

1 Abbreviated title: White matter plasticity in a RCT of reading training

2

3 **Assessing white matter plasticity in a randomized controlled trial of early literacy**
4 **training in preschoolers**

5

6 **Sendy Caffarra^{1,2}, Iliana I. Karipidis^{3,4,5}, John Kruper^{6,7}, Emily Kubota⁸, Adam Richie-**
7 **Halford², Megumi Takada⁸, Ariel Rokem^{6,7}, Jason D. Yeatman^{2,8}**

8

9 ¹ Department of Biomedical, Metabolic and Neural Sciences, University of Modena and
10 Reggio Emilia, Modena, Italy.

11 ² Division of Developmental-Behavioral Pediatrics, Stanford School of Medicine,
12 Stanford, CA, USA.

13 ³ Department of Child and Adolescent Psychiatry and Psychotherapy, University
14 Hospital of Psychiatry Zurich, Zurich, Switzerland.

15 ⁴ Neuroscience Center Zurich, University of Zurich and ETH, Zurich, Switzerland

16 ⁵ Department of Psychiatry and Behavioral Sciences, Stanford School of Medicine,
17 Stanford, CA, USA

18 ⁶ Department of Psychology, University of Washington, Seattle, WA, USA.

19 ⁷ eScience Institute, University of Washington, Seattle, WA, USA.

20 ⁸ Graduate School of Education, Stanford University, Stanford, CA, USA.

21

22 *Corresponding author:

23 Sendy Caffarra

24 caffarra@unimore.it (SC)

25 **Abstract**

26 Reading is a cognitive skill that requires our brain to go through a myriad of changes
27 during learning. While many studies have described how reading acquisition shapes
28 children's brain function, less is known about the impact of reading on brain structure.
29 Here we examined short-term causal effects of reading training on preschoolers' behavior
30 and white matter structure. Forty-eight English-speaking preschoolers (4y10m to 6y2m)
31 participated in a randomized controlled trial where they were randomly assigned to two
32 training programs: the Letter training program was focused on key skills for reading (e.g.,
33 decoding and letter knowledge), while the Language training program strengthened oral
34 language comprehension skills without exposure to text. Longitudinal behavioral data
35 showed that only the Letter Training group increased letter knowledge and decoding
36 skills after the two-week training. Diffusion MRI measures (FA and MD) of eighteen
37 white matter pathways (including the left arcuate and the left inferior longitudinal
38 fasciculus) did not reveal any statistically significant changes for either group despite
39 high degrees of scan-rescan reliability across sessions. These findings suggest that a two-
40 week reading training program can cause changes in preschoolers' letter knowledge and
41 decoding abilities, without being accompanied by measurable changes in the diffusion
42 properties of the major white matter pathways of the reading network. We conclude
43 highlighting possible constraints (i.e., age, training onset and duration, cognitive profile)
44 to reading-related white matter plasticity.

45 **Introduction**

46 Reading is a complex cognitive skill that has an impact on brain structure and function.
47 Learning to decode written language not only changes the way the brain functions [1,2],
48 but is also associated with changes in the structural properties of white matter pathways
49 [3–5]. However, it is still unclear under which conditions (e.g., quantity and quality of the
50 training, developmental stage), and at what time scale experience-dependent structural
51 changes emerge. In addition, the relationship between learning-driven changes in
52 reading behavior and brain plasticity is still underspecified. To deepen our
53 understanding of the short-term effects of the initial phase of reading acquisition on brain
54 structure, we ran a diffusion magnetic resonance study (dMRI) that used a randomized
55 controlled trial in preschoolers to test how training in letter-speech sound knowledge
56 affects behavior, white matter structure, and their relationship over the course of two
57 weeks.

58
59 A growing body of research reports a relationship between reading experience and
60 structural properties of white matter pathways [6]. Among the white matter tracts of the
61 reading circuitry there are the left arcuate (AF), connecting language areas in frontal and
62 temporal lobes [7], and the left inferior longitudinal fasciculus (ILF), which connects the
63 anterior temporal cortex to more posterior ventral-occipital areas (including the visual
64 word form area, [8]). Diffusion properties of left AF and ILF, such as fractional anisotropy

65 (FA) and mean diffusivity (MD), have been related to reading performances in a single
66 time point (concurrent or prior to the dMRI acquisition, [9–11]), as well as over a series
67 of longitudinal observations [4,5,12,13]. Recent studies have started to highlight the
68 presence of a dynamic relationship between the longitudinal trajectories of white matter
69 structure and changes of reading performance over time [14,15]. For instance, studies
70 focused on long-term reading-related structural plasticity reported that typically
71 developing children show increased FA and/or decreased MD in reading white matter
72 pathways (e.g., left AF and left ILF) as reading scores improve. These structural changes
73 are evident over one to four years of formal reading instruction [11,13,16–18]. Moreover,
74 the rate of FA change in the left AF relates to the rate of change in reading performances
75 over a period of five years of reading instruction [15].

76 Longitudinal studies focused on a shorter period of reading training have led to mixed
77 findings. Table 1 summarizes the available longitudinal research on reading-dependent
78 changes of white matter diffusion properties. Unfortunately, the picture provided by
79 these studies is only partial since most research so far has been focused on reading
80 interventions for children with (or at-risk of) reading disorders and its short-term effects
81 on white matter properties.

82 **Table 1. Studies on white matter changes due to short-term reading intervention**
83 **programs.**

Paper	Intervention		Age (y)	Sample Size	Language	Intervention specific effects (Group x Session)			
	Duration (h)	Type				Behavior	dMRI	Behavior - dMRI coupling	Diffusion property examined
[5]	100	g-p	8-10	72 (37 C)	English	x	x (FA, RD)	x (FA, RD)	FA, RD, AD
[4]	160 ^a	i-p	7-12	43 (19 C)	English	x	x (FA, MD)	x (MD)	FA, MD
[19]	160 ^a	i-p	7-12	73 (41 C)	English	x	x (MD, MDe)	N.A.	MD, MDe, AWF, DK, R1
[20]	24 ^b	g-p	8-9	35 (22 C)	English	x	-	N. A.	FA, MD
[21]	18	i-c	5-6	83 (52 C)	Dutch	N.A.	-	N.A.	FA
[3]	18	i-c	5-6	90 (59 C)	Dutch	N.A.	x (MWF)	N.A.	MWF
[12]	100-120	g-p	7-9	41 (15 C)	English	x	N.A.	x ^c (FA, MD)	FA, MD

84

85 Intervention Type: g = intervention was carried out in small groups; i = intervention was
 86 carried out individually; c = intervention was computerized; p = intervention was carried
 87 out in person.

88 Sample Size: overall sample size including the experimental and the control groups. C =
 89 control group sample size. The experimental groups always included children with
 90 reading disorders or at-risk of reading disorders who were enrolled in intensive
 91 remediation programs. The control groups included in the studies included typically

92 developing children, as well as children with or at-risk of reading disorders. They either
93 did not receive any intervention or were enrolled in a non-language specific training
94 program.

95 Intervention specific effects: x: present; - : absent, N.A.: not available. FA = fractional
96 anisotropy; RD = radial diffusivity; AD = axial diffusivity; MD = mean diffusivity; MDe =
97 extra-axonal mean diffusivity; AWF = axonal water fraction; DK = diffusion kurtosis; T1rt
98 = T1 relaxation time; MWF = myelin water fraction.

99 ^a The dMRI effects were visible after 46 hours of intervention.^b 225 hours only for 4 kids.
100 ^c It would not survive a multiple comparison correction.

101

102 Although the number of studies conducted on this topic is still low, some preliminary
103 trends can be highlighted given the available findings. First, all studies consistently
104 showed that only children who received reading intervention improved their reading
105 performance, confirming the efficacy of short-term reading training programs
106 [4,5,12,19,20]. Second, these intervention-specific behavioral changes were not always
107 accompanied by fast learning-driven white matter changes, suggesting that behavioral
108 and brain structure changes do not necessarily co-occur or this co-occurrence might be
109 specific to a subset of white matter diffusion properties [3–5,19–21]. Third, there is scarce
110 and mixed evidence on how effects generalize across different age ranges and types of
111 training, pointing to the need for additional research on short-term coupling between

112 reading behavior and white matter plasticity [4,5,12]. Finally, all studies listed above
113 focused on remediation programs and large effects were seen with intensive (high
114 dosage) training in late elementary school children [4,12]. Hence, they provide insights
115 on rapid white matter changes that might also reflect compensatory mechanisms of the
116 reading circuitry, or other factors that are unique to older children with dyslexia, rather
117 than solely plasticity due to the experience of learning to read.

118 The present exploratory study aims to complement the available research on short-term
119 white matter plasticity by focusing on language and literacy training programs in
120 typically developing preschool children. A randomized controlled trial was conducted
121 with English-speaking preschoolers who were enrolled in a two-week program which
122 either trained reading or spoken language skills (Letter and Language program,
123 respectively). Behavioral and dMRI measures were collected before and after the training
124 to test how preschool learning might affect changes in the white matter. To better
125 characterize the quality and consistency of children's dMRI measures over time, scan-
126 rescan reliability metrics were calculated for each dMRI measure (FA and MD) and white
127 matter tract.

128

129 **Materials and Methods**

130 **Participants**

131 Forty-eight English-speaking preschoolers (5 years of age; range: 58-74 months)
132 participated in a randomized controlled trial in the summer before starting kindergarten
133 (Fig 1). The recruitment period started on June 2nd and ended on November 19th, 2019.
134 An initial behavioral session ensured that all children participating in the study satisfied
135 the following inclusion criteria: not knowing all uppercase letters and their
136 corresponding sounds; having a Peabody Picture Vocabulary Test score higher than 85
137 (PPVT, 4th Edition; [22]); having normal or corrected to normal vision; being able to hold
138 still for 5 minutes during an MRI mock scan. No neuropsychological or psychiatric
139 disorder was reported. All children gave their assent to participate in the study, and their
140 parents (or legal guardians) signed an informed consent form. The study was approved
141 by the Institutional Review Board of the University of Washington.

