning skills in the normal range. Two dyslexic participants did not provide their official diagnosis certificates; however, one was referred by a local dyslexia association, and the other had reading accuracy below the fifth percentile of normative data.

The two groups were not significantly different for sociodemographic and educational variables (Table 1). Thirteen out of the 19 dyslexic participants reported the presence of comorbidities with other learning disorders (Table 1). All participants were right-handed, Italian native-speakers with no reported history of neurological or psychiatric conditions, brain trauma, substance abuse, or pharmacological treatment affecting the nervous system, and no sensory or perceptual deficits except for corrected-to-normal vision. Participants gave written consent to participate in the study after receiving an explanation of the procedures. The study was approved by the Ethical Committee of the University of Trento, Italy.

Experimental design

Cognitive tests

The short version of the Raven's Advanced Progressive Matrices (RAPM; Arthur et al. 1999; Arthur and Day 1994) was administered to assess participants' abstract reasoning and general intelligence. The Italian standardized MT-16-19 reading battery (Cornoldi and Candela 2015) was used to assess text reading speed and accuracy. To examine the cognitive abilities underlying the three alternative theories of dyslexia, we administered phonological, magnocellular, and cerebellar tests. Phonological abilities were assessed using the standardized VALS Italian test battery (George and Pech-Georgel 2017), including the following subtests: phonological awareness (PA) segmentation task, PA fusion task, rapid automatized naming (RAN) task, immediate and delayed recall, digit span forward and backward (see Supplementary Methods S1.1, for a detailed description of the phonological subtests). The coherence motion perception (CMP) task was used for testing processing in the magnocellular system (Boets et al. 2006; Kevan and Pammer 2008; Benassi et al. 2010; Joo et al. 2017). In this task, a kinematogram included a certain proportion of dots moving coherently leftwards or rightwards within the rectangular field, while the remaining dots moved in random directions. Participants were asked to indicate the prevalent direction of the dots' movement by pressing either the left or the right arrow on a computer keyboard. The individual motion coherence threshold was defined as the smallest fraction of coherently moving dots required for correct direction discrimination (see Supplementary Methods S1.2, for a detailed description of the

Table 1. Participants' sociodemographic and educational information.

	Dyslexics	Controls	Statistical test	*P*-value
Number of participants	19	19	–	–
Sex number of Male/Female	10/9	7/12	$\chi^2 = 0.51$	0.50
Age years—mean (SD)	23.67 (4.40)	23.55 (2.46)	t = 0.11	0.91
Education years—mean (SD)	15.00 (2.69)	16.05 (1.68)	t = 1.45	0.16
Parental education years—mean (SD)	15.53 (4.12)	13.39 (2.50)	t = 1.93	0.06
Comorbidities	Dysgraphia: $n = 3$ Disorthography: $n = 6$ Dyscalculia: $n = 10$	–		–

CMP task). Participants' implicit learning skills related to cerebellar activity were tested using a serial reaction time (SRT) task (Vicari et al. 2003; Menghini et al. 2006; Yang et al. 2013). Participants were presented with red, blue, and green colored circles (radius = 100 pixels) that appeared sequentially on a computer screen, and they were asked to press the space bar on the computer keyboard whenever they saw a green circle. The entire set of stimuli was organized into five different blocks of 70 stimuli each, with the first block (I) characterized by a pseudo-random stimulus presentation order, followed by three blocks (II to IV) in which a fixed sequence of presentation was implemented, followed by a final block (V) with again pseudo-random order presentation (see Supplementary Methods S1.3, for a detailed description of the SRT task). Reaction times and accuracy in response to the green circles were measured. Both the CMP and SRT tasks were created and presented using Psychopy (version 2021.1.2; Peirce et al. 2019). The participants sat at a distance of approximately 60 cm from the screen.

Functional Magnetic Resonance Imaging tasks

We designed three fMRI tasks based on an oddball paradigm. In magneto/electroencephalographic studies, mismatch negativity is an effect elicited by the occurrence of an "oddball" (deviant) stimulus in a repetitive stream of frequent (standard) stimuli, reflecting the individual brain's sensitivity to perceive and discriminate stimuli (Kujala and Näätänen 2001; Näätänen et al. 2007; Garrido et al. 2009; Fitzgerald and Todd 2020; Gu and Bi 2020). Several studies provided evidence that fMRI captures an equivalent MMR both in auditory and visual modalities (Schall et al. 2003; Gomot et al. 2006; Kovarski et al. 2021).

Three distinct experimental MMR fMRI tasks were designed for the current study, each associated with a matched control MMR fMRI task: (i) an experimental phonological task matched to a control auditory task; (ii) an experimental magnocellular task matched to a control parvocellular task; and (iii) an experimental cerebellar task matched to a control color task. All MMR tasks were presented using Psychopy 2021.1.2, and conformed to a block design, with four standard blocks alternated with four deviant blocks. The standard block contained only standard stimuli ($n = 26$), while the deviant block comprised 21 standard stimuli (81%) and 5 deviant stimuli (19%) in a pseudo-randomized order. All experimental and control MMR fMRI tasks conformed to an identical design and timing parameters (i.e. number and timing of stimuli, block structure), with the only difference being the nature of the stimuli used (see Supplementary Methods S2 for further details).

Phonological Mismatch Response functional Magnetic Resonance Imaging task

The experimental phonological MMR task featured the */ba/* and */da/* syllables as the standard and deviant stimuli, respectively, to elicit a phonological contrast (Alonso-Búa et al. 2006; Gu and Bi 2020). The syllable stimuli were produced by a female Italian native speaker and recorded in an anechoic room with a condenser microphone. The control auditory MMR task featured non-linguistic tone stimuli generated in MATLAB R2020b. The standard and deviant nonlinguistic tone stimuli had the same fundamental frequencies as the F1 and F2 formants of, respectively, the */ba/* and */da/* syllables. All stimuli lasted 300 ms and were separated by a 700 ms-long Inter-Stimulus Interval (ISI) (Fig. 1A for the experimental task; Fig. S1A for the control task).

Magnocellular Mismatch Response functional Magnetic Resonance Imaging task

The experimental magnocellular MMR task employed horizontal sinusoidal gratings with spatial frequency of 0.1 cycles per degree, a contrast level of 0.1, and a velocity of 50 degrees per second (Kremláček et al. 2016). For the control MMR parvocellular task, we used horizontal sinusoidal gratings with higher spatial frequency (1 cycle per degree) and contrast levels (1.0), along with a lower temporal frequency (velocity = 5 degrees per second). The standard gratings featured an upward-downward motion pattern (150 ms upward followed by 150 ms downward), while the deviant stimuli displayed a reverse down-up motion sequence (150 ms downward, followed by 150 ms upward). All stimuli measured 10x15 degrees, lasted 300 ms, and were separated by a 700 ms ISI (Fig. 1B for the experimental task; Fig. S1B for the control task).

Cerebellar Mismatch Response functional Magnetic Resonance Imaging task

The experimental, cerebellar MMR task featured a blue circle (RGB values = 0, 50, 255; radius = 100 pixels) presented in rhythmic sequences. Standard sequences were constructed with a consistent rhythmic pattern featuring stimulus durations of 200 ms, 100 ms, and 400 ms, separated by a constant 100 ms ISI. The deviant sequences were generated by interchanging the second and last stimulus durations within the rhythmic pattern, resulting in durations of 200 ms, 400 ms, and 100 ms, while still maintaining a constant 100 ms ISI (Fig. 1C for the experimental task). The control color MMR task featured a nonrhythmic presentation of the blue circle used in the experimental task as the standard stimulus, and of a light blue circle (RGB values = 0, 100, 255; radius = 100 pixels) as the deviant stimulus. Both standard and deviant stimuli in the control task were presented for 700 ms, separated by a 300 ms-long ISI (Fig. S1C for the control task).

Experimental procedure

All participants first underwent the MRI data collection, followed by the behavioral testing session. Participants were provided with MRI-compatible earplugs for listening to the auditory stimuli while being isolated from the scan noise. Every participant underwent six fMRI acquisition runs, with each run including either one of the three experimental or one of the three control MMR fMRI tasks. The order of task presentation across the six fMRI runs was

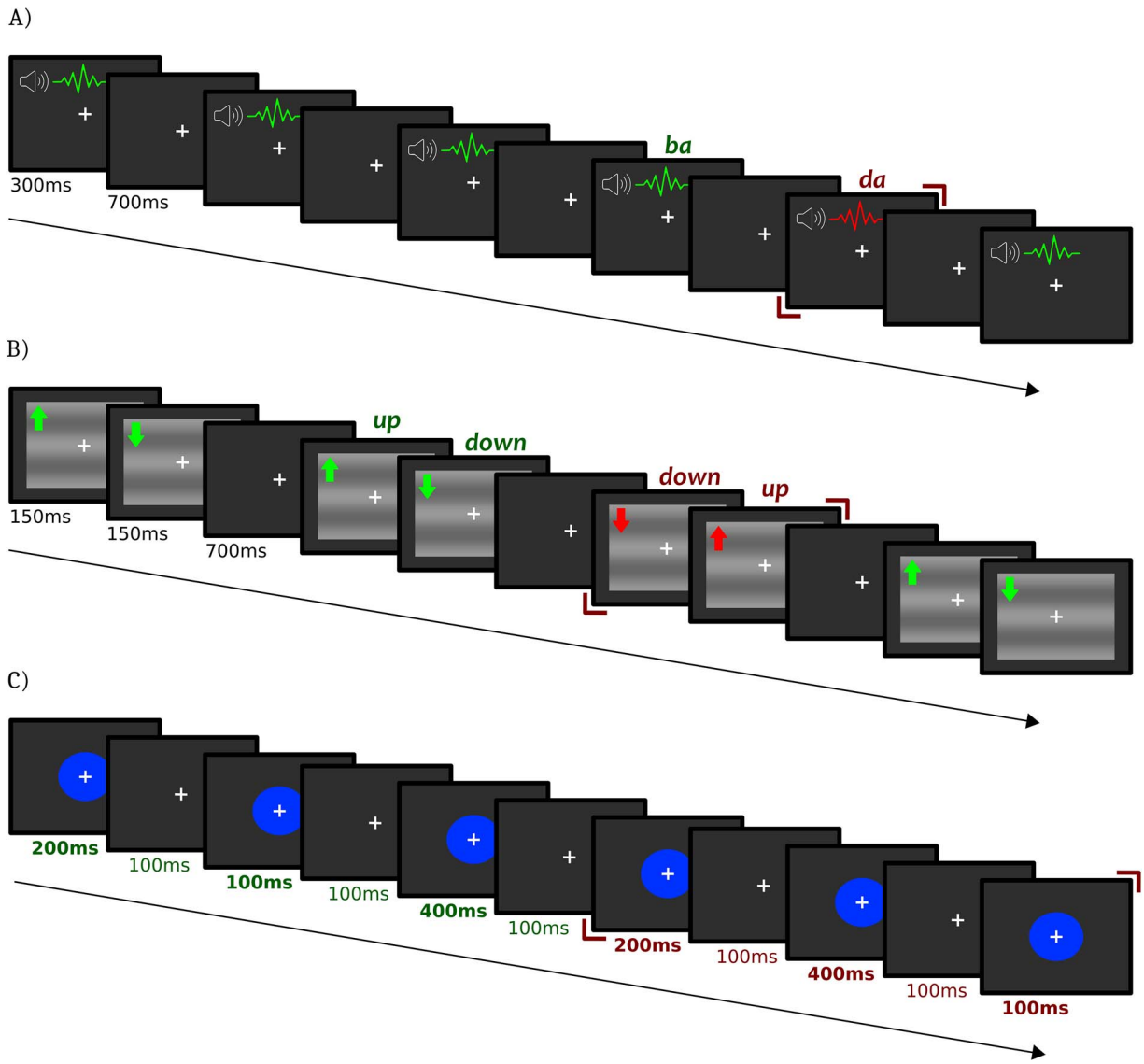

Fig. 1. Mismatch-response (MMR) fMRI task design. Visual representation of standard (green) and deviant (red) stimuli for the phonological, magnocellular and cerebellar experimental MMR fMRI tasks. (A) For the phonological task we used the standard */ba/* and */da/* syllables, with all stimuli lasting 300 ms with an ISI of 700 ms. (B) The magnocellular task included a standard grating with an upward-downward motion and a deviant grating with a downward-upward motion, all stimuli lasting 300 ms (150 ms in one direction and 150 ms in the opposite direction), with an ISI of 700 ms. For illustration purposes only, the spatial frequency and contrast of the stimuli depicted in this figure have been slightly increased (spatial frequency = 0.2, contrast = 0.2) compared to those used in the actual experiment (spatial frequency = 0.1, contrast = 0.1). (C) The cerebellar task involved a visual rhythmic pattern, with the standard rhythm generated by presenting a blue circle on the screen for 200–100–400 ms. For the deviant rhythm the second and last stimulus duration were reversed (i.e. 200–400–100 ms); the ISI was 100 ms.

counter-balanced across participants, except for the experimental phonological and its matched control auditory tasks, which were always presented in the first two runs to minimize earplugs displacement during fMRI data acquisition. At the end of the fMRI session, structural T1 and DWI were acquired.

