Neuron
Clinical Study

# Altering Cortical Connectivity: Remediation-Induced Changes in the White Matter of Poor Readers

Timothy A. Keller[1,*] and Marcel Adam Just[1]
[1]Center for Cognitive Brain Imaging, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213, USA
*Correspondence: tk37@andrew.cmu.edu
DOI 10.1016/j.neuron.2009.10.018

## SUMMARY

**Neuroimaging studies using diffusion tensor imaging (DTI) have revealed regions of cerebral white matter with decreased microstructural organization (lower-fractional anisotropy or FA) among poor readers. We examined whether 100 hr of intensive remedial instruction affected the white matter of 8- to 10-year-old poor readers. Prior to instruction, poor readers had significantly lower FA than good readers in a region of the left anterior centrum semiovale. The instruction resulted in a change in white matter (significantly increased FA), and in the very same region. The FA increase was correlated with a decrease in radial diffusivity (but not with a change in axial diffusivity), suggesting that myelination had increased. Furthermore, the FA increase was correlated with improvement in phonological decoding ability, clarifying the cognitive locus of the effect. The results demonstrate the capability of a behavioral intervention to bring about a positive change in cortico-cortical white matter tracts.**

## INTRODUCTION

A major challenge of cognitive neuroscience is to understand how changes in the structural properties of the brain underpin the plasticity exhibited whenever a person develops, ages, learns a new skill, or adapts to a neuropathology. Longitudinal studies have shown regional changes in the volume of gray matter that co-occur with skill acquisition or learning (Draganski et al., 2004, 2006), but there is also growing acknowledgment that higher-level cognition is based on cofunctioning of a set of cortical areas in a dynamic large-scale network, highlighting the central role of cortical communication. Improved anatomical connectivity in motor tracts as measured by fractional anisotropy (FA) has been associated with enriched experience (extensive childhood piano practice) in a correlational study using diffusion tensor imaging (DTI) (Bengtsson et al., 2005). (FA, which measures the anisotropy of the diffusion of water molecules (Basser and Pierpaoli, 1996), is sensitive to axonal density, size, myelination, and the coherence of organization of fibers within a voxel, and thus provides an index of the structural integrity of white matter).

Functional imaging studies have consistently demonstrated that children with reading disability display under-activation of a network of left-lateralized areas during reading, including occipito-temporal, temporo-parietal, and inferior frontal cortical regions (Hoeft et al., 2006, 2007; Meyler et al., 2007; Shaywitz et al., 2002; Simos et al., 2000a, 2000b), and that effective remedial reading interventions lead to increases in the activation in these same areas (Aylward et al., 2003; Meyler et al., 2008; Shaywitz et al., 2004; Simos et al., 2002; Temple et al., 2003), indicating that effective remediation can lead to a change in the brain functioning of poor readers. However, skilled reading depends not only on the activation of a set of relevant cortical areas, but also on communication among them. Reading difficulty has also been associated with lower functional connectivity (the synchronization of neural activity) across areas of the reading cortical network (Hampson et al., 2006; Horwitz et al., 1998; Pugh et al., 2000). This suggests that reading disability might be associated with structural properties of the white matter that provides the anatomical connectivity among the individual nodes of the reading network. Consistent with this view, several DTI studies of poor readers have found white matter regions with lower FA compared with controls (Beaulieu et al., 2005; Deutsch et al., 2005; Klingberg et al., 2000; Niogi and McCandliss, 2006; Odegard et al., 2009; Richards et al., 2008; Rollins et al., 2009). FA may be reduced in poor readers due to a number of possible differences in the microstructural properties of white matter, including reduced myelination, reduced axonal packing density, decreased axonal diameter, or reduced coherence of the orientation of axons within the region (Beaulieu, 2002; Ben-Shachar et al., 2007), all of which might impact the efficiency of communication (bandwidth) among cortical areas.