142 --- Figure 1 ---

143 **Fig 1. Graphical representation of the randomized controlled trial.**

144

145 **Procedure**

146 Participants were randomly assigned to one of two different training programs: a Letter
147 (n = 24) or a Language (n = 24) organized into a fun and engaging summer camp. The
148 Letter program followed the Slingerland method [23] and was focused on the
149 foundational skills of reading such as letter recognition, letter-speech sound associations,

150 and phonemic awareness (e.g., blending and segmentation of syllables and trigrams;
151 Table 1). The Language program focused on oral linguistic abilities such as recognizing
152 syntactic categories in spoken sentences, listening, comprehending and retelling stories,
153 and learning new vocabulary (Table 1). Critically, the Language program did not include
154 any exposure to written language compared to the Letter program which was almost
155 exclusively focused on written language. Each training program was delivered to a small
156 group of children (n = 6) by two teachers, who had a background in Education or Speech
157 pathology. Each program lasted two weeks (3 hours/day, 5 days/week, 30 total hours)
158 and was based on pedagogical models of Direct Instruction and Gradual Release of
159 Responsibility. Letter and Language activities adopted a multisensory approach
160 involving vision, audition and kinesthetics. Both programs had the same daily schedule
161 and the learning process was scaffolded, so that the content of the activities followed an
162 increasing degree of complexity. Each daily session started with 20 minutes of free play
163 and ended with story time (Table 1). Starting from day 2, each daily activity started with
164 a brief rehearsal of previous lessons.

165

166 **Table 2. Description of the daily activities performed in the Letter and Language**
167 **Training programs.**

Letter Training Program daily activities

Free game (20 min)	Children could freely play with blocks, puzzles, and modeling clay.
Phonological awareness (25 min)	Children were introduced to segmentation and blending of syllables, trigrams (C-VC and CVC) and onset-rime words through songs and interactive games.
Letters (25 min)	Children learned two lowercase letters per day based on letter-picture correspondences and whiteboard writing activities.
Blending and Decoding (25 min)	Children were guided to blend three letters together and then decide whether the outcome was a real word or not.
Center activities (20 min x 2 daily sessions)	Children rotated among four different learning stations to reinforce what was learned in the daily session.
Story time (10 min)	The teachers read a brief (5-10 minutes) story to the kids.
Language Training Program daily activities	
Free game (20 min)	Children could freely play with blocks, puzzles, and modeling clay.
Syntax awareness (25 min)	Children built sentences using a set of picture cards, which represented different words. They were instructed about the function of different words in a sentence. Each card was color coded based on the grammatical category the word belonged to (e.g., noun, verb, adjective),
Listening and comprehension (25 min)	Children listened to a story and were guided to identify different narrative elements (e.g., characters, theme, problem).
Vocabulary (25 min)	Children were introduced to the meaning of new words based on picture cards, context-based information, personal experience, and examples. Simple exercises were proposed where kids had to use the new words in the right context.
Center activities (20 min x 2 daily sessions)	Children rotated among four different learning stations to reinforce what was learned in the daily session.
Story time (10 min)	The teachers read a brief (5-10 minutes) story to the kids, who were then asked to identify story elements based on what they learned in the listening and comprehension daily activity.

169 Behavioral and diffusion magnetic resonance imaging (dMRI) measures were collected
170 before (average time: 19 days, SD: 9.23) and after (average time: 8 days, SD: 3.90) each
171 training program.

172 **Behavioral data acquisition**

173 During the behavioral session, the following standardized tests were administered:
174 Phonological and Print Awareness Scale (tests of Initial sound matching, Final sound
175 matching, and Phonemic awareness were administered from [24]); Phonological
176 Awareness Literacy Screening (the Pseudoword decoding list was administered from
177 [25]); Narrative Language Measures (the Story Retell test was administered from [26]);
178 and the Expressive Vocabulary Test Third Edition [27]. Alphabet knowledge was also
179 tested through flashcards presented in random order. There were 26 flashcards for
180 uppercase letters, and 26 flashcards for lowercase letters. Each of the 26 cards was shown
181 to the child and they were asked “What letter is this?” and “What sound does it make?”.
182 The total score of the alphabet knowledge test was 52, both for upper and lowercase
183 letters.

184 **dMRI data acquisition and preprocessing**

185 MRI data was collected through a 3 T Phillips Achieva scanner with a 32-channel head
186 coil (Philips, Eindhoven, Netherlands). A whole-brain anatomical volume at 1.0 x 1.0 x

187 1.0 mm resolution was acquired using a T1-weighted MPRAGE sequence (TR 9.2 s, TE
188 4.35 ms, matrix size 224 x 224, field of view 224 x 224 x 170, 170 slices). Diffusion-weighted
189 magnetic resonance imaging (dMRI) data of the full brain were acquired with a spatial
190 resolution of 2.0 mm³ (anterior-posterior phase encoding direction). A diffusion-
191 weighted imaging (DWI) scan was acquired with 32 non-collinear directions (b-value =
192 1500 s/mm²; TR = 7200; TE = 83 ms). Four volumes with no diffusion weighting were also
193 acquired (b-value = 0). To correct for echo-planar imaging distortions, one scan with a
194 reversed phase encoding direction (posterior-anterior) and with three non-diffusion-
195 weighted volumes was collected.

196 The T1-weighted (T1w) images were corrected for intensity non-uniformity (INU) using
197 N4BiasFieldCorrection [28], ANTs 2.3.1), and used as T1w-reference throughout the
198 workflow. The T1w-reference was then skull-stripped using antsBrainExtraction.sh
199 (ANTs 2.3.1), using OASIS as target template. Spatial normalization to the ICBM 152
200 Nonlinear Asymmetrical template version 2009c [29] was performed through nonlinear
201 registration with antsRegistration (ANTs 2.3.1, [30], using brain-extracted versions of
202 both T1w volume and template. Brain tissue segmentation of cerebrospinal fluid (CSF),
203 white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w
204 using FAST (FSL 6.0.3, [31]).

205 DMRI preprocessing and reconstruction were carried out using QSIprep 0.13.0RC2 ([32–
206 34]), which is based on *Nipype* 1.6.0[32–34], *Nilearn* 0.7.1 [35] and *Dipy* 1.4.0 [36]. The

207 preprocessing included topup distortion, MP-PCA denoising, motion and Eddy current
208 correction (q -space smoothing factor = 10, 5 iterations; [37–40]). Only experimental
209 sessions with a maximum framewise displacement below 4 mm and an average
210 framewise displacement below 1 mm were further analyzed (Letter group pre-training
211 session: 21; Letter group pre-training session: 22; Language group pre-training session:
212 19; Language group post-training session: 20). Multi-tissue fiber response functions were
213 estimated using the dhollander algorithm as implemented in MRtrix3 [41]. Fiber
214 orientation distributions (FODs) in each voxel were estimated via constrained spherical
215 deconvolution (CSD,[42,43] using an unsupervised multi-tissue method[44,45].
216 Anatomically constrained tracking (ACT) was applied. FODs were intensity-normalized
217 using mtnormalize[46]. Probabilistic tractography was carried out using the following
218 QSIprep parameters: 1M streamlines, minimum length: 30 mm, maximum length: 250
219 mm. Fiber segmentation was carried out using pyAFQ 0.9 default parameters
220 (<https://yeatmanlab.github.io/pyAFQ>:[47,48] cleaning iterations = 5, distance threshold =
221 5 SD, length threshold: 4 SD). Eighteen default tracts were segmented: Left/Right Arcuate,
222 Left/Right Anterior Thalamic Radiation, Left/Right Cingulum, Left/Right Corticospinal
223 Tract, Anterior/Posterior Forceps, Left/Right Inferior Fronto-Occipital Fasciculus,
224 Left/Right Inferior Longitudinal Fasciculus, Left/Right Superior Longitudinal Fasciculus,
225 Left/Right Uncinate. Diffusion metrics were calculated using the constrained spherical
226 deconvolution model (CSD[49,50] and projected onto the tracts. Each streamline was

227 resampled into a fixed number of nodes (n = 100), and average values of fractional
228 anisotropy (FA), and mean diffusivity (MD) were calculated for each node. FA and MD
229 were mapped onto each tract, weighting the values based on the streamline's distance
230 from the core of the tract [47]. Hence, FA and MD values of each white matter tract were
231 calculated as the average across all 100 nodes of the tract profile.

232

233 **Results**

234 **Behavioral results**

235 A linear mixed effect model (LME) was run to test for behavioral effects due to the type
236 of training received. Time (pre vs post session), Training Type (Letter vs Language) and
237 their interaction were included as fixed effects. By-subject random intercepts and slopes
238 were also included. Only models on alphabet knowledge (average accuracy of upper and
239 lower case) and decoding skills showed a significant interaction between Training Type
240 and Time (alphabet knowledge: $\beta = 0.822$, SE = 0.370, $t = 2.224$, $p = 0.026$; decoding skills:
241 $\beta = 0.970$, SE = 0.326, $t = 2.972$, $p = 0.003$) indicating that children participating in the Letter
242 Training improved their letter knowledge ($\beta = 2.770$, SE = 0.784, $t = 3.532$, $p < 0.001$) and
243 decoding ability ($\beta = 1.689$, SE = 0.496, $t = 3.403$, $p = 0.001$), while children participating in
244 the Language Training group did not show such behavioral changes (alphabet

245 knowledge: $\beta = 0.896$, SE = 0.776, $t = 1.154$, $p = 0.248$; decoding skills: $\beta = 0.194$, SE = 0.415,
246 $t = 0.469$, $p = 0.639$; Fig 2).