Neuroimaging data acquisition

MRI scans were acquired at CIMeC, University of Trento, Italy, with a 3-Tesla MRI scanner (MAGNETOM Prisma, Siemens Healthcare, Erlangen, Germany), using a head–neck radio frequency receive coil with 64 channels. Participants first underwent six runs of functional whole-brain image acquisition. Functional images were acquired with a T2*-weighted gradient-echo, simultaneous multislice echoplanar imaging (EPI), and using blood oxygenation-level-dependent (BOLD) contrast (repetition time [TR] = 1,000 ms, echo time [TE] = 28.0 ms, 2 mm-isotropic voxels; multi-band acceleration factor = 5). Each run comprised 245 volumes, preceded by 10 dummy scans excluded from subsequent analysis. Each image volume consisted of 65 contiguous axial slices (2 mm thick) acquired in a bottom-to-top interleaved sequence (field of view = 200 mm, in-plane matrix size = 100 x 100). Each run was followed by the acquisition of two echo-planar spin echo sequences (TR = 4010 ms; TE = 28 ms; 2 mm isotropic voxels; echo spacing = 0.57 ms; Bandwidth 2272 Hz/Px) with reverse phase encoding directions (i.e. anterior–posterior and posterior–anterior, respectively), which served for geometrical distortion correction of functional images.

High-resolution structural 3D T1-weighted MEMPRAGE images (four echoes and one root-mean-square image) were subsequently acquired (TR = 3280 ms, TE's = [2.01, 4.04, 6.07, 8.1 ms], 0.7 mm-isotropic voxels, 288 sagittal slices).

Diffusion-weighted data were acquired using a pulse gradient spin echo sequence with the following parameters: resolution = 2 mm isotropic voxels; TR = 4,900 ms; TE = 75 ms;

shells: b = 0, b = 1000, b = 2000 s/mm^2, with 12/32/64 directions, respectively, uniformly distributed on a semisphere; anterior–posterior phase encoding direction. This resulted in a total of 108 volumes with 12 b = 0 volumes equally interspersed in the acquisition series. For distortion correction, another sequence including 6 b = 0 volumes and with the same parameters but with reversed phase encoding direction (i.e. posterior–anterior) was acquired.

Behavioral data analyses

Independent two-sample t-tests were planned to assess significant differences between dyslexic and control subjects for the following behavioral measures: accuracy scores in the RAPM; number of reading errors and reading speed in the Italian MT-16-19 reading battery; accuracy scores for all the phonological subtests of the VALS battery, time required to perform the PA segmentation, PA fusion, and RAN tasks; CMP threshold; and reaction times in the SRT. The alpha threshold was set at 0.05, and Bonferroni multiple testing correction was applied.

Neuroimaging data preprocessing and analyses

Functional Magnetic Resonance Imaging data

Preprocessing

Brain image preprocessing was performed using SPM12 version 7771 (https://www.fil.ion.ucl.ac.uk/spm/) running on MATLAB R2020b. Structural T1-weighted images were segmented and registered to the Montreal Neurological Institute (MNI) standard space. Functional images were visually inspected for head motion artifacts, and eventual despiking of artifactual volumes by the mean of neighboring non-artifactual volumes: no more than 5% volumes per run were removed for each subject. The preprocessing pipeline for functional images further included spatial realignment, distortion correction, and slice timing. Functional images were then normalized to the MNI space using the segmentation procedure with the subject-specific segmented structural images as customized segmentation priors. A spatial smoothing with a 4 mm full-width half maximum (FWHM) Gaussian kernel was applied for the univariate fMRI analysis but not for the multivariate analyses.

First level analysis

General linear models (GLMs) were specified for each subject with the preprocessed and smoothed functional images, with the time series high-pass filtered at 128 s and globally normalized. We specified three different GLMs. Each model included one experimental MMR fMRI task along with the associated control MMR fMRI task. Each task (either experimental or control task) was modeled with two regressors of interest (standard or deviant blocks), one nuisance regressor (including task instructions and the distractor blocks; see Methods S2), and the six head movement realignment parameters. Despiked volumes were also modeled as confound regressors. Within each of the three estimated first-level GLMs, we specified the following t-contrasts: (i) the simple MMR effect (deviant blocks > standard blocks) was calculated separately for the experimental task and the control task, by specifying a weight of +1 for the deviant condition and a weight of −1 for the standard condition; (ii) the interaction (Task [experimental > control] x MMR [deviant blocks > standard blocks]), testing for stronger MMR in the experimental task compared to the control task (weights: [−1 1 1 −1]; order: [experimental-standard, experimental-deviant, control-standard, control-deviant]).

Second-level region of interest analysis

The simple main effect and interaction contrasts specified at the first level were used to specify a set of second-level, random-effects analyses testing for between-group differences in each task. The first-level simple main effect contrast images were used to specify six second-level, two-sample t-test models, investigating whether dyslexic subjects showed altered MMR compared to control subjects in each of the three experimental (phonological, magnocellular, cerebellar) tasks as well as in each of the three control (auditory, parvocellular, color) tasks. In turn, the first-level interaction contrast images were used to specify three two-sample t-test models, testing for the mixed three-way interaction between the factors Group (dyslexic > controls), Task (experimental > control), and MMR (deviant blocks > standard blocks) in each task. The significance threshold for these analyses was set at peak-level $P < 0.05$, family-wise error (FWE) corrected for multiple comparisons. For each experimental task, we defined a task-specific region of interest (ROI) brain mask that was used to constrain the anatomical space for the second-level analyses, based on literature-driven hypotheses (Menghini et al. 2006; Danelli et al. 2013; Devoto et al. 2022; Fig. 2A, Supplementary Tables S1A, S1B, S1C).

T1-weighted image

Preprocessing

T1-weighted images were preprocessed and analyzed using FreeSurfer, version 7.2.0 (Fischl 2012). The main steps of this process included: intensity normalization, registration to Talairach atlas, skull-stripping, WM segmentation, tessellation of boundaries between gray and WM, deformation to pial surface, automatic topological defects correction (Dale et al. 1999; Fischl et al. 1999). Following the preprocessing, the resultant surfaces were visually inspected to ensure accurate segmentation. CT and GMV measures were computed at each vertex using the standard FreeSurfer procedure (Fischl and Dale 2000; Goto et al. 2022). For univariate analyses only, the individual maps of CT and GMV were smoothed using a FWHM of 10 mm. In addition to the surface-based measures, FreeSurfer also provided segmentation of subcortical regions by assigning labels to each voxel based on probabilistic information. Volumetric measures of the left and right cerebellar cortex were extracted (Fischl et al. 2002).

Cortical thickness and gray matter volume region of interest analysis

Alterations in CT and GMV related to dyslexia diagnosis were investigated with vertex-wise two-sample t-tests comparing dyslexic and control subjects, with the total intracranial volume added as a nuisance covariate. Analyses were conducted by fitting a GLM using the FreeSurfer "mri_glmfit" tool. The resulting maps were corrected for multiple comparisons using cluster-based Monte Carlo simulation with 10,000 iterations of randomly generated z-score maps (cluster-wise alpha = 0.05; Hagler et al. 2006). To constrain the anatomical space for the analyses of CT and GMV, we defined 20 ROIs (Fig. 2B; Supplementary Table S2) based on previous literature (Krafnick et al. 2014; Ma et al. 2015; Ramus et al. 2018; Williams et al. 2018).

Diffusion Weighted Imaging data

Preprocessing

The preprocessing of Diffusion Weighted Imaging (DWI) data was performed using the Functional Magnetic Resonance Imaging of the Brain (FMRIB) Software Library (FSL 6.0.2; Smith et al.

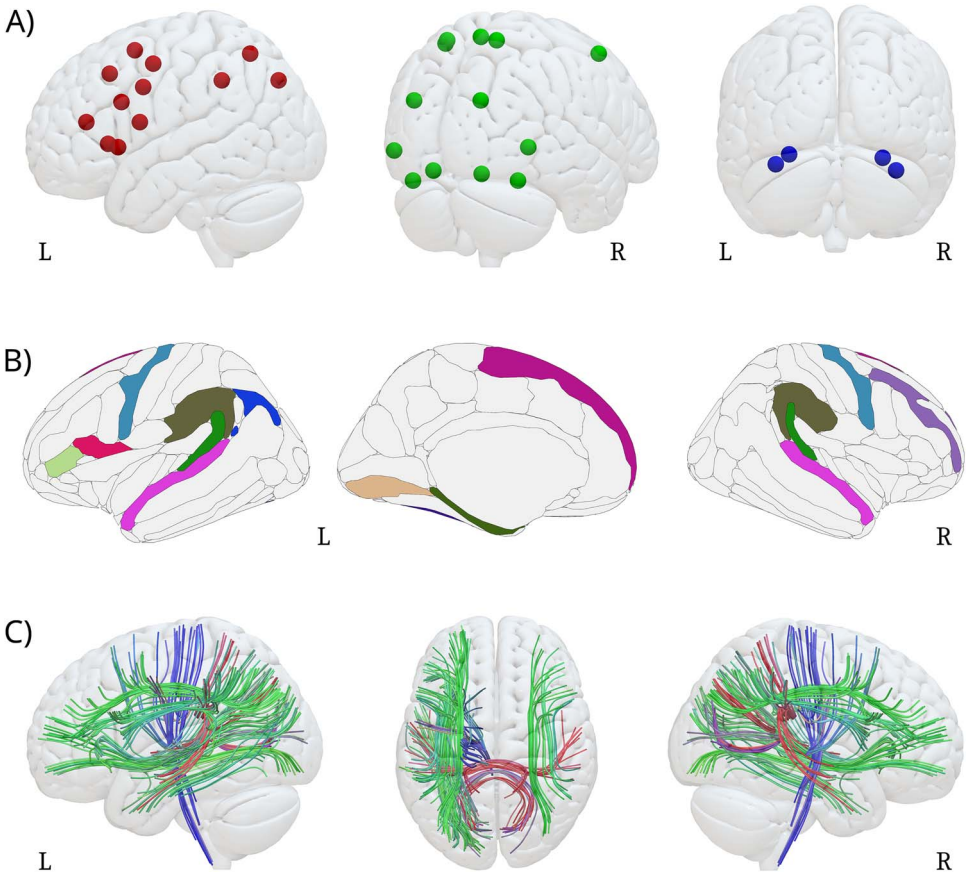

Fig. 2. Regions of interest. (A) Phonological (red), magnocellular (green), and cerebellar (blue) task-specific ROI masks used in the second-level, univariate and multivariate fMRI analyses. (B) Regions of interest of the Destrieux atlas (Destrieux et al. 2010) included in the cortical mask for GMV and CT univariate and multivariate analyses. The different colors identify each single ROI. (C) Tracts of interest of the WM atlas (Radwan et al. 2022) included in the mask used for fractional anisotropy (FA) and mean diffusivity (MD) univariate analyses. Red (left–right), green (antero-posterior), and blue (ventro-dorsal) colors represent the prevalent spatial orientation of the individual tracts. ROIs, regions of interest; GMV, gray matter volume; CT, cortical thickness; L, left; R, right.

2004). After visual inspection for motion artifacts, images were corrected for the effects of eddy currents (Horsfield 1999) and then segmented (nonbrain tissue voxels were removed from the whole head images) by the Brain Extraction Tool (Smith 2002). Diffusion tensor models were then fitted at each voxel with DTIFIT in the FSL Diffusion Toolbox. This procedure allows to derive the tensor eigenvalues that describe the diffusion coefficients in the primary, secondary, and tertiary diffusion directions, leading to the estimation of fractional anisotropy (FA) and mean diffusivity (MD; Pierpaoli and Basser 1996).

Tract-based spatial statistics region of interest analyses

Individual FA images were analyzed using Tract-based spatial statistics (TBSS) in FSL (Smith et al. 2006). First, images were aligned to a standard FA template ("FMRIB58_FA") using a nonlinear registration algorithm. The resulting normalized FA images were averaged to create a mean FA image which was used to generate a mean FA skeleton, representing the centers of all tracts common to all subjects. A threshold of FA > 0.25 was applied for creating the mean FA skeleton to exclude non-WM voxels. All individual FA and MD maps were then projected onto the mean FA skeleton. Tracts of interest were selected based on previous studies showing their involvement in dyslexia etiology (Vandermosten et al. 2012b; Zhao et al. 2016; Zhao et al. 2022) (Fig. 2C; Supplementary Table S3). We defined two GLMs to investigate the effect of dyslexia diagnosis on, respectively, FA and MD measures within the skeletonized masked images. We then applied the FSL randomize permutation-based program with 5,000 permutations. The significance threshold was set at $P < 0.05$ corrected for multiple comparisons by means of threshold-free cluster enhancement (TFCE; Smith and Nichols 2009).