Here we report a longitudinal DTI study indicating that intensive remedial reading instruction (approximately 100 hr) can change the structural integrity of the cortical white matter of children who are poor readers. The children's DTI data were first assessed before instruction began and then a second time after the instruction ended, approximately 6 months later. At the preremediation scan, the poor readers showed significantly reduced fractional anisotropy (FA) in the anterior left centrum semiovale region, relative to a control group of good readers. Subsequent to the instruction, the remediated poor readers had not only made substantial gains in their reading ability, but

 

**Table 1. Changes in Age-Standardized Woodcock Reading Mastery Test—Revised Scores between the Preremediation and Postremediation Scans**

| | Change in Scores Group | | | | Group × Time ANOVA |
|---|---|---|---|---|---|
| | Poor Readers (PR) | | Poor Reader Controls (PC) | | Interaction Effect |
| Measure | Time 2 − Time1 | t(34) | Time 2 − Time1 | t(11) | $F_{(1, 45)}$ |
| WRMT-R word attack | 5.5 | 3.98[a] | −0.3 | −0.17 | 5.22[b] |
| WRMT-R word identification | 2.5 | 2.50[b] | 2.3 | 0.65 | 0.01 |
| WRMT-R passage comprehension | 1.1 | 0.86 | −3.0 | −1.04 | 2.18 |
| WRMT-R basic skills cluster | 4.2 | 5.06[c] | 2.6 | 0.87 | 0.48 |
| WRMT-R total reading cluster | 2.2 | 2.51 [b] | 1.5 | 0.47 | 0.08 |

[a] $p < 0.005$.
[b] $p < 0.05$.
[c] $p < 0.0005$.

also showed significantly increased FA in the anterior left centrum semiovale, in contrast to good readers and to a control group of untreated poor readers.

To help determine which microstructural properties had changed during remediation, we also examined the diffusivity in directions that are perpendicular to the principal axis of diffusion in anisotropic regions of white matter (radial diffusivity, $(\lambda_2 + \lambda_3)/2$), or parallel to it (axial diffusivity, $\lambda_1$). For example, changes in radial diffusivity in the absence of changes in axial diffusivity have been associated with changes in myelin (Beaulieu, 2002; Song et al., 2002, 2005), whereas changes in axial diffusivity in the absence of changes in radial diffusivity have been associated with an increase in axon diameter (Dougherty et al., 2007; but also see Wheeler-Kingshott and Cercignani [2009] for caveats about these measures). The results analyzed this way indicate that a behavioral intervention can bring about a positive change in the microstructure of human cortico-cortical white matter tracts, demonstrating the malleability of the anatomical connectivity that supports human cortical network function.

## RESULTS

Forty-seven children (8–12 years old) who were poor readers were randomly assigned to either an intensive 100 hr program of systematic and explicit remedial reading instruction focused primarily on developing word-level decoding skills (n = 35), or they were assigned to a control group that received normal classroom instruction (n = 12). There was also a control group of good readers (n = 25) of the same age. The remedial instruction was distributed over about 6 months of schooling, with instruction occurring in groups of three children with one teacher. (Although the remedial instruction came in one of four alternative forms (see Experimental Procedures), there were no reliable differences among the children assigned to the different forms in either initial behavioral measures or DTI measures, nor in the impacts of the instruction (see Supplemental Results and Discussion, available online). Hence the data reported here are collapsed across the children in the four forms of remedial reading instruction.) The remediated and unremediated poor readers scored equivalently at the preinstruction scan on multiple measures of reading ability, whereas the group of good readers scored significantly better than both groups of poor readers on every reading ability measure (see Table S1). The behavioral results indicated that the poor readers who received the remedial instruction showed significant improvement on most of the age-standardized Woodcock Reading Mastery Test – Revised (WRMT-R, Woodcock et al., 1998) reading ability measures when retested following the instruction period, but that the control poor readers did not show improvement on these measures, indicated by a reliable overall group by time effect ($F_{1, 45} = 4.36$, $p < 0.05$), with means shown in Table 1. Individual ANOVAs for each measure indicated that the interaction between group and time was reliable only for the subtest measuring non-word reading ability (Word Attack scores, $F_{1, 45} = 5.22$, $p < 0.05$), but not for the subtests measuring real word reading ability (Word Identification) or passage comprehension ability (Passage Comprehension). This pattern of outcomes suggests that the instruction specifically improved phonological decoding skills more than the standard reading curricula did. This conclusion was also supported by an analysis of changes in raw scores on all ability measures collected from the poor readers before and after the treatment phase (see Supplemental Results and Discussion and Table S2).