247 --- Figure 2 ---

248 **Fig 2. Training-related behavioral changes.**

249 First two columns: behavioral changes in alphabet knowledge from the Letter and
250 Language Training groups for each experimental session. The third column shows the
251 distribution of alphabet knowledge changes (i.e. difference between the individual scores
252 obtained in the post and pre-training sessions) for each group.

253

254 **dMRI results**

255 **Scan Rescan reliability**

256 For each white matter tract and diffusion property (FA and MD), we calculated scan-
257 rescan reliability metrics to quantify the consistency of two types of dMRI measures
258 between experimental sessions: profile and subject reliability (as in [51]). For the profile
259 reliability, a Pearson correlation was calculated between the pre and post session tract
260 profiles of each participant. To calculate the subject reliability, a Pearson correlation was
261 calculated between the pre and post sessions median values of the tract profiles of each
262 participant. Individual correlation coefficients were averaged for each tract to obtain a

263 reliability estimate. Our dMRI measures showed high degrees of scan-rescan reliability
264 between the two experimental sessions (profile reliability: FA, median $r = 0.99$, range 0.93-
265 0.99; MD, median $r = 0.92$, range: 0.65-0.99; subject reliability: FA, median $r = 0.83$, range
266 0.62-0.90; MD, median $r = 0.87$, range: 0.55-0.92; Fig 3, for Subject reliability of each
267 experimental group, see S1 and S2 Figs).

268 --- Figure 3 ---

269 **Fig 3. Scan Rescan reliability.**

270 The two columns show the profile and subject reliability estimates of each white matter
271 tract examined in the study. The two rows show reliability estimates for FA and MD,
272 respectively. ARC: Arcuate Fasciculus; ATR: Anterior Thalamic Radiation; CGC:
273 Cingulum Cingulate; CST: Corticospinal Tract; FA: Anterior Forceps; FP: Posterior
274 Forceps; IFO: Inferior Longitudinal Fasciculus; ILF: Inferior Longitudinal Fasciculus; SLF:
275 Superior Longitudinal Fasciculus; UNC: Uncinate.

276

277 **Training effects on dMRI measures**

278 An LME model was run on the average FA and MD values of each tract profile to test for
279 structural changes due to the type of training received. Time (pre vs post session),
280 Training Type (Letter vs Language) and their interaction were included as fixed factors.
281 By-subject random intercepts were also included. No structural changes were observed
282 between experimental sessions (FA: all $ts < 2$; MD: all $ts < 2.5$) or between the two groups

283 (FA: all $ts < 2.8$; MD: all $ts < 2$). Interactions between Training Type and Time were not
284 significant (FA: all $ts < 2.12$; MD: all $ts < 2$), suggesting that no statistically significant
285 changes were observed for either group over the 2-week training period (Fig 4 shows the
286 results for the left arcuate; for similar results on the left ILF, see S3 Fig).

287 --- Figure 4 ---

288 **Fig 4. Training-related dMRI changes of the left arcuate.**

289 First two columns: structural changes of the left arcuate are shown for the Letter and
290 Language Training groups and each experimental session. The third column shows the
291 distribution of FA and MD changes (i.e. difference between the individual profiles
292 observed in the post and pre-training sessions) for each group.

293 Similar LME models were fitted for each single node of each tract profile and they
294 confirmed this pattern of results (Figs 5 and 6); there were no observable changes in white
295 matter properties over the two-week training period.

296 --- Figure 5 ---

297 **Fig 5. FA tract profile for each experimental group and training session.**

298 The plots show FA values estimated based on the beta coefficients extracted from node-
299 by-node LME models. Shaded areas represent ± 2 SE. ARC: Arcuate Fasciculus; ATR:
300 Anterior Thalamic Radiation; CGC: Cingulum Cingulate; CST: Corticospinal Tract; FA:

301 Anterior Forceps; FP: Posterior Forceps; IFO: Inferior Longitudinal Fasciculus; ILF:
302 Inferior Longitudinal Fasciculus; SLF: Superior Longitudinal Fasciculus; UNC: Uncinate.

303 --- Figure 6 ---

304 **Fig 6. MD tract profile for each experimental group and training session.**

305 The plots show MD values estimated based on the beta coefficients extracted from node-
306 by-node LME models. Shaded areas represent +/- 2 SE. ARC: Arcuate Fasciculus; ATR:
307 Anterior Thalamic Radiation; CGC: Cingulum Cingulate; CST: Corticospinal Tract; FA:
308 Anterior Forceps; FP: Posterior Forceps; IFO: Inferior Longitudinal Fasciculus; ILF:
309 Inferior Longitudinal Fasciculus; SLF: Superior Longitudinal Fasciculus; UNC: Uncinate.

310 To estimate the evidence supporting the null hypothesis (H_0 : no plasticity), additional
311 Bayesian analyses were run on each tract to compare the dMRI training effect (post-pre
312 mean profile difference) between the two groups. Bayes factors supported small-to-
313 moderate evidence for the null effect in the majority of the tracts, including all tracts that
314 are part of the reading circuitry (FA: $BFs < 1$ in 16 of the 18 tracts; MD: $BFs < 1$ in 12 of the
315 18 tracts; Fig 7).

316 --- Figure 7 ---

317 **Fig 7. Bayes factors relative to the group comparison of the dMRI training effect for**
318 **each tract.**

319 ARC: Arcuate Fasciculus; ATR: Anterior Thalamic Radiation; CGC: Cingulum Cingulate;
320 CST: Corticospinal Tract; FA: Anterior Forceps; FP: Posterior Forceps; IFO: Inferior
321 Longitudinal Fasciculus; ILF: Inferior Longitudinal Fasciculus; SLF: Superior
322 Longitudinal Fasciculus; UNC: Uncinate.

323

324 **Linking training effects between behavioral and dMRI measures**

325 Despite the lack of experience-driven white matter plasticity at the group level, it is
326 possible that longitudinal change in alphabet knowledge might relate to longitudinal
327 white matter changes. To test this hypothesis, the two groups were combined and
328 Pearson correlations were calculated to check whether individual changes in alphabet
329 knowledge (average of lower and upper case knowledge) could be mapped onto
330 structural changes of left AF and ILF. This analysis did not show significant effects after
331 Bonferroni correction (see Table 3).

	FA		MD	
	<i>r</i>	<i>p corrected</i>	<i>r</i>	<i>p corrected</i>
Left ARC	-0.38	0.104	0.12	0.99
Left ILF	-0.41	0.060	0.36	0.124

332 **Table 3.** Pearson correlations between post-pre differences in reading performance and
333 structural properties.

334

335 Discussion

336 This randomized controlled trial examined short-term effects of a Letter and a Language
337 training program on preschoolers' reading performance and brain structure. The findings
338 suggest that a two-week Letter training program causes improvements in preschoolers'
339 letter knowledge and decoding skills. However, this behavioral effect was not
340 accompanied by short-term changes in the diffusion properties (i.e., FA and MD) of white
341 matter pathways, within or outside the reading circuitry. The presence of quick
342 behavioral changes as a result of Letter training confirms previous findings on the
343 effectiveness of short-term reading instruction, which has been observed in children with
344 and without reading disorders [52–55].
345 Our dMRI findings further complement the existing literature on short-term reading-
346 relating brain plasticity by showing that reading performance improvements are not
347 always accompanied by changes in diffusion properties of white matter pathways
348 [4,5,12,20,21]. The high reliability estimates for both FA and MD scores across sessions
349 ensure that this null effect could not be accounted for by low dMRI data quality. Bayesian
350 analyses provided support for the null hypothesis (no change in white matter diffusion)
351 for all major white matter tracts of the reading network. In addition, correlation analyses
352 confirmed the lack of a clear correspondence between preschoolers' individual
353 behavioral changes and variations in structural properties of reading white matter
354 pathways.

355 One aspect that can account for the lack of structural changes is the type and intensity of
356 the reading program. Previous studies have mainly focused on effects of reading
357 intervention in children diagnosed with dyslexia, which can have an intense and
358 profound impact on struggling readers' cognitive and social lives. In the current study,
359 our reading training proposed preschool/kindergarten activities that are usually carried
360 out in a classroom setting. Since these training programs are similar to common preschool
361 and kindergarten classrooms, they might not represent a dramatic enough environmental
362 change to cause large-scale remodeling of the white matter. Related to this point, another
363 variable that can account for our results is the cognitive profile of the trainees. This is the
364 first randomized controlled trial on the effects of a short-term reading training with
365 typically developing preschoolers. Previous experimental evidence collected so far (Table
366 1) refers to the effects of short-term remediation programs on children with reading
367 disorders or at-risk of reading disorders. Hence, the large effects that have been reported
368 so far might reflect the dramatic environmental change of entering an intensive
369 intervention environment after struggling in school for years. This experience is quite
370 different than typically developing children beginning formal reading instruction [3–
371 5,19].

372 Another possible explanation to consider is the type of diffusion properties examined
373 here. Recent dMRI findings on the short-term effects of reading intervention programs in
374 preschoolers reported structural changes only in myelin water fraction, but not in FA and

375 MD scores [3,21]. This might suggest that MRI measures more specifically related to
376 myelination would better reflect reading-related short-term plasticity around 5 years of
377 age. However, within 7 and 12 years of age an opposite pattern of results have been
378 reported, with MD and FA providing evidence for rapid structural plasticity while no
379 training-dependent changes were reported for more myelin-specific correlates, such as
380 axonal water fraction and R1 [4,19]. These results are still compatible with the idea that
381 short-term plasticity due to reading training might affect different structural properties
382 of white matter depending on the developmental time window (e.g., there might be a
383 higher degree of plasticity for myelin-specific indexes in the early stages of life).
384 Additional research is needed in order to clarify which type of diffusion properties can
385 be shaped by experience as a function of age (e.g., neural or non-neuronal plasticity, intra
386 or extra axonal plasticity, [3,19,56]).
387 Finally, another potential explanation for our dMRI findings regards the presence of a
388 possible time shift between the training effects on behavior and brain structure, with
389 white matter changes happening over a larger temporal scale compared to behavioral
390 changes. For instance, our findings are still compatible with the idea that at this early age
391 the amount of training received is not sufficient to shape what will become the reading
392 circuit later on. Although some studies have shown no time lag between behavioral and
393 structural changes in response to a short reading intervention program [3–5,12], the exact

394 time course of reading-related neuroplasticity is still understudied and needs further
395 investigation.