Correlation analyses

Significant neuroimaging results for the three different MRI modalities (fMRI, GM, WM) were correlated with the behavioral measures that showed significant between-group differences (see Results 3.1). In order to avoid a statistical circularity fallacy, the analyses were performed at the regional level, namely if the between-group differences fell within a specific ROI or set of ROIs, the whole ROI(s)—and not only the significant voxels—were used for computing the correlations with the behavioral measures.

Multivariate pattern analyses

Multivariate classification between the two groups was performed with the PyMVPA toolbox (version 2.6.5) (www.pymvpa.org; Hanke et al. 2009) running on Python 2.7.18 (www.python.org). The input data for the multivariate pattern analyses (MVPAs) were the same fMRI, GM, and diffusion measures examined at the univariate level.

For each of the three experimental MMR fMRI tasks, one MVPA was performed on the t-statistical images (spmT) derived from the respective simple MMR contrast (deviant blocks > standard blocks) and was anatomically restricted to the task-specific ROI mask (Tables S1A, S1B, S1C). One MVPA model was specified for each of the two GM measures (CT, GMV), by first converting the corresponding surface-based maps to volume-based maps and restricting the analysis to the cortical ROI mask (Table S2). Another MVPA model was defined for each of the two diffusion indices (FA, MD), using the corresponding skeletonized images, and restricting the analysis to the WM tracts of interest (Table S3).

For all MVPAs, we used a C-Support Vector Machine (C-SVM) classifier, based on a leave-one-pair-of-subjects-out (LOPSO) cross-validation procedure. In case of a resulting classification accuracy above chance (> 50%), we also performed a searchlight MVPA (Kriegeskorte et al. 2006), with a LOPSO cross-validation procedure and a Gaussian Naive Bayes (GNB) classifier. The searchlight spheres had a radius of 4 mm and were centered at each voxel comprised within the specific mask. The statistical significance was determined with a GNB permutation test (Stelzer et al. 2013).

Multimodal machine learning model

We implemented a ML model to evaluate whether combining multimodal features could effectively discriminate between subjects with and without dyslexia. The features used in the model corresponded to the fMRI, GM, and diffusion brain measures also analyzed with univariate and MVPA approaches. For the three experimental MMR fMRI tasks, GMV, CT, FA, and MD, mean values at the single-subject level were extracted from individual ROI in the respective sets (Tables S1A, S1B, S1C, S2, S3). Total brain volume and mean thickness values of the left and right hemispheres were also extracted (Ramus et al. 2018).

All features for all subjects were concatenated into a matrix with 38 rows (number of subjects) and 102 columns (overall number of ROI features), which was used as input for the ML analysis. The model was implemented using the code released by Cui and colleagues (Cui et al. 2016). The linear C-SVM algorithm with the regularization parameter (C) set to 1 was used to discriminate between the two groups. This discrimination task was performed using a nested-leave-one-out approach, consisting of inner and outer loops. The inner loop served for performing feature selection, and the outer loop served to evaluate classification accuracy. The significance of the accuracy result was determined with permutation tests ($n = 1{,}000$). The most discriminative features were defined as those that had been used in the outer loop of all folds. Classification scores were computed as the distance from the hyperplane and used to determine the class label of each subject (Cui et al. 2016). To compare the performance of combined neuroimaging modalities with individual ones, we implemented four different models: (i) the multimodal model including features of all modalities (fMRI, GM, WM), and three single-modality models including (ii) only fMRI features, (iii) only GM features, (iv) only WM features.

For the only purpose of confirming the trend-wise reliability of the ML classification approach with respect not only to the diagnostic labels (dyslexic vs. controls), but also to the individual participants' reading skills, we correlated the classification scores of the winning model with the behavioral performance at reading tests. Since this approach entails a circular inference, we only report the effect size magnitude (Spearman's rho), without *P*-value. In addition, to circumvent the problem of circularity altogether, we also implemented a different ML analysis, using a regression model with the full sample of participants ($n = 38$), instead of the binary group classification approach. The ML regression model was trained to predict the number of reading errors from the multimodal neuroimaging features, without relying on diagnostic group labels. Details and results of the ML regression analysis are provided in Supplementary Methods S3, Supplementary Results S2, and Supplementary Table S7.

Results

Behavioral results

In the RAPM all participants obtained a score above the 10th percentile of the normative data (Arthur et al. 1999), with no significant differences between the two groups (Supplementary Materials: Table S4). As for the MT-16-19 battery, reading skills differed significantly between the two groups both in terms of the number of reading errors ($T(1,36) = 4.35$, $P = 0.0002$) and of the reading time parameter ($T(1,36) = 4.37$, $P = 0.0002$) (Supplementary Materials: Table S4).

As far as the phonological skills were concerned, dyslexic participants performed more poorly than controls in the following tests: PA segmentation time ($T(1,36) = 3.35$, $P = 0.018$); PA fusion time ($T(1,36) = 3.23$, $P = 0.026$); RAN time ($T(1,36) = 4.47$, $P = 0.0007$); digit span forward score ($T(1,36) = -3.75$, $P = 0.006$); digit span backward score ($T(1,36) = -3.54$, $P = 0.011$). Nonsignificant results emerged from the immediate and delayed recall tasks (Fig. 3A; Table S4). Accuracy in the segmentation and fusion PA, and the RAN tasks were not analyzed, due to ceiling effects present in at least one of the two groups. No significant between-group differences emerged in the CMP and the SRT tasks (Fig. 3B, 3C; Table S4).

Univariate functional Magnetic Resonance Imaging results

Regardless of the (dyslexic, control) group, the fMRI tasks elicited detectable MMR activations (Supplementary Table S5).

When testing for between-group (dyslexic vs. controls) differences in each of the experimental tasks, no significant MMR differences emerged in the phonological and the magnocellular tasks. In turn, a significant MMR difference was found in the rhythmic cerebellar task: dyslexic participants showed reduced MMR compared to controls in the right cerebellar lobule VI (peak-level $P_{FWE} = 0.033$, $Z(1,36) = 3.77$, 1 voxel, $x = 34$, $y = -58$, $z = -26$; Fig. 4A). Given this significant result, we correlated the cerebellar MMR activation (averaged over all voxels of the anatomical ROI, see Methods 2.7) with the participants' performance in the behavioral tests. The results showed a significant positive correlation between the digit span forward task and the MMR in the right cerebellum, lobule VI ($P_{FWE} = 0.004$, $Z(1,36) = 4.33$, 5 voxels, $x = 36$, $y = -58$, $z = -26$; Fig. 4B).

Results for the control tasks are reported in Supplementary Results S1. Data analyses investigating the interaction effect between the factors Group, Task, and MMR revealed no significant results.

Univariate T1-weighted results

We found no significant between-group differences.

Univariate Tract-based spatial statistics results

Between-group FA and MD differences were found in WM tracts of interest. More specifically, dyslexic subjects showed reduced FA compared to controls in the third segment of the left superior longitudinal fasciculus - segment III (SLF III) (Fig. 5A; Table 2A).

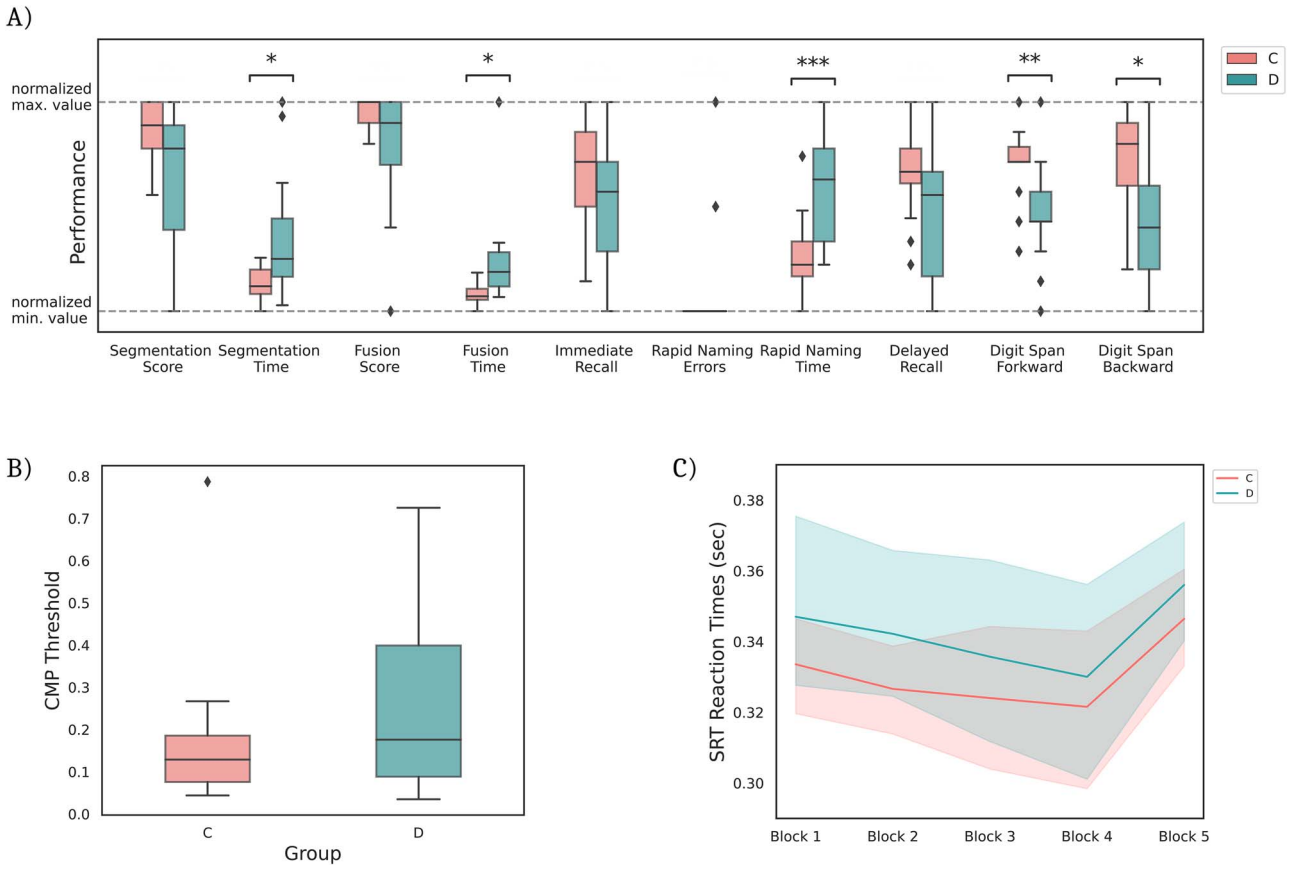

Fig. 3. Behavioral results. (A) Significant between groups differences emerged for the following phonological tests: segmentation-time, fusion-time, rapid naming-time, digit span forward, digit span backward. A Bonferroni correction for multiple comparisons (number of tests = 7) was applied. *: $P < 0.05$; **: $P < 0.01$; ***: $P < 0.001$. Segmentation and fusion scores, as well as rapid naming errors, were not analyzed due to ceiling effects in at least one of the two groups, and are presented here only for illustration purposes. Performance values of all the tasks were normalized to bring them onto a common scale. No significant differences were found for the magnocellular (B) and cerebellar (C) tasks. C, controls; D, dyslexics; CMP, coherent motion perception; SRT, serial reaction time task; sec, seconds.