The DTI results indicated that poor readers who received the remedial instruction showed a reliable increase in FA between the preremediation and postremediation scans, with a peak difference in the left anterior centrum semiovale, as shown in Figure 1A. Corresponding contrasts conducted for the two control groups that received no remedial instruction found no areas showing either an increase or decrease in FA between the two scans, indicating that the change in FA among the remediated poor readers was not due to maturational changes over the 6 month interval between the two scans. This same region also showed significantly reduced FA at the preremediation scan among all poor readers relative to the group of good readers (Figure 1B). The reliable increase in FA between the two scans among the poor readers, but no change in FA between

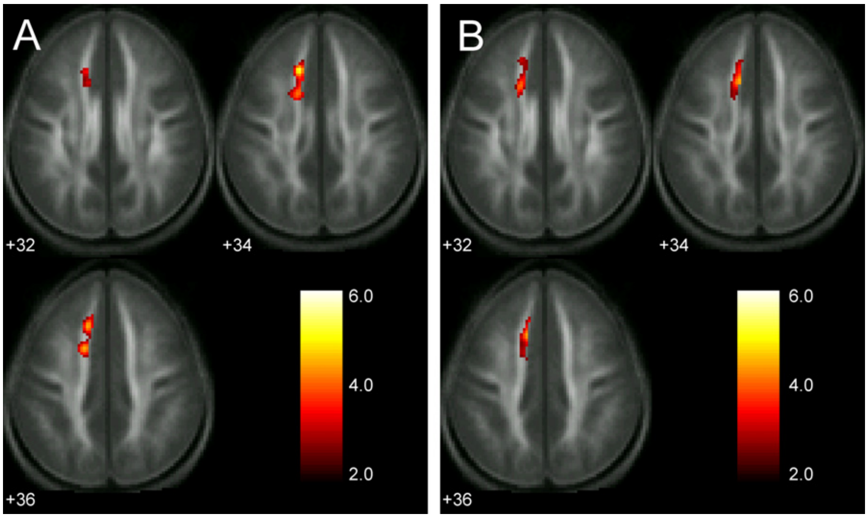


**Figure 1. Fractional Anisotropy Increases following Remediation in Poor Readers in the Same Region of the Left Anterior Centrum Semiovale that Showed Reduced Fractional Anisotropy Relative to Good Readers Prior to the Instruction**

(A) Region where the poor reader group showed an increase in FA between the preremediation and postremediation scans (peak $t(34) = 5.12$, at Montreal Neurological Institute [MNI] coordinates $-12$ 28 36, spatial extent = 450 voxels, $p < 0.05$ corrected for multiple comparisons). There were no areas where poor readers showed a decrease in FA between phases, nor were there any areas where the control group of good readers or the control group of unremediated poor readers showed either an increase or decrease in FA.

(B) Region showing a significant difference in FA between good readers and all poor readers at the first scan (peak $t(70) = 4.66$, at MNI coordinates $-10$ 20 38, spatial extent = 418 voxels, $p < 0.05$ corrected for multiple comparisons). Statistical maps are overlaid on a normalized FA image averaged across all participants and both scans. The MNI z-coordinate is shown at the bottom left of each axial slice. Color scale represents t values.

the scans among the good readers, nor among the unremediated poor reader controls, resulted in a significant group (3) by time (2) interaction with a peak F value in the same region of the left anterior centrum semiovale (Figure 2A), strongly suggesting that intensive remedial reading instruction led to changes in some microstructural property of white matter in a region of left frontal white matter, a region that differed between good and poor readers prior to the treatment. Additional analyses presented in the Supplemental Results confirmed that these findings were not due to the particular voxel-based analysis methods that were used; essentially identical results were obtained using unsmoothed data and nonparametric statistical inference methods (see Supplemental Results and Discussion and Figures S1 and S2).