396

397 Overall, this heterogeneous picture of findings on short-term reading-related structural
398 neuroplasticity highlights the need to better define the conditions under which white
399 matter can be shaped by experience. Several experiential and developmental factors
400 might modulate the degree of white matter plasticity exhibited in response to reading
401 training or intervention. Research evidence coming from other cognitive domains might
402 give us some insights on the critical constraining variables to be considered. For instance,
403 studies testing for the presence of a sensitive period of sensory and motor white matter
404 circuits suggest that the time onset of the environmental exposure is a key factor to
405 establish whether white matter structure is stable or plastic [57,58]. Other factors that
406 have been suggested to modulate the balance between structural plasticity and stability
407 are the type and the duration of experiential exposure [59–63] and the individual
408 cognitive health and lifestyle risk factors [61,64].

409

410 In conclusion, this randomized controlled trial highlights that a two-week literacy
411 training can cause fast behavioral changes in preschoolers' reading performance, without
412 being accompanied by fast FA and MD changes of the reading circuitry. These findings
413 highlight that rapid diffusion properties variations are not always observed in response

414 to short-term reading training and point to the need of specifying the conditions under
415 which white matter structure is plastic versus stable.

416 References

417 1. Dehaene S, Pegado F, Braga LW, Ventura P, Nunes Filho G, Jobert A, et al. How
418 learning to read changes the cortical networks for vision and language. *Science*.
419 2010;330: 1359–1364.

420 2. Brem S, Bach S, Kucian K, Kujala JV, Guttermann TK, Martin E, et al. Brain sensitivity to
421 print emerges when children learn letter–speech sound correspondences.
422 *Proceedings of the National Academy of Sciences*. 2010;107: 7939–7944.

423 3. Economou M, Bempp FV, Van Herck S, Wouters J, Ghesquière P, Vanderauwera J, et
424 al. Myelin plasticity during early literacy training in at-risk pre-readers. *Cortex*.
425 2023;167: 86–100.

426 4. Huber E, Donnelly PM, Rokem A, Yeatman JD. Rapid and widespread white matter
427 plasticity during an intensive reading intervention. *Nat Commun*. 2018;9: 2260.

428 5. Keller TA, Just MA. Altering cortical connectivity: remediation-induced changes in
429 the white matter of poor readers. *Neuron*. 2009;64: 624–631.

430 6. Chyl K, Fraga-González G, Brem S, Jednoróg K. Brain dynamics of (a)typical reading
431 development-a review of longitudinal studies. *NPJ Sci Learn*. 2021;6: 4.

432 7. Catani M, Mesulam M. The arcuate fasciculus and the disconnection theme in

433 language and aphasia: history and current state. *Cortex*. 2008;44: 953–961.

434 8. Herbet G, Zemmoura I, Duffau H. Functional Anatomy of the Inferior Longitudinal
435 Fasciculus: From Historical Reports to Current Hypotheses. *Front Neuroanat*.
436 2018;12: 77.

437 9. Hoeft F, McCandliss BD, Black JM, Gantman A, Zakerani N, Hulme C, et al. Neural
438 systems predicting long-term outcome in dyslexia. *Proc Natl Acad Sci U S A*.
439 2011;108: 361–366.

440 10. Vanderauwera J, Wouters J, Vandermosten M, Ghesquière P. Early dynamics of
441 white matter deficits in children developing dyslexia. *Dev Cogn Neurosci*. 2017;27:
442 69–77.

443 11. Wang Y, Mauer MV, Raney T, Peysakhovich B, Becker BLC, Sliva DD, et al.
444 Development of Tract-Specific White Matter Pathways During Early Reading
445 Development in At-Risk Children and Typical Controls. *Cereb Cortex*. 2017;27: 2469–
446 2485.

447 12. Meisler SL, Gabrieli JDE, Christodoulou JA. White matter microstructural plasticity
448 associated with educational intervention in reading disability. *Imaging*
449 *Neuroscience*. 2024;2: 1–18.

450 13. Yeatman JD, Dougherty RF, Ben-Shachar M, Wandell BA. Development of white

451 matter and reading skills. *Proc Natl Acad Sci U S A*. 2012;109: E3045–53.

452 14. Bonte M, Brem S. Unraveling individual differences in learning potential: A dynamic
453 framework for the case of reading development. *Dev Cogn Neurosci*. 2024;66:
454 101362.

455 15. Roy E, Richie-Halford A, Kruper J, Narayan M, Bloom D, Nedelec P, et al. White
456 matter and literacy: A dynamic system in flux. *Dev Cogn Neurosci*. 2024;65: 101341.

457 16. Lebel C, Benischek A, Geeraert B, Holahan J, Shaywitz S, Bakhshi K, et al.
458 Developmental trajectories of white matter structure in children with and without
459 reading impairments. *Dev Cogn Neurosci*. 2019;36: 100633.

460 17. Moulton E, Bouhali F, Monzalvo K, Poupon C, Zhang H, Dehaene S, et al.
461 Connectivity between the visual word form area and the parietal lobe improves after
462 the first year of reading instruction: a longitudinal MRI study in children. *Brain*
463 *Struct Funct*. 2019;224: 1519–1536.

464 18. Vanderauwera J, De Vos A, Forkel SJ, Catani M, Wouters J, Vandermosten M, et al.
465 Neural organization of ventral white matter tracts parallels the initial steps of
466 reading development: A DTI tractography study. *Brain Lang*. 2018;183: 32–40.

467 19. Huber E, Mezer A, Yeatman JD. Neurobiological underpinnings of rapid white
468 matter plasticity during intensive reading instruction. *Neuroimage*. 2021;243:

469 118453.

470 20. Partanen M, Kim DHC, Rauscher A, Siegel LS, Giaschi DE. White matter but not grey
471 matter predicts change in reading skills after intervention. *Dyslexia*. 2021;27: 224–
472 244.

473 21. Economou M, Van Herck S, Vanden Bempt F, Glatz T, Wouters J, Ghesquière P, et
474 al. Investigating the impact of early literacy training on white matter structure in
475 prereaders at risk for dyslexia. *Cereb Cortex*. 2022;32: 4684–4697.

476 22. Dunn LM, Dunn DM. Peabody Picture Vocabulary Test--Fourth Edition. PsycTESTS
477 Dataset. 2012. doi:10.1037/t15144-000

478 23. Slingerland B. Pre-reading screening procedures. Educators Publishing Service;
479 1968.

480 24. Williams K. Phonological and print awareness scale. Torrence, CA: Western
481 Psychological Services.

482 25. Invernizzi M, Meier JD. PALS 1-3: Phonological awareness literacy screening;
483 technical reference. Virginia State Department of Education; 2003.

484 26. Petersen DB, Spencer TD. The Narrative Language Measures: Tools for Language
485 Screening, Progress Monitoring, and Intervention Planning. 2012 [cited 10 Jan 2024].

486 doi:10.1044/lle19.4.119

487 27. Williams KT. Expressive Vocabulary Test Kit. American Guidance Services,

488 Incorporated; 1997.

489 28. Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, et al. N4ITK:

490 improved N3 bias correction. IEEE Trans Med Imaging. 2010;29: 1310–1320.

491 29. Fonov VS, Evans AC, McKinstry RC, Almlí CR, Collins DL. Unbiased nonlinear

492 average age-appropriate brain templates from birth to adulthood. Neuroimage.

493 2009;Supplement 1: S102.

494 30. Avants BB, Epstein CL, Grossman M, Gee JC. Symmetric diffeomorphic image

495 registration with cross-correlation: evaluating automated labeling of elderly and

496 neurodegenerative brain. Med Image Anal. 2008;12: 26–41.

497 31. Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden

498 Markov random field model and the expectation-maximization algorithm. IEEE

499 Trans Med Imaging. 2001;20: 45–57.

500 32. Cieslak M, Cook PA, He X, Yeh F-C, Dhollander T, Adebimpe A, et al. QSIprep: An

501 integrative platform for preprocessing and reconstructing diffusion MRI. Cold

502 Spring Harbor Laboratory. 2020. p. 2020.09.04.282269. doi:10.1101/2020.09.04.282269

503 33. Gorgolewski K, Burns CD, Madison C, Clark D, Halchenko YO, Waskom ML, et al.

504 Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing

505 Framework in Python. *Frontiers in Neuroinformatics*. 2011.

506 doi:10.3389/fninf.2011.00013

507 34. Gorgolewski KJ, Esteban O, Markiewicz CJ, Ziegler E, Ellis DG, Notter MP, et al.

508 Nipype. *Softw Pract Exp*. 2018.

509 35. Abraham A, Pedregosa F, Eickenberg M, Gervais P, Mueller A, Kossaifi J, et al.

510 Machine learning for neuroimaging with scikit-learn. *Front Neuroinform*. 2014;8: 14.

511 36. Garyfallidis E, Brett M, Amirbekian B, Rokem A, van der Walt S, Descoteaux M, et

512 al. Dipy, a library for the analysis of diffusion MRI data. *Front Neuroinform*. 2014;8:

513 8.

514 37. Andersson JLR, Sotropoulos SN. An integrated approach to correction for off-

515 resonance effects and subject movement in diffusion MR imaging. *Neuroimage*.

516 2016;125: 1063–1078.

517 38. Andersson JLR, Skare S, Ashburner J. How to correct susceptibility distortions in

518 spin-echo echo-planar images: application to diffusion tensor imaging. *Neuroimage*.

519 2003;20: 870–888.

520 39. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg

521 H, et al. Advances in functional and structural MR image analysis and
522 implementation as FSL. *Neuroimage*. 2004;23 Suppl 1: S208–19.