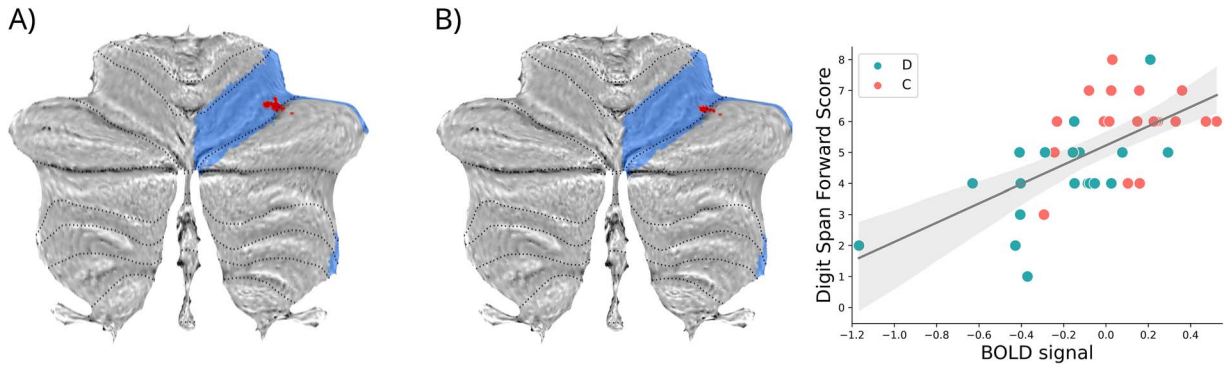

Fig. 4. Univariate fMRI results. (A) MMR activation in control > dyslexic participants in the experimental cerebellar task. Significant voxels (in red; $P < 0.05$, FWE-corrected at cluster-level for illustration purposes) are superimposed on the anatomical region corresponding to cerebellar lobule VI (in light blue) extracted from the cerebellar atlas by Diedrichsen et al. (2009). (B) Positive correlation between the cerebellar MMR and the digit span forward scores (significant voxels in red, $P < 0.05$, peak-level FWE-corrected; cerebellar lobule VI, in light blue). The scatter plot shows the average BOLD signal across significant voxels against the digit span forward scores. BOLD, blood oxygenation level dependent; C, controls; D, dyslexics.

MD was significantly higher in the dyslexic group in several WM tracts of the left hemisphere, including the SLF, arcuate fasciculus (AF), cortico-spinal tract (CST), thalamo-cortical radiation (ThR), middle longitudinal fasciculus (MdLF), inferior longitudinal fasciculus (ILF), inferior fronto-occipital fasciculus (IFOF), and the posterior portions of the corpus callosum (Fig. 5B; Table 2A). Based on these results, we correlated FA and MD in the significant WM tracts (see Methods 2.7) with the participants' performance in the behavioral tests. The results showed significant negative correlations between FA in the SLF III and: (i) the number of reading errors; (ii) the time required for performing the PA fusion task. Significant positive correlations were found between FA in the SLF III and the time required for performing (i) the phonological fusion task, and (ii) the PA segmentation task (Fig. 5C; Table 2B). MD across a wide set of tracts, including the left AF, SLF, CST, and ThR, was positively correlated with the number of reading errors. In turn, a negative correlation was found between the accuracy in the digit span backward task and MD in several tracts, including

Table 2A. Univariate between-group differences in FA and MD of WM tracts.

		Cluster size	Peak MNI coordinate (x, y, z, in mm)	Tract	*P*-value
Fractional anisotropy					
	C > D	10	−26 −18 29	L-SLF III	0.049
Mean diffusivity					
	D > C	4672	−33 −12 40	L-SLF III (30%) L-SLF IId (29%) L-SLF IIv (27%) L-CST (26%) L-AF (24%) CC-parietal (4%) L-ThR (3%)	0.012
	D > C	1624	−24 −74 19	L-MdLF (13%) L-IFOF (8%) CC-occipital (8%) CC-parietal (6%) L-SLF IId (6%) L-ILF (4%) L-SLF IIv (4%) L-SLF III (3%) CC-temporal (1%)	0.029

the left SLF, AF, ThR, IFOF, and the posterior part of the corpus callosum (Fig. 5D; Table 2B).

Multivariate Pattern Analyses results

The mean classification accuracy of the MVPA classifier was below the chance level (50%) for both the phonological (48.8%) and magnocellular (40.16%) tasks. In turn, the mean accuracy for the cerebellar task was 62.46% (Fig. 6A). The searchlight analysis on the cerebellar task indicated that the sphere yielding the most accurate between-group classification accuracy (74.79%; Fig. 6B) was located in the right cerebellar lobule VI ($P = 0.0014$ against 20,500 permutations, not passing Bonferroni correction for the number of spheres = 1,007; x = 34, y = −48, z = −28; Fig. 6C).

As for GM measures, the MVPA classifier reached a mean classification accuracy of 55% for GMV and 57% for CT (Fig. 6A). The searchlight analysis on GMV revealed 80 searchlight spheres with a mean between-group classification accuracy of 77.8% (Fig. 6B), which were located in left inferior frontal and occipito-temporal gyri, right middle frontal and precentral cortices, and bilateral superior frontal and temporal gyri ($P < 0.0001$ against 10,000 permutations; Bonferroni correction could not be assessed due to the exceedingly high computational load for the number of tested spheres = 80,928; Fig. 6C, Supplementary Table S6). The searchlight analysis on CT yielded 68 significant spheres with a mean between-group classification accuracy of 62.4% (Fig. 6B), which were located in the left inferior frontal gyrus, left parietal cortices, bilateral superior frontal, precentral and superior temporal gyri ($P < 0.0001$ against 10,000 permutations, no Bonferroni correction, number of spheres = 80,928; Fig. 6C, Supplementary Table S6).

As for WM diffusion indices, the mean classification accuracy was 57% for both FA and MD (Fig. 6A). The searchlight analysis on FA revealed 62 spheres with mean between-group classification accuracy of 77.7% (Fig. 6B), overlapping with all WM tracts of interest ($P < 0.0001$ against 10,000 permutations, no Bonferroni correction, number of spheres = 53,863; Fig. 6C, Supplementary Table S6). The searchlight analysis on MD yielded 147 spheres with a mean accuracy of 78.1% (Fig. 6B), overlapping with all WM tracts of interest except the left uncinate fasciculus ($P < 0.0001$ against 10,000 permutations, no Bonferroni correction, number of spheres = 53,863; Fig. 6C, Supplementary Table S6).

Machine learning results

The ML model including features from all modalities discriminated between the two groups with an accuracy of 66%. The area under the curve (AUC) was 0.62, whereas sensitivity and specificity were 68% and 63%, respectively. However, accuracy in this model did not reach significance in the permutation test ($P = 0.1$). Compared to the all-modality model, the three single-modality models (fMRI, GM, WM) yielded worse performances with, respectively, an accuracy of 31%, 34%, and 55% (Fig. 7).

Within the all-modalities model, the algorithm selected the set of multi-modal features leading to between-group separation. These included three fMRI features, related to MMR in the bilateral cerebellum during the experimental cerebellar task and in the left inferior frontal gyrus during the phonological task; three GM features, including CT in the left occipito-temporal and inferior frontal gyri, and GMV in the planum temporale of the left superior temporal gyrus; and 25 WM features, including both FA and MD measures (Fig. 7, Table 3).

By correlating the classification score of each participant in the all-modality model with the behavioral performance at reading tests, we found a positive correlation with the number of reading errors (Spearman's rho = 0.58; Fig. 7). As this classification approach involves circular inference, we do not report the *P*-value. To further address this issue of circularity, we also conducted a post hoc ML regression analysis, which helped circumventing this problem. In brief, this analysis revealed that the all-modality model significantly predicts the number of reading errors per subject ($P = 0.046$). For further details, see Supplementary Methods S3, Supplementary Results S2, and Supplementary Table S7.

Discussion

The current study aimed to investigate dyslexia in adult participants by combining multimodal neuroimaging and behavioral data through an innovative approach. To account for dyslexia

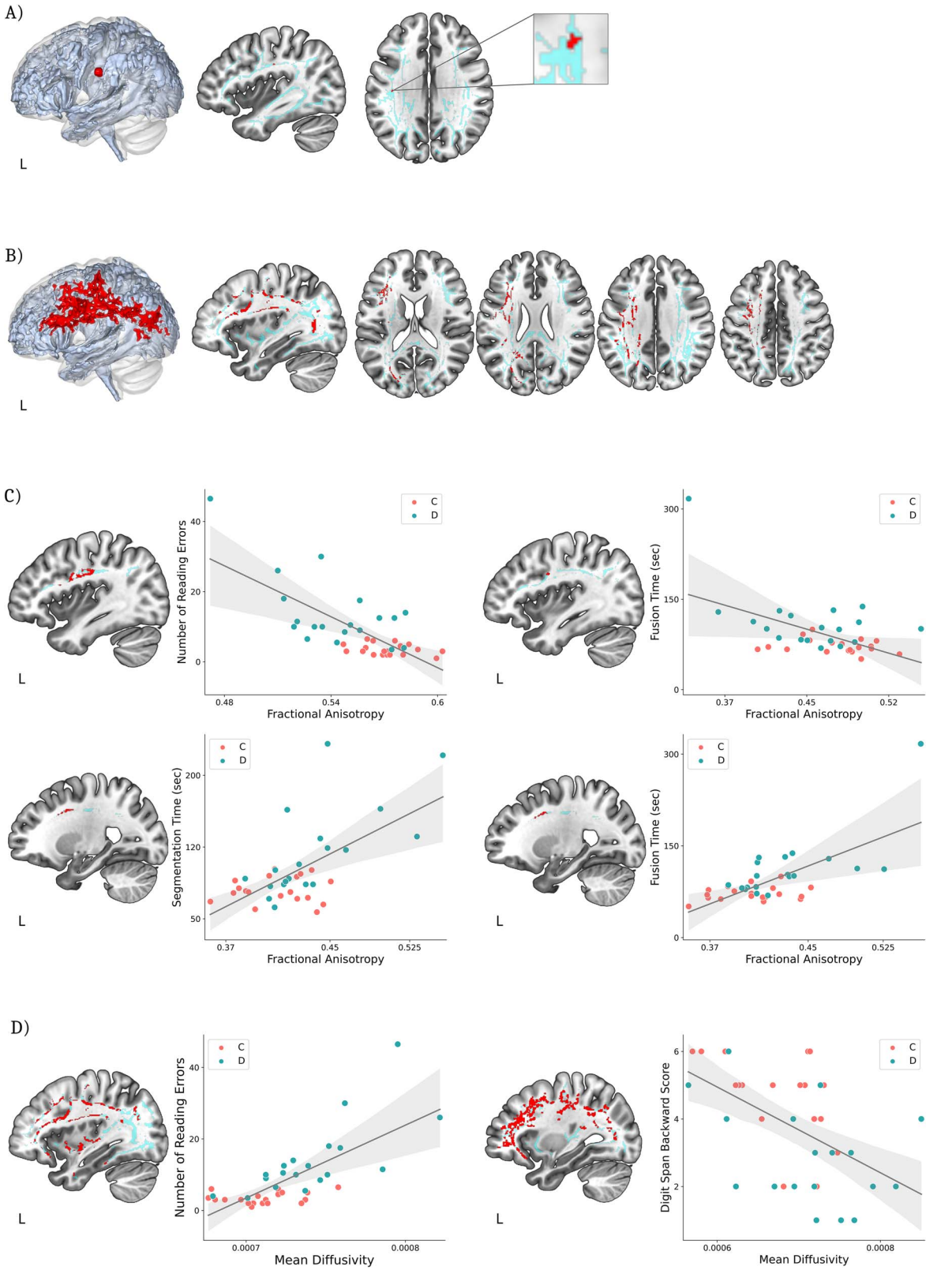

Fig. 5. Univariate FA and MD results. Compared to controls, dyslexic participants showed (A) reduced FA in the superior longitudinal fasciculus (SLFIII) and (B) increased MD in several WM tracts of the left hemisphere. The significant voxels ($P < 0.05$, TFCE-corrected) are superimposed (in red) on the tract based spatial statistics (TBSS) skeleton masked with the a-priori selected WM regions of interest (ROIs, in light blue). (C) Correlations between FA and behavioral tests. The significant voxels ($P < 0.05$, TFCE-corrected) are superimposed (in red) on the TBSS skeleton masked with the SLFIII tract (in light blue). The scatter plots show FA averages in the significant clusters (x-axis) against behavioral performances (y-axis). (D) Correlations between MD and behavioral tests. The significant voxels ($P < 0.05$, TFCE-corrected) are superimposed (in red) on the TBSS skeleton masked with the WM tracts showing significant MD differences between groups (in light blue). The scatter plots show MD averages in the significant cluster (x-axis) against behavioral performances (y-axis). C, controls; D, dyslexics; L, left; sec, seconds.

Table 2B. Correlations between FA or MD and performance in the behavioral tests.