Because increased FA in highly organized white matter can occur due to either a relative decrease in radial diffusivity or a relative increase in axial diffusivity (or both), a further analysis examined the remediation effect in each of these components separately in the region shown in Figure 2A. It was the radial diffusivity that had changed in the remediated poor readers subsequent to the instruction. There was a reliable group by time interaction for radial diffusivity in this same region ($F_{2,\,69} = 5.92$, $p < 0.005$); this measure reliably decreased among the remediated poor readers ($t(34) = 3.98$, $p < 0.0005$), but showed no change in either the good readers or the poor reader controls, as shown in Figure 2C. This pattern of radial diffusivity effects mirrors the findings for FA (a reliable increase in FA among poor readers who received remedial instruction but no reliable change in FA among the two unremediated groups; Figure 2B). By contrast, the other component of FA, axial diffusivity, showed no significant changes between phases for any group at this location, nor was there a reliable interaction (Figure 2D). The pattern of diffusivity effects indicates that the difference in FA between poor and good readers before remediation is due to initially higher radial diffusivity in the poor readers, and that the change in FA results from a change in some microstructural feature (e.g., myelination, packing density, or axon diameter) that affects radial diffusivity. The pattern of results also argues against the preremediation differences in FA between good and poor readers being due to the existence of more crossing fibers or smaller diameter axons in the poor readers in the area, and argues against the proposition that the changes in FA resulting from remediation were due to changes in either of these microstructural features, both of which would be expected to affect axial diffusivity. This leaves increased myelination as a plausible mechanism of the microstructural change.

The findings of increased reading ability and increased FA strongly suggest that the remedial instruction brought about a change in both variables, but say little about the relation between the two variables. To investigate this relation in more detail and to assess which aspects of reading ability were associated with increased FA, an exploratory stepwise hierarchical multiple regression analysis examined how well the *change* in raw reading scores of an individual poor reader could account for that individual's *change* in FA in the region. This analysis (which also took the change in age between scans into account) indicated that a model including the change in raw scores on two subtests from the Test of Word Reading Efficiency (TOWRE, Torgesen et al., 1999) provided the best fit to the change in FA data among poor readers ($R^2 = 0.10$, $F_{2,\,43} = 2.36$, $p = 0.11$). The change in Phonemic Decoding Efficiency (PDE, a measure of non-word reading fluency similar to the WRMT-R WA subtest) was positively associated with change in FA (partial $r = 0.23$, $p = 0.06$). In contrast, the change in the Sight-Word Efficiency (SWE, a measure of real word reading fluency similar to the WRMT-R Word ID subtest) showed a negative partial correlation with change in FA ($pr = -0.21$). No other variables met the criteria for entry into the model. An identical analysis conducted with radial diffusivity in the region as the dependent measure also showed that these same two measures provided the best fit to the data ($R^2 = 0.13$, $F_{2,\,43} = 3.41$, $p < 0.05$) with change in PDE significantly negatively associated with