523 40. Andersson JLR, Graham MS, Zsoldos E, Sotiropoulos SN. Incorporating outlier
524 detection and replacement into a non-parametric framework for movement and
525 distortion correction of diffusion MR images. *Neuroimage*. 2016;141: 556–572.

526 41. Veraart J, Novikov DS, Christiaens D, Ades-Aron B, Sijbers J, Fieremans E. Denoising
527 of diffusion MRI using random matrix theory. *Neuroimage*. 2016;142: 394–406.

528 42. Tournier J-D, Calamante F, Gadian DG, Connelly A. Direct estimation of the fiber
529 orientation density function from diffusion-weighted MRI data using spherical
530 deconvolution. *Neuroimage*. 2004;23: 1176–1185.

531 43. Tournier J-D, Yeh C-H, Calamante F, Cho K-H, Connelly A, Lin C-P. Resolving
532 crossing fibres using constrained spherical deconvolution: validation using
533 diffusion-weighted imaging phantom data. *Neuroimage*. 2008;42: 617–625.

534 44. Dhollander T, Mito R, Raffelt D, Connelly A. Improved white matter response
535 function estimation for 3-tissue constrained spherical deconvolution. *Proc Intl Soc
536 Mag Reson Med*. 2019.

537 45. Dhollander T, Raffelt D, Connelly A. Unsupervised 3-tissue response function
538 estimation from single-shell or multi-shell diffusion MR data without a co-registered

539 T1 image. ISMRM Workshop on Breaking the Barriers of Diffusion MRI. 2016. p. 5.

540 46. Raffelt D, Dhollander T, Tournier J-D, Tabbara R, Smith RE, Pierre E, et al. Bias field
541 correction and intensity normalisation for quantitative analysis of apparent fibre
542 density. Proc Intl Soc Mag Reson Med. 2017. p. 3541.

543 47. Yeatman JD, Dougherty RF, Myall NJ, Wandell BA, Feldman HM. Tract profiles of
544 white matter properties: automating fiber-tract quantification. PLoS One. 2012;7:
545 e49790.

546 48. Kruper J, Yeatman JD, Richie-Halford A, Bloom D, Grotheer M, Caffarra S, et al.
547 Evaluating the reliability of human brain white matter tractometry. bioRxiv. 2021. p.
548 2021.02.24.432740. doi:10.1101/2021.02.24.432740

549 49. Henriques RN, Correia MM, Marrale M, Huber E, Kruper J, Koudoro S, et al.
550 Diffusional Kurtosis Imaging in the Diffusion Imaging in Python Project. Front Hum
551 Neurosci. 2021;15: 675433.

552 50. Jensen JH, Helpern JA, Ramani A, Lu H, Kaczynski K. Diffusional kurtosis imaging:
553 the quantification of non-gaussian water diffusion by means of magnetic resonance
554 imaging. Magn Reson Med. 2005;53: 1432–1440.

555 51. Kruper J, Yeatman JD, Richie-Halford A, Bloom D, Grotheer M, Caffarra S, et al.
556 Evaluating the reliability of human brain white matter tractometry.

557 doi:10.1101/2021.02.24.432740

558 52. Burns MK, Wagner D. Determining an effective intervention within a brief
559 experimental analysis for reading: A meta-analytic review. *School Psych Rev.*
560 2008;37: 126–136.

561 53. Gao T, Zhao J, Li X, Mao Y, Chen Q, Harrison SE. Impact of rapid reading skills
562 training on reading rate and reading achievement among primary school students
563 in China. *Educ Psychol Rev.* 2020;40: 42–61.

564 54. Gersten R, Haymond K, Newman-Gonchar R, Dimino J, Jayanthi M. Meta-Analysis
565 of the Impact of Reading Interventions for Students in the Primary Grades. *J Res
566 Educ Eff.* 2020;13: 401–427.

567 55. Scammacca NK, Roberts GJ, Cho E, Williams KJ, Roberts G, Vaughn SR, et al. A
568 Century of Progress: Reading Interventions for Students in Grades 4-12, 1914-2014.
569 *Rev Educ Res.* 2016;86: 756–800.

570 56. Yeatman JD, Huber E. Sensitive periods for white matter plasticity and reading
571 intervention. *bioRxiv.* 2019. p. 346759. doi:10.1101/346759

572 57. Li Y, Ding G, Booth JR, Huang R, Lv Y, Zang Y, et al. Sensitive period for white-
573 matter connectivity of superior temporal cortex in deaf people. *Hum Brain Mapp.*
574 2012;33: 349–359.

575 58. Steele CJ, Bailey JA, Zatorre RJ, Penhune VB. Early musical training and white-
576 matter plasticity in the corpus callosum: evidence for a sensitive period. *J Neurosci*.
577 2013;33: 1282–1290.

578 59. Bengtsson SL, Nagy Z, Skare S, Forsman L, Forssberg H, Ullén F. Extensive piano
579 practicing has regionally specific effects on white matter development. *Nat Neurosci*.
580 2005;8: 1148–1150.

581 60. Colcombe SJ, Erickson KI, Scalf PE, Kim JS, Prakash R, McAuley E, et al. Aerobic
582 exercise training increases brain volume in aging humans. *J Gerontol A Biol Sci Med*
583 *Sci*. 2006;61: 1166–1170.

584 61. Sampaio-Baptista C, Johansen-Berg H. White Matter Plasticity in the Adult Brain.
585 *Neuron*. 2017;96: 1239–1251.

586 62. Scholz J, Klein MC, Behrens TEJ, Johansen-Berg H. Training induces changes in
587 white-matter architecture. *Nat Neurosci*. 2009;12: 1370–1371.

588 63. Wandell BA, Smirnakis SM. Plasticity and stability of visual field maps in adult
589 primary visual cortex. *Nat Rev Neurosci*. 2009;10: 873–884.

590 64. Peel NM, McClure RJ, Bartlett HP. Behavioral determinants of healthy aging. *Am J*
591 *Prev Med*. 2005;28: 298–304.

592 **Supporting information**

593 **S1 Fig. Subject reliability of FA for each experimental group.**

594 **S2 Fig. Subject reliability of MD for each experimental group.**

595 **S3 Fig. Training-related dMRI changes of the left inferior fasciculus.**

596 First two columns: structural changes of the left ILF are shown for the Letter and

597 Language Training groups and each experimental session. The third column shows the

598 distribution of FA and MD changes (i.e. difference between the individual profiles

599 observed in the post and pre-training sessions) for each group.