		DoE	Cluster size	Peak MNI coordinate (x, y, z, in mm)	Tract	P-value
Fractional anisotropy						–
Reading errors		neg	279	−40 1 22	L-SLF III	0.002
		neg	95	−32 22 18	L-SLF III	0.030
PA fusion time		neg	11	−36 1 29	L-SLF III	0.040
		neg	1	−36 1 27	L-SLF III	0.049
		pos	65	−24 7 35	L-SLF III	0.006
PA segmentation time		pos	82	−26 8 33	L-SLF III	0.008
Mean diffusivity						
	Reading errors	pos	12,288	−36 −9 34	L-AF (50%) L-SLF IId (50%) L-SLF III (50%) L-ThR (47%) L-SLF IIv (46%) L-CST (45%) L-IFOF (26%) L-ILF (22%) L-MdLF (21%) CC-temporal (11%) CC-parietal (10%) CC-occipital (1%)	0.003
	Digit span backward	neg	10,158	−18 −39 38	L-SLF IId (54%) L-SLF III (51%) L-SLF IIv (49%) L-ThR (42%) L-AF(37%) L-CST (31%) L-IFOF (25%) L-MdLF (23%) CC-parietal (18%) CC-occipital (10%) CC-temporal (9%) L-ILF (6%)	0.007
		neg	326	20 −38 40	CC-parietal	0.012
		neg	159	−31 24 −14	L-IFOF	0.037
		neg	55	19 −55 35	CC-parietal	0.047
		neg	23	−19 11 −14	L-IFOF	0.049
		neg	11	25 −52 22	CC-parietal	0.045
		neg	5	20 −42 33	CC-parietal	0.048

The column "DoE" in Table 2B indicates the direction of the effect. The column "Tract" reports the WM tracts with which the significant correlation clusters overlap. If more than one, the numbers in parentheses indicate the percentage of the tract included in the corresponding cluster. Results are TFCE-corrected for multiple comparisons ($P < 0.05$). C = controls; D = dyslexics; MNI, Montreal Neurological Institute; PA, phonological awareness; pos, positive; neg, negative; L, Left; SLF, Superior Longitudinal Fasciculus; CST, Cortico Spinal Tract; AF, Arcuate Fasciculus; CC, Corpus Callosum; ThR, Thalamic Radiation; MdLF, Middle Longitudinal Fasciculus; IFOF, Inferior Fronto-Occiptal Fasciculus; ILF, Inferior Longitudinal Fasciculus.

heterogeneity, the predictions of some of the most influential theories about the disorder were tested both at the behavioral and brain levels. We first investigated whether dyslexic subjects showed worse behavioral performance in phonological, magnocellular, and cerebellar cognitive tests compared to controls. We then developed an MMR fMRI paradigm to evaluate whether dyslexic subjects showed reduced responses in tasks requiring phonological processing, motion perception, or implicit sequence learning. Structural gray and WM brain alterations associated with dyslexia diagnosis were also investigated. Between-group differences in each of the examined measures were tested with both univariate and multivariate approaches. As the main aim of the study, we implemented a multimodal machine learning algorithm (Cui et al. 2016) to test whether the combination of multimodal MRI-derived features accurately discriminated between subjects with and without dyslexia.

Behavioral data show phonological impairment in the dyslexic group

The cognitive assessment revealed that dyslexics were significantly slower than controls in PA and rapid automatized naming tasks. However, accuracy data from these tasks could not be analyzed due to ceiling effects observed in both groups for the RAN and fusion tasks, and in the control group for the segmentation task. Overall, the results of the cognitive assessment align with previous literature demonstrating that PA and RAN are robust cognitive predictors of dyslexia (Carioti et al. 2022; Devoto et al. 2022), with fluency being a more robust indicator than accuracy (Carioti et al. 2022). While our results are in line with the double-deficit hypothesis, which posits that dyslexia can stem from distinct impairments in phonological processing and naming speed (Vukovic and Siegel 2006; Stein 2023), the overlapping cognitive mechanisms between these two tasks prevent

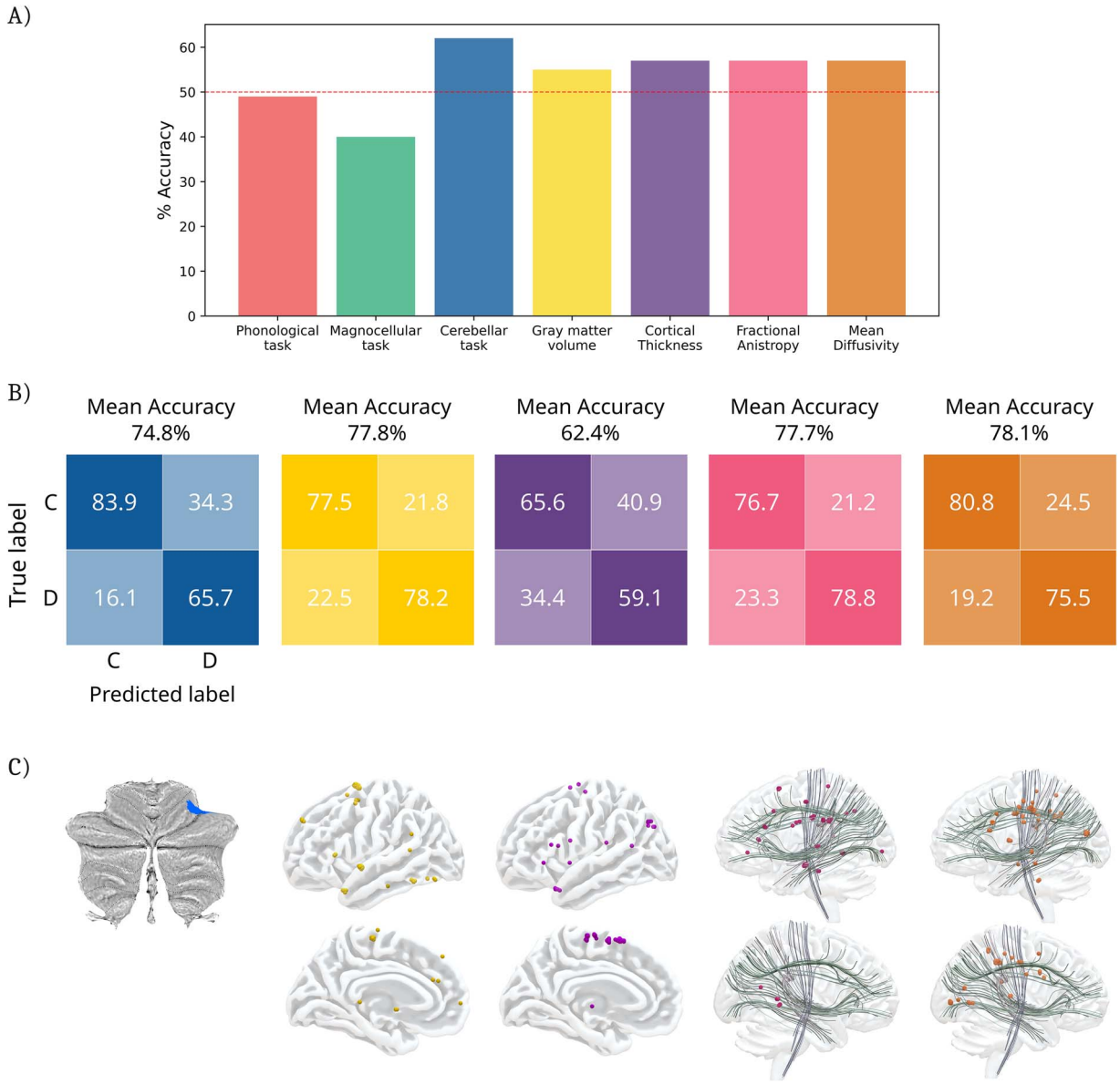

Fig. 6. Multivariate pattern analyses results. (A) Mean accuracy of the MVPA classifier trained with features from the following modalities: experimental phonological task (red), experimental magnocellular task (green), experimental cerebellar task (blue), GMV (yellow), CT (purple), FA (pink), MD (orange). The red dotted line represents the chance-level accuracy of 50%. (B) Confusion matrices displaying the mean percent classification accuracies across all significant ($P < 0.0001$) searchlight spheres. (C) Visual representation of the significant ($P < 0.0001$) searchlight spheres for each modality (see supplementary table S6). The colors in panels (B) and (C) correspond to those used in (A). C, controls; D, dyslexics.

us from drawing firm conclusions about the independence of these deficits. Dyslexic participants showed worse performance in phonological working memory skills, assessed through the digit span tasks, consistent with prior research suggesting that working memory training could be a potential treatment for reading and learning disorders (Maehler et al. 2019). In contrast, tests of immediate and delayed recall targeting semantic memory resulted in no significant differences between the two groups, confirming that semantic processing is preserved in dyslexia (Rasamimanana et al. 2020). Unexpectedly, we found no significant between-group differences in the CMP (testing for the magnocellular hypothesis) and SRT (testing for the cerebellar hypothesis) tasks. Despite the lack of statistical significance, the CMP (Fig. 3B) and SRT (Fig. 3C) response profiles revealed a greater data dispersion in the dyslexic group compared to controls, with a more heterogeneous distribution of data in the upper boundary (i.e. worse performance). This observation may indicate insufficient statistical power to reveal subtle effects, or large within-group variability in our dyslexic sample, with a subgroup of participants with magnocellular or cerebellar impairments, and another subgroup without, in line with the existence of different subtypes of the disorder (Menghini et al. 2010; Jednoróg et al. 2014; Wolf et al. 2024).

Functional Magnetic Resonance Imaging data show cerebellar impairments in the dyslexic group

Univariate analyses showed reduced MMR in dyslexics compared to controls in the right cerebellar lobule VI during the cerebellar task. MVPA showed that the cerebellar task also provided between-group classification accuracy above chance level. The searchlight analysis revealed that the brain region displaying the highest between-group classification accuracy was located in the

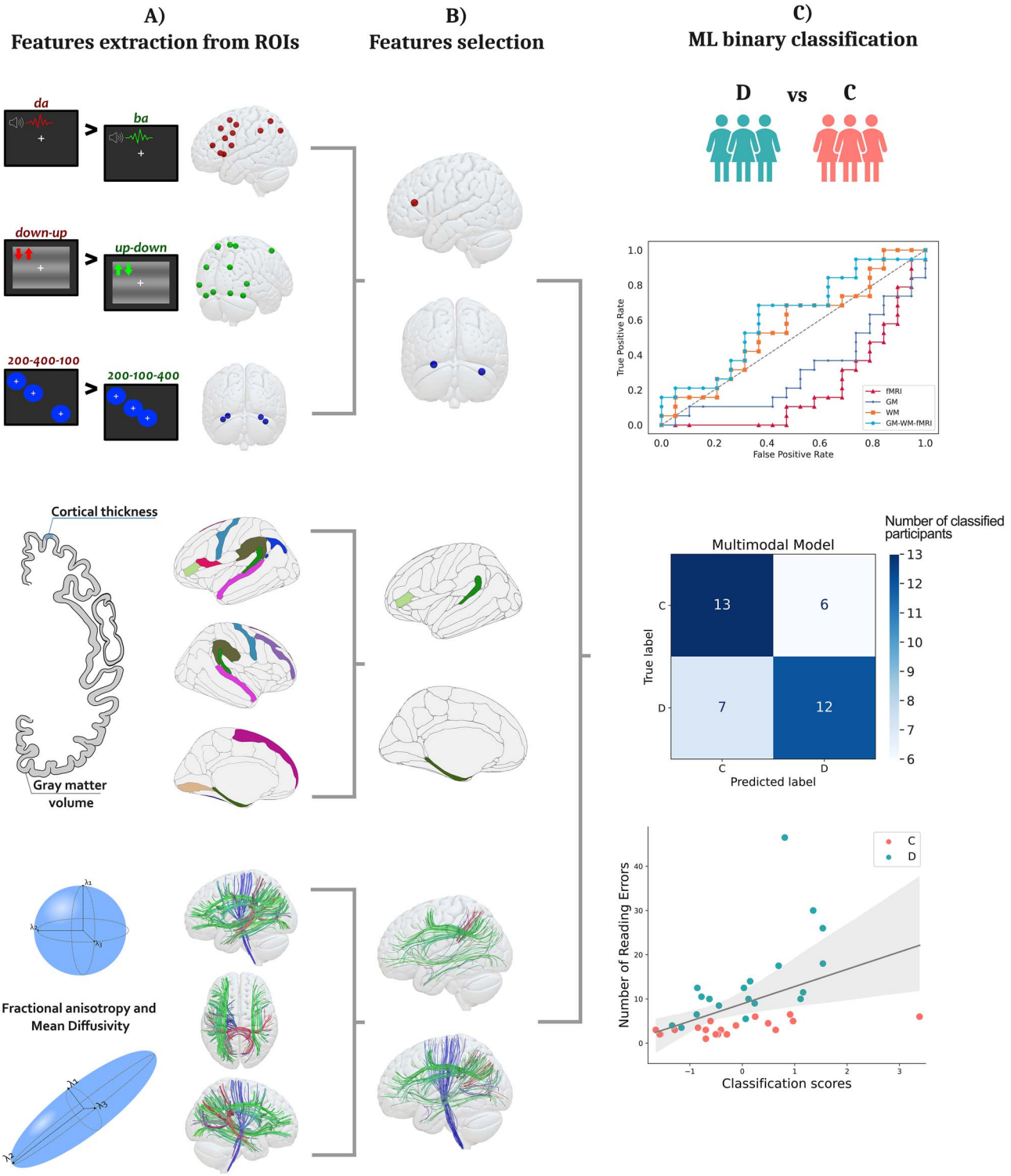

Fig. 7. Machine learning model. (A) Top: Mean BOLD signal extracted from single-subject data for the three experimental tasks (simple MMR contrast [deviant blocks > standard blocks]) from task-specific regions of interest. Middle: Mean CT and GMV extracted from anatomical regions of interest. Bottom: Mean FA and MD extracted from tracts of interest. (B) Discriminative features (i.e. features used in all folds) are shown. Top: BOLD signal in the left inferior frontal gyrus during the phonological task (red) and in the left and right cerebellar lobule VI during the cerebellar task (blue). Middle: CT in the left occipito-temporal and inferior frontal gyri and GMV in the left superior temporal gyrus. Bottom: FA (top brain) and MD (bottom brain) in several WM tracts. (C) Top: ROC curve comparison between the multimodal model and single-modality models. Middle: Confusion matrix showing correctly (diagonal) and incorrectly (off-diagonal) classified participants (controls vs. dyslexics). Bottom: Correlation between classification scores (x-axis) and the number of reading errors (y-axis). ROIs, regions of interest; C, controls; D, dyslexic participants; fMRI, functional magnetic resonance imaging; GM, gray matter; WM, white matter.