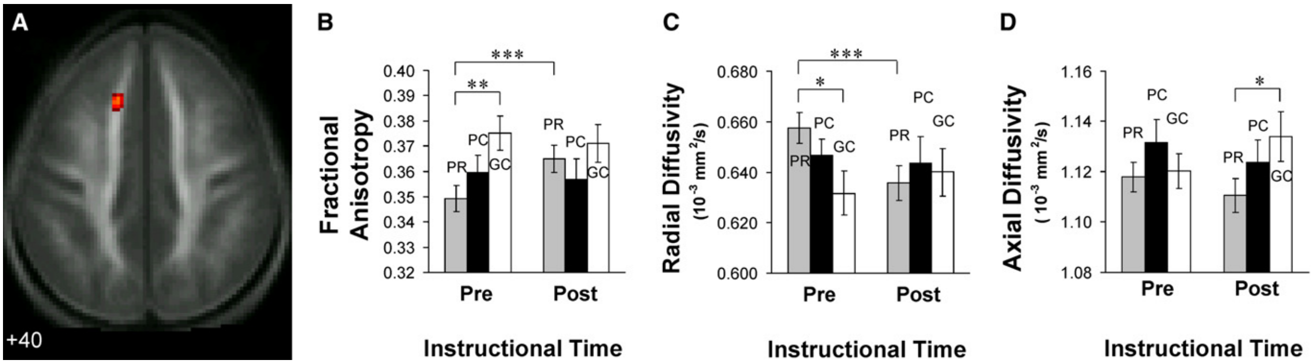


**Figure 2. Differential Changes in Fractional Anisotropy as a Function of Group Are Due to Differences in Diffusivity Perpendicular to the Principal Diffusion Direction**

(A) Location of the cluster of voxels with the maximum F value (peak $F_{2, 69} = 9.66$, spatial extent = 49 voxels, $p < 0.0005$ uncorrected, at MNI coordinates −12 26 40) for a test of the group by time interaction.

(B) Mean FA for this cluster in each group at each phase of the study.

(C) Mean radial diffusivity for this cluster in each group and at each phase.

(D) Mean axial diffusivity for this cluster in each group and at each phase. Error bars represent the standard error of the mean. PR = poor readers who received remediation, PC = poor reader control group, GC = good reader control group. *$p < 0.05$, **$p < 0.005$, ***$p < 0.0005$.

change in radial diffusivity ($pr = -0.23$, $p < 0.05$) and change in SWE positively associated with the change ($pr = 0.29$). In contrast, for axial diffusivity, an identical stepwise regression analysis indicated that no change in any of the raw ability measures explained enough variance for entry into the model (p's > 0.15). The outcome of these analyses indicate that there is a coupling between the behavioral change in reading and the anatomical change measured by FA and radial diffusivity, and indicates that increased phonological decoding ability provides the best predictor of increased FA and decreased radial diffusivity.

These results and conclusions are further supported by additional analyses (described in the Supplemental Results and Discussion) of the relationships between individual differences in various reading abilities and various diffusion tensor measures in the entire sample of good and poor readers, (using reading and diffusion measures obtained *prior* to the remediation), in the cluster that eventually showed an increase in FA among the remediated poor readers. Multiple regression analyses indicated that individual differences in phonological decoding ability (as measured by WRMT-R WA scores) were strongly positively related to FA (see Figure S3), strongly negatively related to radial diffusivity, and only weakly negatively related to axial diffusivity at the time of the preremediation scan (see Figure S4). These findings suggest that radial diffusivity drives the positive relationship between FA and individual differences in reading ability measured at the initial scan. In addition, both FA and radial diffusivity were more strongly related to Word Attack scores than to Word ID scores, suggesting that connections passing through the cluster area may be more important for phonological processing than for direct access to meaning via a direct orthographic route (see Supplemental Results and Discussion).

To determine the orientation of the tracts showing the remediation-related change and to identify the cortical areas that they likely connect, fiber tractography was carried out on group-averaged diffusion tensor data, using as a seed region the cluster of voxels showing a reliable group by time interaction. These group-averaged tracts were remarkably similar in their gross morphology between the good and poor readers and also across the two scanning sessions, as shown in Figure 3A, indicating the reliability of the data and the tracking methods. The principal direction of diffusion in the region showing a group difference in FA at the preremediation scan remained the same at the follow-up scan, and the fibers identified as passing through the regions were remarkably consistent between the two scans for both groups of subjects, suggesting that microstructural changes in the white matter within the region, rather than changes in the orientation of fibers, are responsible for the remediation effect and for the relationship of reading ability to the diffusion measures. The principal diffusion direction was anterior-posterior in all groups, and fibers passing through this region extended anteriorly and medially to terminate in a medial region of the superior frontal gyrus (Figure 3B) and extended posteriorly and superiorly to terminate in the left paracentral lobule (Figure 3C).