RANDOMIZED CONTROLLED TRIAL

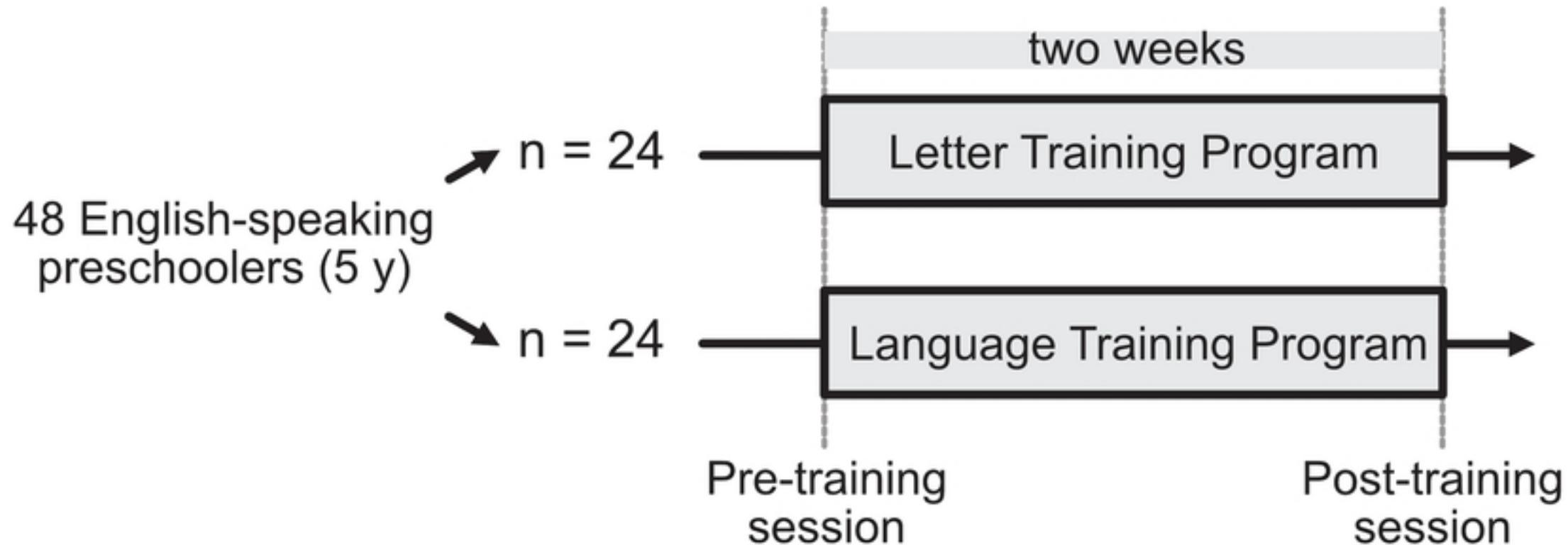


Fig 1

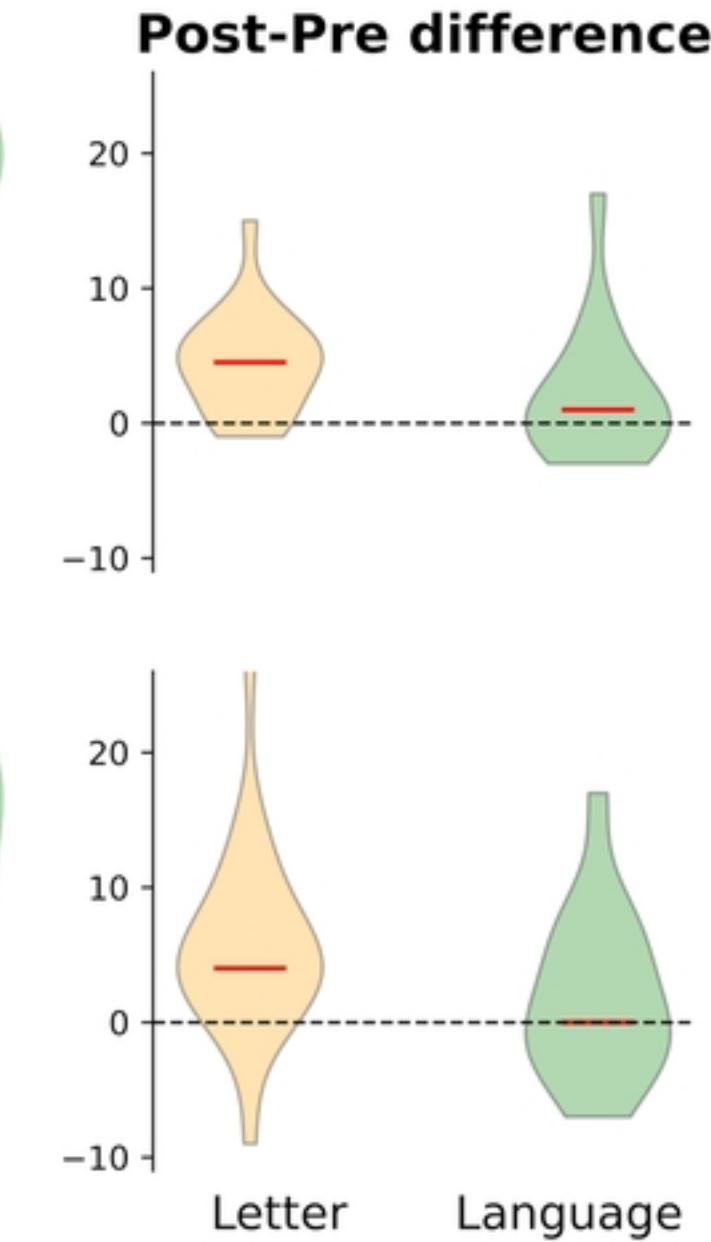
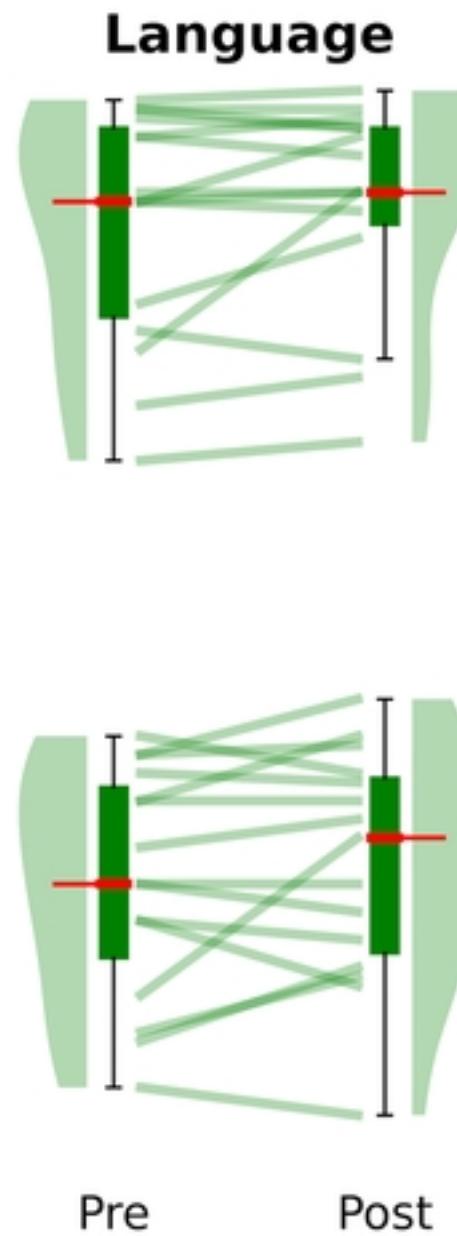
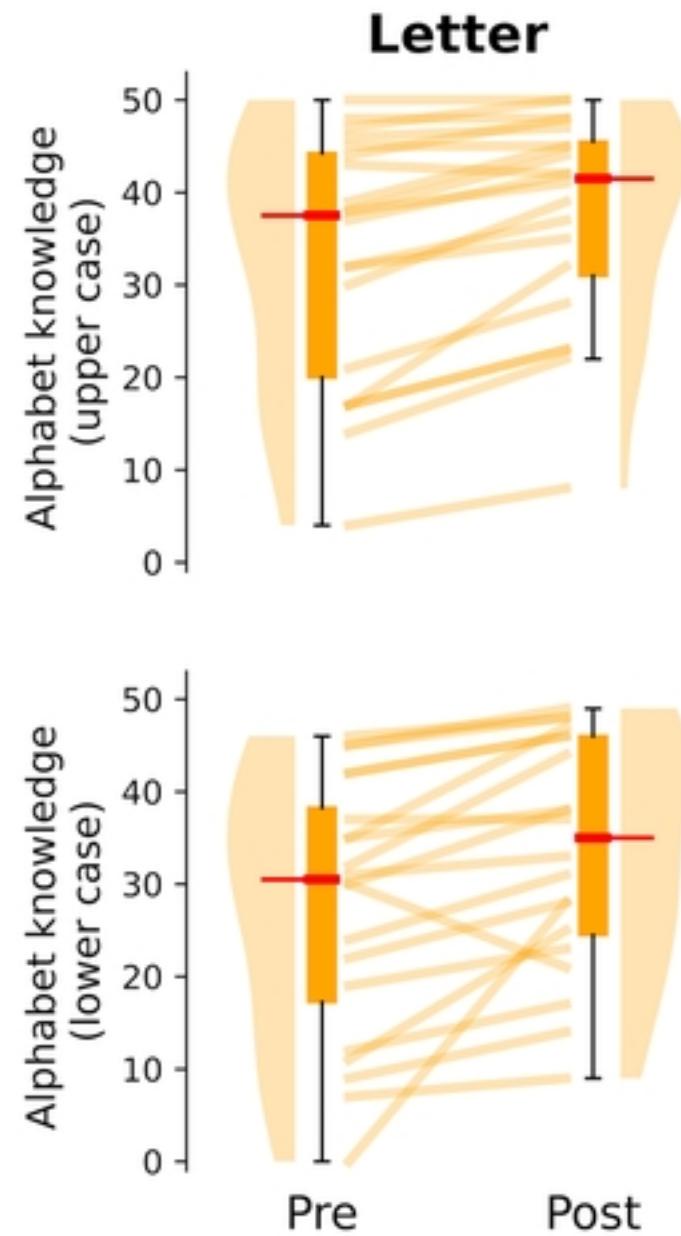