Table 3. Discriminant features of the all-modality ML model.

	Modality	Measures	Features
1	fMRI	cerebellar-rhythm task	Left cerebellum—Lobule VI
2	fMRI	phonological-syllables task	Left inferior frontal gyrus—pars triangularis
3	fMRI	cerebellar-rhythm task	Right cerebellum—Lobule VI
4	GM	CT	L occipito-temporal gyrus medial / parahippocampus
5	GM	CT	L inferior frontal gyrus—pars triangularis
6	GM	GMV	L_superior temporal gyrus—planum temporale
7	WM	FA	L superior longitudinal fasciculus—III
8	WM	FA	R superior longitudinal fasciculus—II dorsal
9	WM	FA	L superior longitudinal fasciculus—II ventral
10	WM	FA	L superior longitudinal fasciculus—II dorsal
11	WM	FA	L middle longitudinal fasciculus
12	WM	FA	Corpus callosum—parietal
13	WM	FA	L inferior fronto-occipital fasciculus
14	WM	FA	L uncinate fasciculus
15	WM	FA	L inferior longitudinal fasciculus
16	WM	FA	L arcuate fasciculus
17	WM	MD	L cortico spinal tract
18	WM	MD	R superior longitudinal fasciculus—II ventral
19	WM	MD	L arcuate fasciculus
20	WM	MD	L superior longitudinal fasciculus—II ventral
21	WM	MD	R superior longitudinal fasciculus—II dorsal
22	WM	MD	L superior longitudinal fasciculus—III
23	WM	MD	Corpus callosum—parietal
24	WM	MD	R superior longitudinal fasciculus—III
25	WM	MD	L inferior fronto-occipital fasciculus
26	WM	MD	L uncinate fasciculus
27	WM	MD	L superior longitudinal fasciculus—II dorsal
28	WM	MD	L thalamic radiation
29	WM	MD	L middle longitudinal fasciculus
30	WM	MD	L inferior longitudinal fasciculus
31	WM	MD	Corpus callosum—occipital

WM, white matter; GM, gray matter; fMRI, functional magnetic resonance imaging; FA, fractional anisotropy; MD, mean diffusivity; CT, cortical thickness; GMV, gray matter volume; L, left; R right.

right cerebellar lobule VI. This pattern of results is consistent with previous studies showing abnormal activation and reduced GMV in the right cerebellar lobule VI of dyslexic individuals (Nicolson et al. 2001; Menghini et al. 2006; Pernet et al. 2009; Stoodley 2016). This finding has been linked to the key role of the cerebellum in the predictive brain system (Leggio and Molinari 2015; Gatti et al. 2021). The cerebellum is known to support the learning of patterns of structured events, and the generation of predictions about the future based on these patterns. It also produces signals when an event deviates from these expectations (Leggio and Molinari 2015; Gatti et al. 2021). In light of this view, the cerebellum's involvement aligns with the MMR oddball paradigm, which is thought to reflect the brain's automatic detection of a prediction error when an incoming stimulus does not match an expected pattern (Fong et al. 2020). In the context of our fMRI study, the cerebellar task was the only one constituted by stimuli intrinsically characterized by a predictable rhythmic structure, thus likely engaging the cerebellar predictive system. Furthermore, the correlation analyses revealed that increased cerebellar activity was associated with better performance at the digit span forward test. This is consistent with the previous literature showing cerebellar involvement in phonological short-term memory (Marvel and Desmond 2010; Ashida et al. 2019; Peterburs et al. 2019) and outlines a probable connection between cerebellar impairment and phonological deficits (Nicolson et al. 2001; Nicolson and Fawcett 2019). Against our prediction, we found no significant results for the phonological and magnocellular tasks using either univariate or multivariate models. It remains unclear whether the neurocognitive processes underlying these tasks were compensated to the normal range in our sample of dyslexic adults, or whether there were subtle effects that could not be detected due to insufficient statistical power.

Structural data show widespread gray and white matter differences in the dyslexic group

We investigated the effects of dyslexia diagnosis on GMV and CT of several regions of interest. No significant between-group differences emerged from the univariate analyses. These null results may be linked to the inconsistent evidence about reliable structural markers of dyslexia emerged so far (Ramus et al. 2018). Inconsistencies across studies may be due to several factors, including dyslexia heterogeneity, gender differences, small sample size, and varying methodological approaches (Ramus et al. 2018). However, our multivariate searchlight analyses indicated gray matter measures as potential markers for distinguishing between dyslexic and control groups. More specifically, GMV in frontal, superior temporal, and occipito-temporal regions, along with CT in the inferior and superior frontal gyri and temporo-parietal regions, showed above-chance between-group classification accuracy. The differing results obtained from univariate and multivariate approaches may suggest that multivariate analyses

possess higher sensitivity for detecting subtle and distributed patterns of brain differences. Overall, these results confirm previous evidence that dyslexia is associated with widespread neuroanatomical variations, particularly in regions involved in reading and language processing (Ramus et al. 2018), reinforcing the multifactorial nature of the disorder.

DTI univariate analyses revealed widespread disorganization of WM microstructure associated with dyslexia diagnosis. Between-group comparisons highlighted reduced FA and increased MD in several left-hemispheric tracts of dyslexics compared to controls. Consistently, multivariate searchlight analyses showed that diffusion indices in WM tracts could also discriminate dyslexic from control subjects with accuracy above chance level. The involvement of these WM bundles in reading and language processing was corroborated by correlation analyses between diffusion indices and behavioral performance in reading and phonological tests. The results showed that higher FA and reduced MD were associated with better reading and phonological performance. These findings align with previous research (Vandermosten et al. 2012a; Vandermosten et al. 2012b; Cui et al. 2016; Ramus et al. 2018; Sihvonen et al. 2021), providing further support for the idea of dyslexia as a *disconnection syndrome*. According to this perspective, the reading deficit is characterized by disorganized WM connectivity across distributed brain networks, rather than localized impairments in isolated cortical areas (Paulesu et al. 1996; Vandermosten et al. 2012a; Habib 2021).

Machine learning model

The machine learning model trained with features from all modalities discriminated between dyslexic and control subjects with an accuracy of 66%, though this did not reach significance in the permutation test ($P = 0.1$). The single-modality models yielded lower accuracies: 31% for the fMRI model, 34% for the GM model, and 55% for the WM model. The multimodal model showed a positive correlation between the classification scores and the number of reading errors. The finding that the multimodal model outperformed single-modality models supports the multifactorial nature of dyslexia, suggesting that different structural and functional brain characteristics may be involved across individuals. Accordingly, the features selected by the all-modality model that mostly contributed to the between-group classification included measures from all three modalities. Consistently with the univariate results, the greatest proportion of selected features was represented by WM measures, including FA and MD in several associative, commissural, and projection tracts. Three fMRI features also contributed to the between-group classification, including MMR in the cerebellar region during the cerebellar task, with good agreement with the univariate and multivariate analyses, but also in the inferior frontal gyrus during the phonological syllable task, in line with previous evidence indicating abnormal activation in the inferior frontal gyrus during phonological tasks in dyslexic subjects compared to controls (Georgiewa et al. 2002; Devoto et al. 2022). The ML model further revealed that CT in the left occipito-temporal and inferior frontal regions, as well as GMV in the planum temporale of the superior temporal gyrus substantially contributed to between-group classification. This finding aligns with a large body of literature, showing CT and GMV alterations in occipito-temporal regions of dyslexic subjects (Kronbichler et al. 2008; Frye et al. 2010; Ramus et al. 2018). These regions include the so-called visual word form area, a brain region responsible for letters and word recognition (Kronbichler et al. 2008; Caffarra et al. 2021). As for the inferior frontal gyrus, although no study so far reported a thinner cortex in this region in dyslexic subjects compared to controls, GMV reductions have been previously observed (Brown et al. 2001; Norton et al. 2015; Ramus et al. 2018). Initial evidence of structural abnormalities in the planum temporale of dyslexic brains was provided by Galaburda's post-mortem studies (Galaburda et al. 1985) and has been largely confirmed by recent evidence (Bloom et al. 2013; Altarelli et al. 2014; Ramus et al. 2018). However, these results should be taken with caution, as it is worth emphasizing that the model did not reach statistical significance in the permutation testing. The classification accuracy of 66% is relatively low compared to similar studies, which have reported accuracies exceeding 80% (Cui et al. 2016; Tamboer et al. 2016; Langer et al. 2017; Zahia et al. 2020). This discrepancy may be due to the limited size of our samples, consistent with the notion that machine learning classifiers take particular advantage of larger datasets for effective training. Nonetheless, the classification scores for the individual participants derived from the all-modality model correlated with individual reading accuracy. This finding is particularly relevant, as it suggests that the combination of multimodal MRI features may offer insights into the individual severity of the reading deficits.

Limitations and future directions

The current study suffers from some important limitations. First, the relatively small sample size may determine reduced statistical power, potentially hindering the detection of subtle effects and introducing noise in effect estimation.

Second, investigating a sample of adult participants with high educational level for studying developmental learning disorders may be sub-optimal, as we cannot exclude compensatory mechanisms developed over time. Moreover, the fact that some dyslexic participants presented comorbidities with other learning disorders may represent a source of confound. Indeed, the cognitive and brain measures here examined are not exclusively related to dyslexia but also show correlations with other learning difficulties (Landerl et al. 2009; Willcutt et al. 2013). Consequently, some of the effects associated with dyslexia may be contaminated by other learning disorders.

Given these limitations, it would be valuable for future research to explore the current paradigm in a pediatric sample. The use of the mismatch negativity is particularly suitable for studying young children, as it captures an implicit and automatic response without requiring active task-related compliance. Furthermore, several studies suggested that phonological mismatch negativity and structural brain measures in preschool children could serve as potential predictors of future reading skills (Raschle et al. 2011; Raschle et al. 2012; van Zuijen et al. 2013; Vandermosten et al. 2015; Kraft et al. 2016). Consequently, longitudinal studies should evaluate whether the integration of multimodal neuroimaging and cognitive measures can predict the outcome of cognitive development and dyslexic impairments. This approach could offer important insights into the early identification and intervention for specific reading disorders.

Conclusions

Extensive research has shown that dyslexia involves widespread impairments across various cognitive, functional, and structural brain measures. The current study integrates these diverse elements, combining cognitive assessments with functional and structural MRI data to provide a more comprehensive view of the reading disorder. Our findings indicate that our sample of dyslexic

adults experience significant phonological difficulties. Functional abnormalities in the right cerebellum, along with widespread disorganization in gray and WM regions, likely contribute to reading and phonological difficulties. Crucially, this study demonstrated that using a multivariate model to combine diverse neuroimaging measures is the most effective approach to uncovering the complexity of developmental dyslexia.

Acknowledgments

We thank Ricardo De Haro Mancha for experimental assistance, and we also thank the volunteers who kindly participated in our study. M.G. was supported by a postdoctoral researcher grant by the University of Trento, Italy (Decree No. 48, 23.05.2019) awarded to M.T.

Author contributions

Cara Cristina (Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing—original draft, Writing—review & editing), Zantonello Giulia (Formal analysis, Investigation, Methodology, Writing—original draft, Writing—review & editing), Marta Ghio (Conceptualization, Methodology, Writing—original draft, Writing—review & editing), and Marco Tettamanti (Conceptualization, Formal analysis, Funding acquisition, Methodology, Project administration, Software, Supervision, Writing—original draft, Writing—review & editing).

Supplementary material

Supplementary material is available at *Cerebral Cortex* online.