To check for consistency with previous DTI studies of white matter abnormalities among poorer readers in a left temporo-parietal region (Beaulieu et al., 2005; Deutsch et al., 2005; Klingberg et al., 2000), we tested for group differences and a remediation effect in this region that had shown a relation to reading ability in these previous studies. Although there were no statistically reliable effects in the voxel-wise analyses, the FA was reliably lower among poor readers at the initial scan when the average FA across the entire region of interest was examined and the specific analysis was closely matched to those previous studies. A review of diffusion studies of this region (Ben-Shachar et al., 2007) suggests that the reduced FA among poorer readers is probably due to increased fiber crossings, and if this is indeed case, then intensive reading remediation would not be expected to change the coherence or the orientation of the fibers. Consistent with this expectation, there was no remediation effect in the region (see Supplementary Results and Discussion).

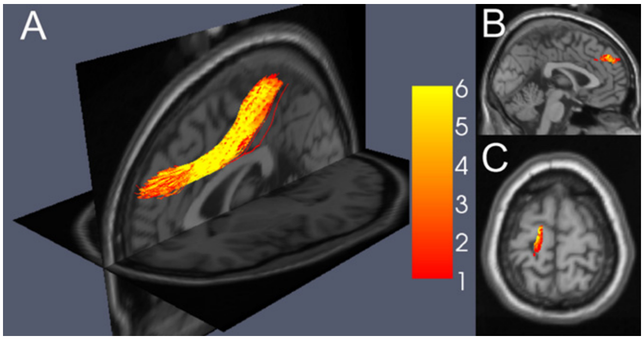


**Figure 3. Similarity of Estimated Fiber Orientation and Location across Groups**
(A) Consistency of the group-averaged tractography for PR, PC, and GC groups at each of two scans, using a seed region based on the cluster in Figure 2A. Color scale indicates the consistency of the tracking across the groups and phases, with red indicating voxels entered by only one of the groups at one scan, and yellow indicating voxels entered by all three groups at both scans.
(B) Location of the anterior termination of the estimated fibers in the medial superior frontal gyrus.
(C) Location of the posterior termination of the estimated fibers in the left paracentral lobule.

## DISCUSSION

The finding of longitudinally measured, experimentally mediated changes in the structural properties of left hemisphere white matter in children with reading problems reveals the considerable potential of behavioral remediation and rehabilitation programs, and furthers the understanding of reading disability and brain plasticity. The most important finding is clearly that both reading ability and the structural integrity of left hemisphere white matter can be increased by extensive practice in word decoding skills. This finding suggests that whatever the cause of abnormally low FA among poor readers may be, the abnormality is amenable to behavioral treatment when provided within an age window in which maturation, experience, and development are still capable of influencing FA.

The precise microstructural properties underlying both the initial group differences in FA and radial diffusivity and the remediation-related changes in these measures may be identifiable by further research. Among the factors influencing radial diffusivity are myelination, axonal packing density, and axon diameter (Beaulieu, 2002). One reason that myelination is a particularly attractive potential mechanism for future exploration is that myelin is known to affect primarily radial diffusivity (Song et al., 2002; Song et al., 2005). In addition, neuronal firing has been shown to affect myelination in central nervous system axons (Demerens et al., 1996; Ishibashi et al., 2006; Stevens et al., 2002). Although it is unknown whether such a mechanism could increase myelination in humans at the ages examined in the current study, it is possible that intensive training in word-decoding skills increases the communication among left hemisphere cortical areas, which may in turn increase the myelination along the axons connecting these regions, decrease radial diffusivity along these axons, and increase FA. Methods exist for investigating this hypothesis concerning the role of myelination in the remediation effect using techniques such as magnetization transfer or T2 relaxation imaging for directly measuring myelin content.