Fig 2

Test-Retest Reliability

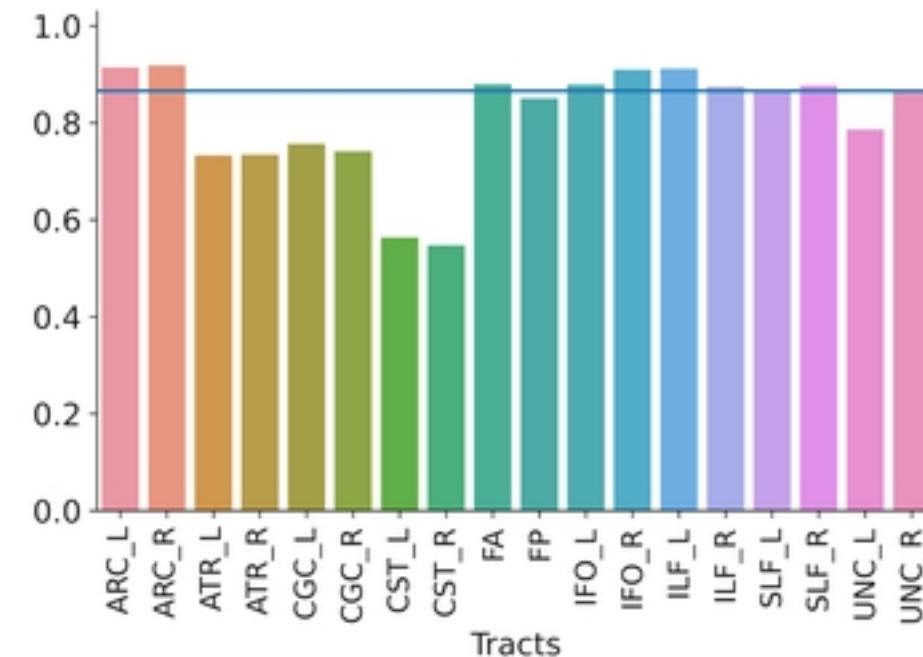
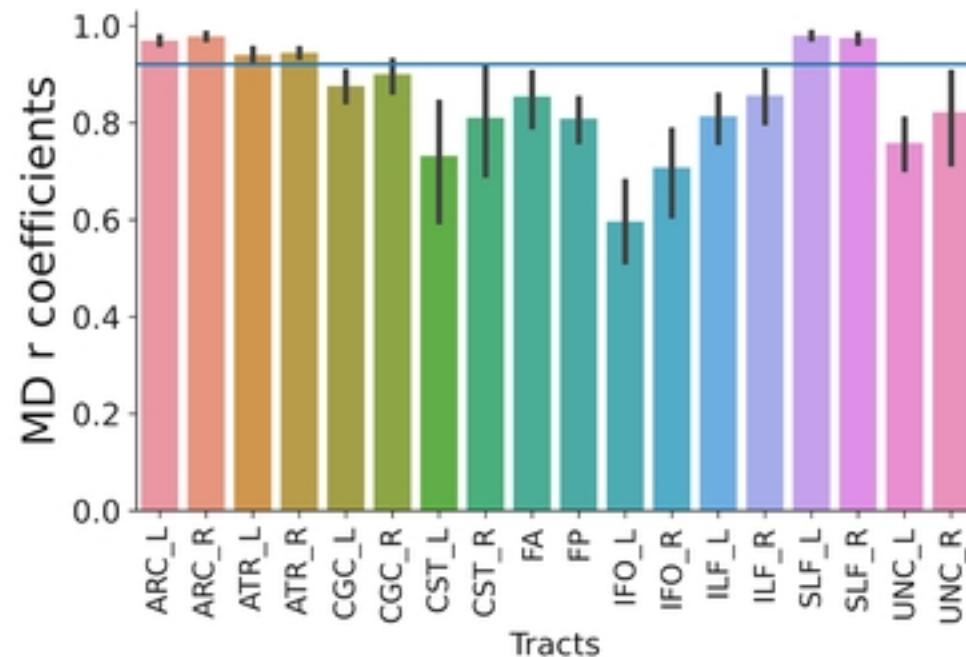
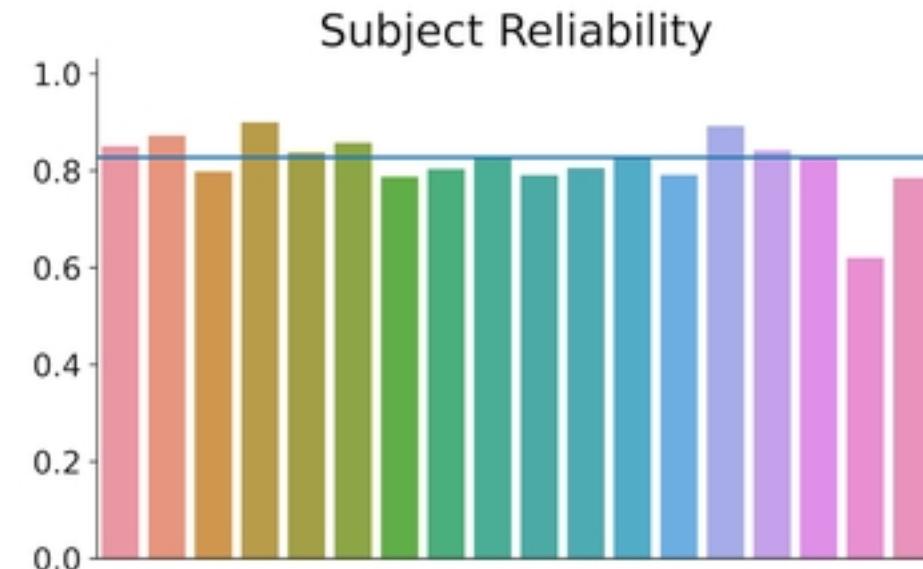
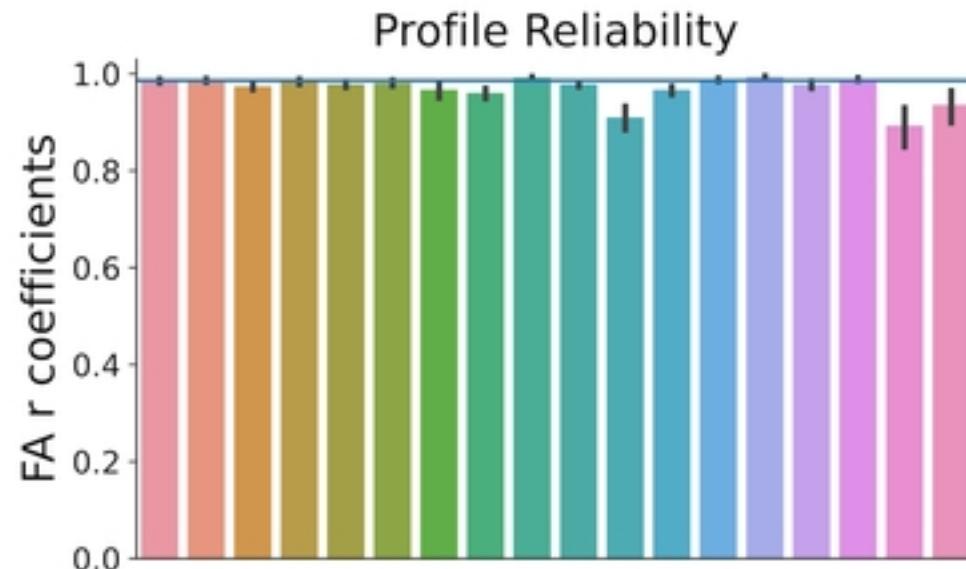