Conflict of interest statement. None declared.

Data availability

The data that support the findings of this study are available on request from the corresponding author.

Ethics approval statement

The study was approved by the Ethical Committee of the University of Trento, Italy (Protocol no. 2019–014).

References

Alonso-Búa B, Díaz F, Ferraces MJ. 2006. The contribution of AERPs (MMN and LDN) to studying temporal vs. linguistic processing deficits in children with reading difficulties. *Int J Psychophysiol*. 59: 159–167. https://doi.org/10.1016/j.ijpsycho.2005.03.020.

Altarelli I et al. 2014. Planum temporale asymmetry in developmental dyslexia: revisiting an old question. *Hum Brain Mapp*. 35: 5717–5735. https://doi.org/10.1002/hbm.22579.

Arthur W, Day DV. 1994. Development of a short form for the raven advanced progressive matrices test. *Educ Psychol Measure*. 54:394–403. https://doi.org/10.1177/0013164494054002013.

Arthur W, Tubre TC, Paul DS, Sanchez-Ku ML. 1999. College-sample psychometric and normative data on a short form of the raven advanced progressive matrices test. *J Psychoeduc Assess*. 17: 354–361. https://doi.org/10.1177/073428299901700405.

Ashida R, Cerminara NL, Edwards RJ, Apps R, Brooks JCW. 2019. Sensorimotor, language, and working memory representation within the human cerebellum. *Hum Brain Mapp*. 40:4732–4747. https://doi.org/10.1002/hbm.24733.

Benassi M, Simonelli L, Giovagnoli S, Bolzani R. 2010. Coherence motion perception in developmental dyslexia: a meta-analysis of behavioral studies. *Dyslexia*. 16:341–357. https://doi.org/10.1002/dys.412.

Beyer M et al. 2022. Structural gray matter features and behavioral preliterate skills predict future literacy–a machine learning approach. *Front Neurosci*. 16:920150. https://doi.org/10.3389/fnins.2022.920150.

Bloom JS, Garcia-Barrera MA, Miller CJ, Miller SR, Hynd GW. 2013. Planum temporale morphology in children with developmental dyslexia. *Neuropsychologia*. 51:1684–1692. https://doi.org/10.1016/j.neuropsychologia.2013.05.012.

Boets B, Wouters J, van Wieringen A, Ghesquière P. 2006. Coherent motion detection in preschool children at family risk for dyslexia. *Vis Res*. 46:527–535. https://doi.org/10.1016/j.visres.2005.08.023.

Borghesani V et al. 2021. Functional and morphological correlates of developmental dyslexia: a multimodal investigation of the ventral occipitotemporal cortex. *J Neuroimaging*. 31:962–972. https://doi.org/10.1111/jon.12892.

Brown WE et al. 2001. Preliminary evidence of widespread morphological variations of the brain in dyslexia. *Neurology*. 56:781–783. https://doi.org/10.1212/WNL.56.6.781.

Bucci MP, Brémond-Gignac D, Kapoula Z. 2008. Poor binocular coordination of saccades in dyslexic children. *Graefes Arch Clin Exp Ophthalmol*. 246:417–428. https://doi.org/10.1007/s00417-007-0723-1.

Caffarra S, Karipidis II, Yablonski M, Yeatman JD. 2021. Anatomy and physiology of word-selective visual cortex: from visual features to lexical processing. *Brain Struct Funct*. 226:3051–3065. https://doi.org/10.1007/s00429-021-02384-8.

Carioti D et al. 2022. Rapid automatized naming as a universal marker of developmental dyslexia in Italian monolingual and minority-language children. *Front Psychol*. 13:783775. https://doi.org/10.3389/fpsyg.2022.783775.

Cornoldi C, Candela M. 2015. *Prove di lettura e scrittura MT-16-19 battery for the assessment of reading and writing in high school students*. Centro studi Erickson, Trento.

Cui Z, Xia Z, Su M, Shu H, Gong G. 2016. Disrupted white matter connectivity underlying developmental dyslexia: a machine learning approach. *Hum Brain Mapp*. 37:1443–1458. https://doi.org/10.1002/hbm.23112.

Dale AM, Fischl B, Sereno MI. 1999. Cortical surface-based analysis: I. Segmentation and surface reconstruction. *NeuroImage*. 9:179–194. https://doi.org/10.1006/nimg.1998.0395.

Danelli L et al. 2013. 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. https://doi.org/10.1002/hbm.22098.

Destrieux C, Fischl B, Dale A, Halgren E. 2010. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *NeuroImage*. 53:1–15. https://doi.org/10.1016/j.neuroimage.2010.06.010.

Devoto F, Carioti D, Danelli L, Berlingeri M. 2022. A meta-analysis of functional neuroimaging studies on developmental dyslexia across European orthographies: the ADOD model. *Lang Cogn Neurosci*. 37:285–314. https://doi.org/10.1080/23273798.2021.1970200.

Diedrichsen J, Balsters JH, Flavell J, Cussans E, Ramnani N. 2009. A probabilistic MR atlas of the human cerebellum. *Neuroimage*. 46:39–46. https://doi.org/10.1016/j.neuroimage.2009.01.045.

Fischer B, Hartnegg K. 2000. Stability of gaze control in dyslexia. *Strabismus*. 8:119–122. https://doi.org/10.1076/0927-3972(200006)821-2FT119.

Fischl B. 2012. FreeSurfer. *NeuroImage*. 62:774–781. https://doi.org/10.1016/j.neuroimage.2012.01.021.

Fischl B, Dale AM. 2000. Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc Natl Acad Sci*. 97:11050–11055. https://doi.org/10.1073/pnas.200033797.

Fischl B, Sereno MI, Dale AM. 1999. Cortical surface-based analysis: II: inflation, flattening, and a surface-based coordinate system. *NeuroImage*. 9:195–207. https://doi.org/10.1006/nimg.1998.0396.

Fischl B et al. 2002. Whole brain segmentation. *Neuron*. 33:341–355. https://doi.org/10.1016/S0896-6273(02)00569-X.

Fitzgerald K, Todd J. 2020. Making sense of mismatch negativity. *Front Psychiatry*. 11:468. https://doi.org/10.3389/fpsyt.2020.00468.

Fong CY, Law WHC, Uka T, Koike S. 2020. Auditory mismatch negativity under predictive coding framework and its role in psychotic disorders. *Front Psychiatry*. 11:557932. https://doi.org/10.3389/fpsyt.2020.557932.

Frye RE et al. 2010. Surface area accounts for the relation of gray matter volume to reading-related skills and history of dyslexia. *Cereb Cortex*. 20:2625–2635. https://doi.org/10.1093/cercor/bhq010.

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

Garrido MI, Kilner JM, Stephan KE, Friston KJ. 2009. The mismatch negativity: a review of underlying mechanisms. *Clin Neurophysiol*. 120:453–463. https://doi.org/10.1016/j.clinph.2008.11.029.

Gatti D, Rinaldi L, Ferreri L, Vecchi T. 2021. The human cerebellum as a hub of the predictive brain. *Brain Sciences*. 11:1492. https://doi.org/10.3390/brainsci11111492.

George F, Pech-Georgel C. 2017. *VALS-Valutazione delle difficoltà di lettura e scrittura in età adulta*. Erickson, Trento.

Georgiewa P et al. 2002. Phonological processing in dyslexic children: a study combining functional imaging and event related potentials. *Neurosci Lett*. 318:5–8. https://doi.org/10.1016/S0304-3940(01)02236-4.

Gomot M et al. 2006. Change detection in children with autism: an auditory event-related fMRI study. *NeuroImage*. 29:475–484. https://doi.org/10.1016/j.neuroimage.2005.07.027.

Gori S, Facoetti A. 2015. How the visual aspects can be crucial in reading acquisition: the intriguing case of crowding and developmental dyslexia. *J Vision*. 15:8. https://doi.org/10.1167/15.1.8.

Gori S, Seitz AR, Ronconi L, Franceschini S, Facoetti A. 2016. Multiple causal links between magnocellular-dorsal pathway deficit and developmental dyslexia. *Cereb Cortex*. 26:4356–4369. https://doi.org/10.1093/cercor/bhv206.

Goto M et al. 2022. Advantages of using both voxel- and surface-based morphometry in cortical morphology analysis: a review of various applications. *Magn Reson Med Sci*. 21:41–57. https://doi.org/10.2463/mrms.rev.2021-0096.

Gu C, Bi H-Y. 2020. Auditory processing deficit in individuals with dyslexia: a meta-analysis of mismatch negativity. *Neurosci Biobehav Rev*. 116:396–405. https://doi.org/10.1016/j.neubiorev.2020.06.032.

Habib M. 2021. The neurological basis of developmental dyslexia and related disorders: a reappraisal of the temporal hypothesis, twenty years on. *Brain Sci*. 11:708. https://doi.org/10.3390/brainsci11060708.

Hagler DJ, Saygin AP, Sereno MI. 2006. Smoothing and cluster thresholding for cortical surface-based group analysis of fMRI data. *NeuroImage*. 33:1093–1103. https://doi.org/10.1016/j.neuroimage.2006.07.036.

Hanke M et al. 2009. PyMVPA: a python toolbox for multivariate pattern analysis of fMRI data. *Neuroinform*. 7:37–53. https://doi.org/10.1007/s12021-008-9041-y.

Hoeft F et al. 2011. Neural systems predicting long-term outcome in dyslexia. *Proc Natl Acad Sci USA*. 108:361–366. https://doi.org/10.1073/pnas.1008950108.

Horsfield MA. 1999. Mapping eddy current induced fields for the correction of diffusion-weighted echo planar images. *Magn Reson Imaging*. 17:1335–1345. https://doi.org/10.1016/S0730-725X(99)00077-6.

Jednoróg K, Gawron N, Marchewka A, Heim S, Grabowska A. 2014. Cognitive subtypes of dyslexia are characterized by distinct patterns of grey matter volume. *Brain Struct Funct*. 219:1697–1707. https://doi.org/10.1007/s00429-013-0595-6.

Joo SJ, Donnelly PM, Yeatman JD. 2017. The causal relationship between dyslexia and motion perception reconsidered. *Sci Rep*. 7:4185. https://doi.org/10.1038/s41598-017-04471-5.

Kaisar S. 2020. Developmental dyslexia detection using machine learning techniques : a survey. *ICT Express*. 6:181–184. https://doi.org/10.1016/j.icte.2020.05.006.

Kevan A, Pammer K. 2008. Visual deficits in pre-readers at familial risk for dyslexia. *Vis Res*. 48:2835–2839. https://doi.org/10.1016/j.visres.2008.09.022.

Kovarski K et al. 2021. Emotional visual mismatch negativity: a joint investigation of social and non-social dimensions in adults with autism. *Transl Psychiatry*. 11:10–12. https://doi.org/10.1038/s41398-020-01133-5.

Krafnick AJ, Flowers DL, Luetje MM, Napoliello EM, Eden GF. 2014. An investigation into the origin of anatomical differences in dyslexia. *J Neurosci*. 34:901–908. https://doi.org/10.1523/JNEUROSCI.2092-13.2013.

Kraft I et al. 2016. Predicting early signs of dyslexia at a preliterate age by combining behavioral assessment with structural MRI. *NeuroImage*. 143:378–386. https://doi.org/10.1016/j.neuroimage.2016.09.004.

Kremláček J et al. 2016. Visual mismatch negativity (vMMN): a review and meta-analysis of studies in psychiatric and neurological disorders. *Cortex*. 80:76–112. https://doi.org/10.1016/j.cortex.2016.03.017.

Kriegeskorte N, Goebel R, Bandettini P. 2006. Information-based functional brain mapping. *Proc Natl Acad Sci*. 103:3863–3868. https://doi.org/10.1073/pnas.0600244103.

Kronbichler M et al. 2008. Developmental dyslexia: gray matter abnormalities in the occipitotemporal cortex. *Hum Brain Mapp*. 29:613–625. https://doi.org/10.1002/hbm.20425.

Kujala T, Näätänen R. 2001. The mismatch negativity in evaluating central auditory dysfunction in dyslexia. *Neurosci Biobehav Rev*. 25:535–543. https://doi.org/10.1016/S0149-7634(01)00032-X.

Landerl K, Fussenegger B, Moll K, Willburger E. 2009. Dyslexia and dyscalculia: two learning disorders with different cognitive profiles. *J Exp Child Psychol*. 103:309–324. https://doi.org/10.1016/j.jecp.2009.03.006.

Langer N et al. 2017. White matter alterations in infants at risk for developmental dyslexia. *Cereb Cortex*. 27:1027–1036. https://doi.org/10.1093/cercor/bhv281.

Leggio M, Molinari M. 2015. Cerebellar sequencing: a trick for predicting the future. *Cerebellum*. 14:35–38. https://doi.org/10.1007/s12311-014-0616-x.