Fig 3

Left Arcuate - dMRI results

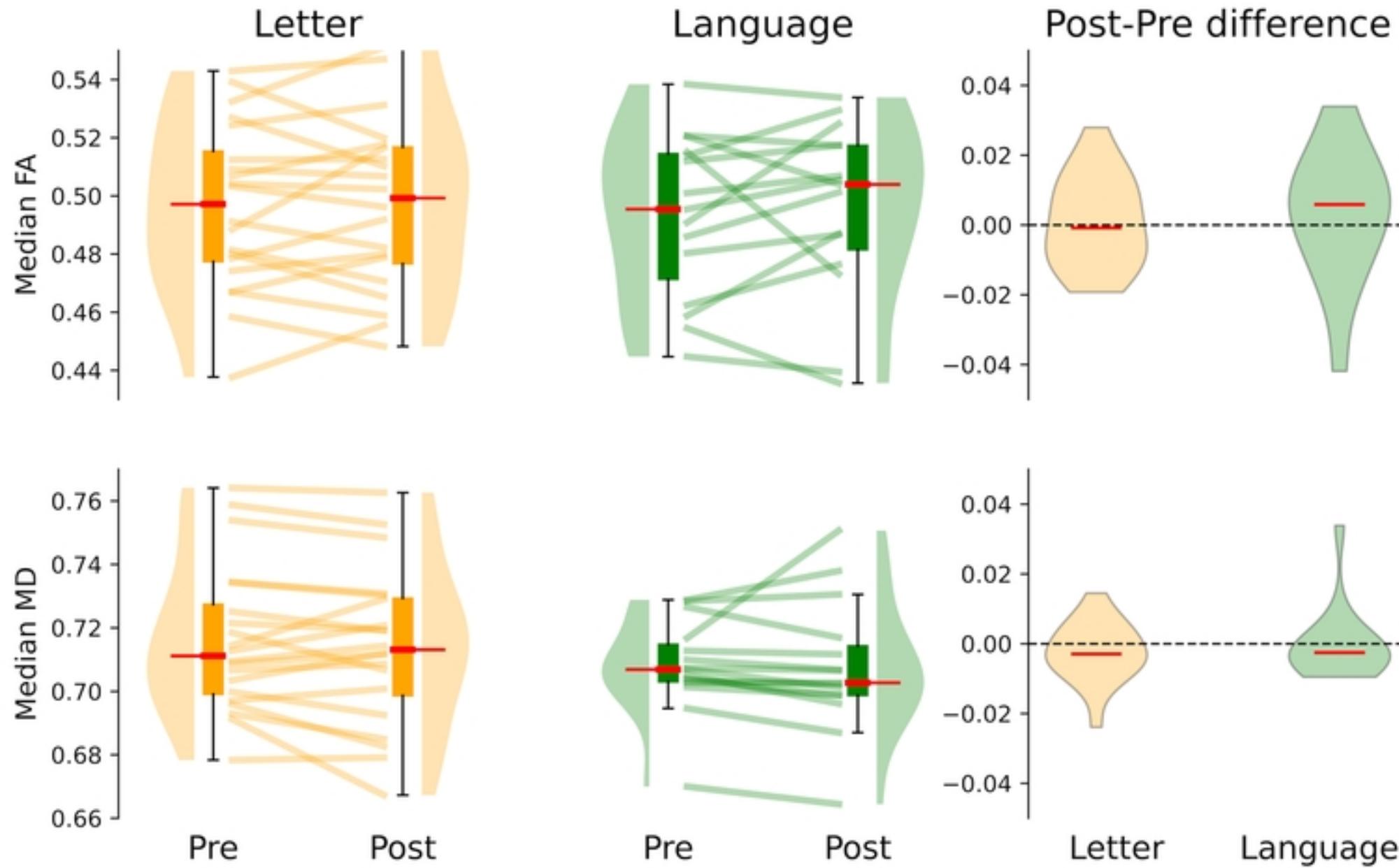


Fig 4

FA Tract profile comparisons

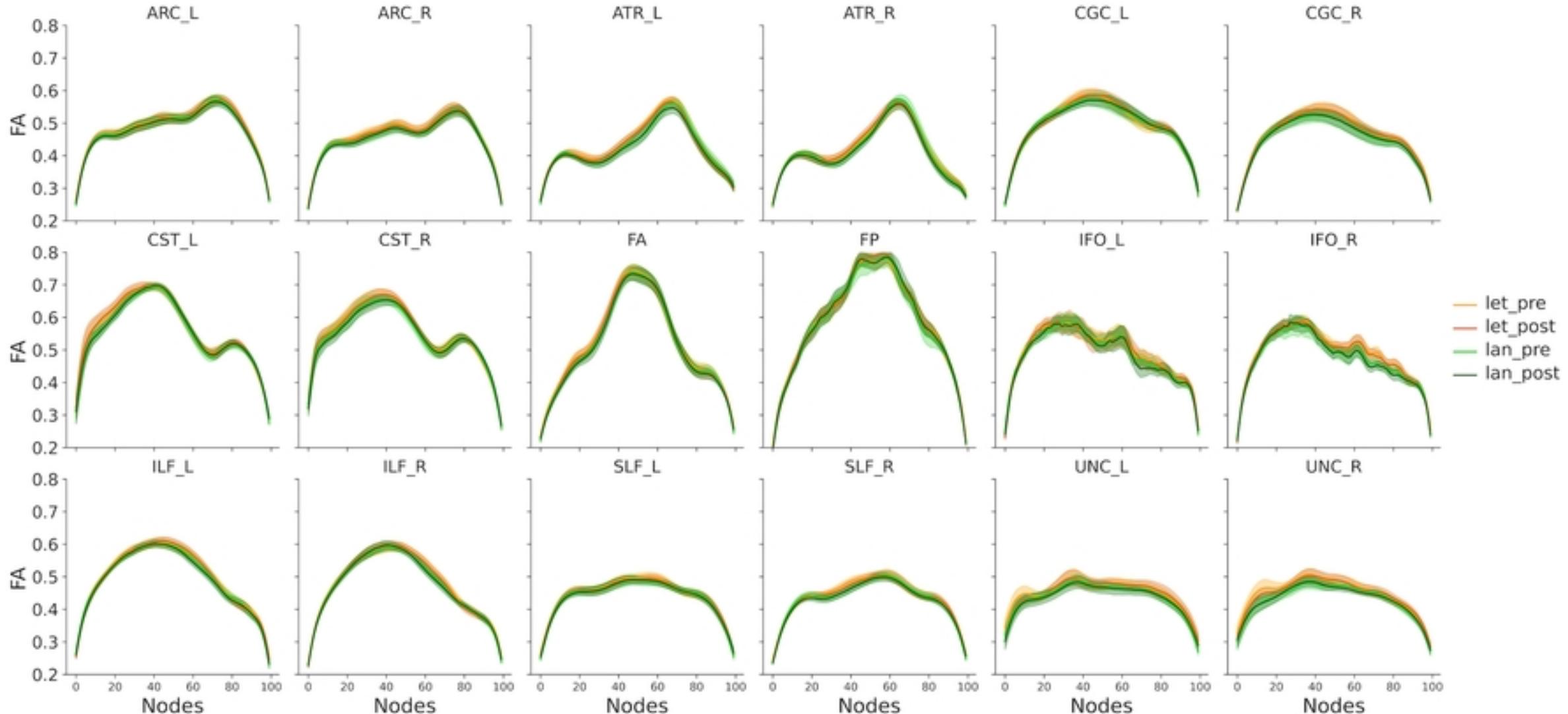


Fig 5

MD Tract profile comparisons

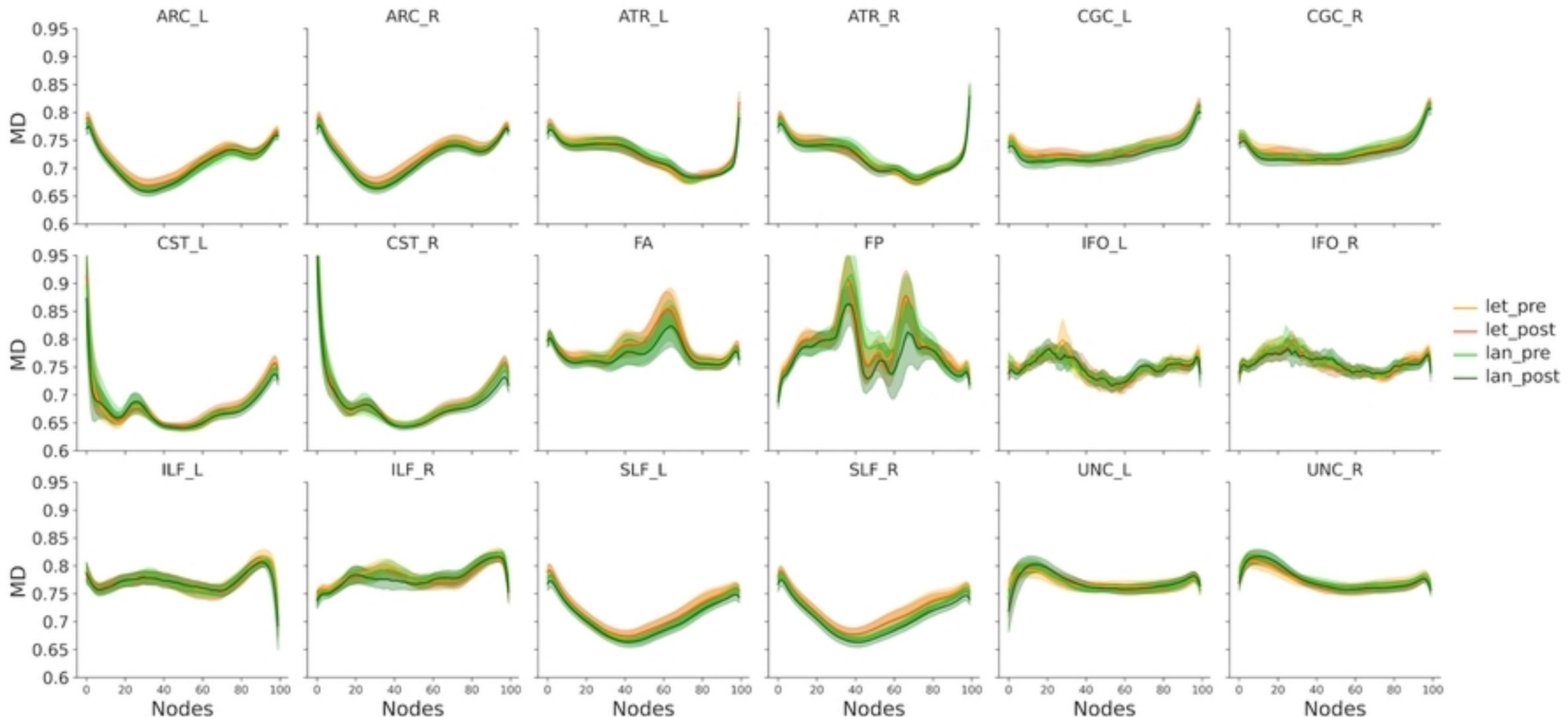


Fig 6

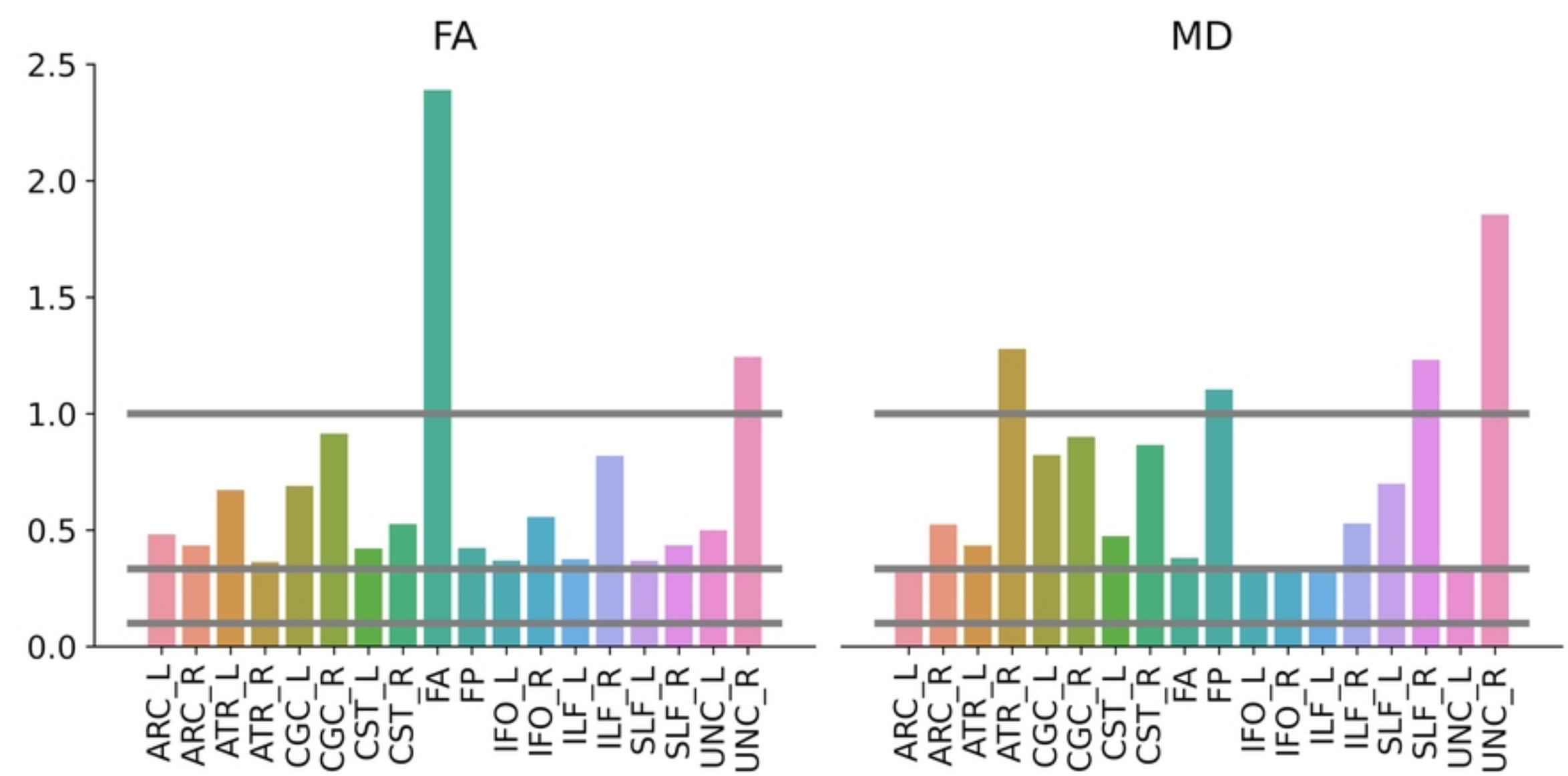


Fig 7