Ma Y et al. 2015. Cortical thickness abnormalities associated with dyslexia, independent of remediation status. *NeuroImage Clin*. 7:177–186. https://doi.org/10.1016/j.nicl.2014.11.005.

Maehler C, Joerns C, Schuchardt K. 2019. Training working memory of children with and without dyslexia. *Children*. 6:47. https://doi.org/10.3390/children6030047.

Marvel CL, Desmond JE. 2010. Functional topography of the cerebellum in verbal working memory. *Neuropsychol Rev*. 20:271–279. https://doi.org/10.1007/s11065-010-9137-7.

McGrath LM, Peterson RL, Pennington BF. 2020. The multiple deficit model: progress, problems, and prospects. *Sci Stud Read*. 24:7–13. https://doi.org/10.1080/10888438.2019.1706180.

Menghini D, Hagberg GE, Caltagirone C, Petrosini L, Vicari S. 2006. Implicit learning deficits in dyslexic adults: an fMRI study. *NeuroImage*. 33:1218–1226. https://doi.org/10.1016/j.neuroimage.2006.08.024.

Menghini D et al. 2010. Different underlying neurocognitive deficits in developmental dyslexia: a comparative study. *Neuropsychologia*. 48:863–872. https://doi.org/10.1016/j.neuropsychologia.2009.11.003.

Näätänen R, Paavilainen P, Rinne T, Alho K. 2007. The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin Neurophysiol*. 118:2544–2590. https://doi.org/10.1016/j.clinph.2007.04.026.

Nakai T, Tirou C, Prado J. 2024. From brain to education through machine learning: predicting literacy and numeracy skills from neuroimaging data. *Imaging Neurosci*. 2:1–24. https://doi.org/10.1162/imag_a_00219.

Nemmi F et al. 2023. Developmental dyslexia, developmental coordination disorder and comorbidity discrimination using multimodal structural and functional neuroimaging. *Cortex*. 160:43–54. https://doi.org/10.1016/j.cortex.2022.10.016.

Nenning K-H, Langs G. 2022. Machine learning in neuroimaging: from research to clinical practice. *Radiologie*. 62:1–10. https://doi.org/10.1007/s00117-022-01051-1.

Nicolson RI, Fawcett AJ. 2019. Development of dyslexia: the delayed neural commitment framework. *Front Behav Neurosci*. 13:112. https://doi.org/10.3389/fnbeh.2019.00112.

Nicolson RI, Fawcett AJ, Dean P. 2001. Developmental dyslexia: the cerebellar deficit hypothesis. *Trends Neurosci*. 24:508–511. https://doi.org/10.1016/S0166-2236(00)01896-8.

Norton ES, Beach SD, Gabrieli JDE. 2015. Neurobiology of dyslexia. *Curr Opin Neurobiol*. 30:73–78. https://doi.org/10.1016/j.conb.2014.09.007.

Paulesu E et al. 1996. Is developmental dyslexia a disconnection syndrome?: evidence from PET scanning. *Brain*. 119:143–157. https://doi.org/10.1093/brain/119.1.143.

Paulesu E, Danelli L, 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. https://doi.org/10.3389/fnhum.2014.00830.

Peirce J et al. 2019. PsychoPy2: experiments in behavior made easy. *Behav Res Methods*. 51:195–203. https://doi.org/10.3758/s13428-018-01193-y.

Pernet CR, Poline JB, Demonet JF, Rousselet GA. 2009. Brain classification reveals the right cerebellum as the best biomarker of dyslexia. *BMC Neurosci*. 10:67. https://doi.org/10.1186/1471-2202-10-67.

Peterburs J, Blevins LC, Sheu Y-S, Desmond JE. 2019. Cerebellar contributions to sequence prediction in verbal working memory. *Brain Struct Funct*. 224:485–499. https://doi.org/10.1007/s00429-018-1784-0.

Peterson RL, Pennington BF. 2012. Developmental dyslexia. *Lancet*. 379:1997–2007. https://doi.org/10.1016/S0140-6736(12)60198-6.

Peterson RL, Pennington BF. 2015. Developmental dyslexia. *Annu Rev Clin Psychol*. 11:283–307. https://doi.org/10.1146/annurev-clinpsy-032814-112842.

Pierpaoli C, Basser PJ. 1996. Toward a quantitative assessment of diffusion anisotropy. *Magn Reson Med*. 36:893–906. https://doi.org/10.1002/mrm.1910360612.

Płoński P et al. 2017. Multi-parameter machine learning approach to the neuroanatomical basis of developmental dyslexia. *Hum Brain Mapp*. 38:900–908. https://doi.org/10.1002/hbm.23426.

Radwan AM et al. 2022. An atlas of white matter anatomy, its variability, and reproducibility based on constrained spherical deconvolution of diffusion MRI. *NeuroImage*. 254:119029. https://doi.org/10.1016/j.neuroimage.2022.119029.

Ramus F. 2001. Outstanding questions about phonological processing in dyslexia. *Dyslexia*. 7:197–216. https://doi.org/10.1002/dys.205.

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

Rasamimanana M, Barbaroux M, Colé P, Besson M. 2020. Semantic compensation and novel word learning in university students with dyslexia. *Neuropsychologia*. 139:107358. https://doi.org/10.1016/j.neuropsychologia.2020.107358.

Raschle NM, Chang M, Gaab N. 2011. Structural brain alterations associated with dyslexia predate reading onset. *NeuroImage*. 57:742–749. https://doi.org/10.1016/j.neuroimage.2010.09.055.

Raschle N et al. 2012. Pediatric neuroimaging in early childhood and infancy: challenges and practical guidelines. *Ann N Y Acad Sci*. 1252:43–50. https://doi.org/10.1111/j.1749-6632.2012.06457.x.

Schall U, Johnston P, Todd J, Ward PB, Michie PT. 2003. Functional neuroanatomy of auditory mismatch processing: an event-related fMRI study of duration-deviant oddballs. *NeuroImage*. 20:729–736. https://doi.org/10.1016/S1053-8119(03)00398-7.

Shaywitz SE, Shaywitz BA. 2005. Dyslexia (specific reading disability). *Biol Psychiatry*. 57:1301–1309. https://doi.org/10.1016/j.biopsych.2005.01.043.

Sihvonen AJ, Virtala P, Thiede A, Laasonen M, Kujala T. 2021. Structural white matter connectometry of reading and dyslexia. *NeuroImage*. 241:118411. https://doi.org/10.1016/j.neuroimage.2021.118411.

Smith SM. 2002. Fast robust automated brain extraction. *Hum Brain Mapp*. 17:143–155. https://doi.org/10.1002/hbm.10062.

Smith S, Nichols T. 2009. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage*. 44:83–98. https://doi.org/10.1016/j.neuroimage.2008.03.061.

Smith SM et al. 2004. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*. 23:S208–S219. https://doi.org/10.1016/j.neuroimage.2004.07.051.

Smith SM et al. 2006. Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *NeuroImage*. 31:1487–1505. https://doi.org/10.1016/j.neuroimage.2006.02.024.

Snowling MJ, Hulme C, Nation K. 2020. Defining and understanding dyslexia: past, present and future. *Oxf Rev Educ*. 46:501–513. https://doi.org/10.1080/03054985.2020.1765756.

Stein J. 2014. Dyslexia: the role of vision and visual attention. *Curr Dev Disord Rep*. 1:267–280. https://doi.org/10.1007/s40474-014-0030-6.

Stein J. 2019. The current status of the magnocellular theory of developmental dyslexia. *Neuropsychologia*. 130:66–77. https://doi.org/10.1016/j.neuropsychologia.2018.03.022.

Stein J. 2023. Theories about developmental dyslexia. *Brain Sci*. 13:208. https://doi.org/10.3390/brainsci13020208.

Stelzer J, Chen Y, Turner R. 2013. Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (MVPA): random permutations and cluster size control. *NeuroImage*. 65:69–82. https://doi.org/10.1016/j.neuroimage.2012.09.063.

Stoodley CJ. 2016. The cerebellum and neurodevelopmental disorders. *Cerebellum*. 15:34–37. https://doi.org/10.1007/s12311-015-0715-3.

Tamboer P, Vorst HCM, Ghebreab S, Scholte HS. 2016. Machine learning and dyslexia: classification of individual structural neuroimaging scans of students with and without dyslexia. *NeuroImage Clin*. 11:508–514. https://doi.org/10.1016/j.nicl.2016.03.014.

Tiadi A, Gérard C-L, Peyre H, Bui-Quoc E, Bucci MP. 2016. Immaturity of visual fixations in dyslexic children. *Front Hum Neurosci*. 10:58. https://doi.org/10.3389/fnhum.2016.00058.

Usman OL, Muniyandi RC, Omar K, Mohamad M. 2021. Advance machine learning methods for dyslexia biomarker detection: a review of implementation details and challenges. *IEEE Access*. 9:36879–36897. https://doi.org/10.1109/ACCESS.2021.3062709.

van Zuijen TL, Plakas A, Maassen BAM, Maurits NM, van der Leij A. 2013. Infant ERPs separate children at risk of dyslexia who become good readers from those who become poor readers. *Dev Sci*. 16:554–563. https://doi.org/10.1111/desc.12049.

Vandermosten M et al. 2012a. A tractography study in dyslexia: neuroanatomic correlates of orthographic, phonological and speech processing. *Brain*. 135:935–948. https://doi.org/10.1093/brain/awr363.

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

Vandermosten M et al. 2015. A DTI tractography study in pre-readers at risk for dyslexia. *Dev Cogn Neurosci*. 14:8–15. https://doi.org/10.1016/j.dcn.2015.05.006.

Vandermosten M et al. 2020. Brain activity patterns of phonemic representations are atypical in beginning readers with family risk for dyslexia. *Dev Sci*. 23:e12857. https://doi.org/10.1111/desc.12857.

Vicari S, Marotta L, Menghini D, Molinari M, Petrosini L. 2003. Implicit learning deficit in children with developmental dyslexia. *Neuropsychologia*. 41:108–114. https://doi.org/10.1016/S0028-3932(02)00082-9.

Vogt N. 2018. Machine learning in neuroscience. *Nat Methods*. 15:33–33. https://doi.org/10.1038/nmeth.4549.

Vu M-AT et al. 2018. A shared vision for machine learning in neuroscience. *J Neurosci*. 38:1601–1607. https://doi.org/10.1523/JNEUROSCI.0508-17.2018.

Vukovic RK, Siegel LS. 2006. The double-deficit hypothesis: a comprehensive analysis of the evidence. *J Learn Disabil*. 39:25–47. https://doi.org/10.1177/00222194060390010401.

Willcutt EG et al. 2013. Comorbidity between reading disability and math disability: concurrent psychopathology, functional impairment, and neuropsychological functioning. *J Learn Disabil*. 46:500–516. https://doi.org/10.1177/0022219413477476.

Williams VJ, Juranek J, Cirino P, Fletcher JM. 2018. Cortical thickness and local Gyrification in children with developmental dyslexia. *Cereb Cortex*. 28:963–973. https://doi.org/10.1093/cercor/bhx001.

Wolf M et al. 2024. Towards a dynamic, comprehensive conceptualization of dyslexia. *Ann of Dyslexia*. 74:303–324. https://doi.org/10.1007/s11881-023-00297-1.

Yang Y, Bi H-Y, Long Z-Y, Tao S. 2013. Evidence for cerebellar dysfunction in Chinese children with developmental dyslexia: an fMRI study. *Int J Neurosci*. 123:300–310. https://doi.org/10.3109/00207454.2012.756484.

Yu X et al. 2020. Putative protective neural mechanisms in prereaders with a family history of dyslexia who subsequently develop typical reading skills. *Hum Brain Mapp*. 41:2827–2845. https://doi.org/10.1002/hbm.24980.

Zahia S, Garcia-Zapirain B, Saralegui I, Fernandez-Ruanova B. 2020. Dyslexia detection using 3D convolutional neural networks and functional magnetic resonance imaging. *Comput Methods Prog Biomed*. 197:105726. https://doi.org/10.1016/j.cmpb.2020.105726.

Zhao J, Thiebaut de Schotten M, Altarelli I, Dubois J, Ramus F. 2016. Altered hemispheric lateralization of white matter pathways in developmental dyslexia: evidence from spherical deconvolution tractography. *Cortex*. 76:51–62. https://doi.org/10.1016/j.cortex.2015.12.004.

Zhao J et al. 2022. White matter connectivity in uncinate fasciculus accounts for visual attention span in developmental dyslexia. *Neuropsychologia*. 177:108414. https://doi.org/10.1016/j.neuropsychologia.2022.108414.

Cerebral Cortex, 2025, **35**, bhaf193
https://doi.org/10.1093/cercor/bhaf193
Original Article