It is tempting to ask about the causal directionality between the reading effects and the diffusion effects: does an increase in the efficiency of neural transmission resulting from remediation produce an increase in phonological decoding ability, or does increased phonological decoding ability produce increased reading behavior and consequent increases in the efficiency of the neural transmission? Both alternatives are possible, but it is also possible that the two types of changes develop interactively, as one might expect in a dynamic system like the brain. If the latter is the case, then it may be more fruitful to investigate factors that can accelerate or more finely control both the neuroplastic changes in white matter and the changes in reading processes, rather than attempting to determine the casual directionality.

The functional role in the reading process of the modified left anterior centrum semiovale white matter is not well understood, but it may pertain to the control processes of reading, rather than to word decoding itself. Activation in the left medial superior frontal gyrus occurs in normal children when processing orthographic and phonological forms of stimuli that are mutually inconsistent (Bitan et al., 2007), suggesting a response selection role for this area that may have been repeatedly evoked in the remedial phonological decoding tasks. The paracentral lobule has been found to activate more to phonologically dissimilar items than to similar items in a verbal memory task in adults (Sweet et al., 2008). Another control function associated with the paracentral lobule is as hub controller in the "structural core" of cortico-cortical axonal communication pathways (Hagmann et al., 2008), the nodes of which correspond to the "default mode" network (Raichle et al., 2001). It is possible that the repeated phonological processing in the remediation strengthened inhibitory connections between the paracentral lobule and medial frontal cortex, leading to reduced default network activity during reading. Although our findings do not illuminate the roles of the areas whose connectivity was improved, they nevertheless establish a structural change that could only have been brought about because of changes in activity in these areas or in secondarily connected areas.

The methodological question of how to accurately align the data from different participants for group analysis remains a topic of interest because of inherent limitations in regularizing unsystematic morphological variation. The limitations of the voxel-based approach used in the current study lie in its dependence on the accuracy of the coregistration algorithm and the amount of smoothing subsequently applied to the data to compensate for the inaccuracy (Jones et al., 2005; Smith et al., 2006). Recently developed alternative methods that attempt to avoid these particular concerns (Lee et al., 2009; Oakes et al., 2007; Smith et al., 2006) merit further evaluation, which is beyond the scope of the present paper. To address these limitations of the voxel-based approach, we have demonstrated that the main conclusions of the present study are also supported by analyses that do not use spatial smoothing or parametric assumptions (see Supplemental Results and Discussion and Figures S1 and S2).

The capability to improve white matter provides a possible remediation not only for reading difficulty but also for other neurological abnormalities believed to be underpinned by deficits in anatomical connectivity, such as autism (Just et al., 2007). Although the basic computing power of the brain surely lies in individual neurons, it is only their *collective* action, made possible by white matter connectivity, that enables the multicentered large-scale brain networks that characterize human thought. For this reason, modest modifications in white matter may enable major changes in cognitive ability.

## EXPERIMENTAL PROCEDURES

### Participants

Seventy-two participants were included in the analyses (35 poor readers that received the treatment, 12 poor readers that did not receive the treatment, and 25 good readers that did not receive the treatment). They were selected from a larger sample on the basis of their having provided functional and behavioral data used in an fMRI study of sentence comprehension (Meyler et al., 2007), and on their having artifact-free DTI data at both the preremediation and postremediation phases. The children gave verbal informed consent in the presence of a parent or guardian, who gave signed informed consent. The children were paid for their participation. A parent questionnaire was used to verify that all participants met inclusion criteria. All protocols were approved by the University of Pittsburgh and Carnegie Mellon University Institutional Review Boards.

The participants were all right-handed, native English-speaking children, with normal vision and hearing. Children were excluded from the study if they had brain injury, sensory disorders, psychiatric disorders, attention deficit disorder, metal in their bodies, were on medication, or were claustrophobic. The poor readers were participants in the Power4Kids Reading Initiative, a randomized-trial field study of remedial instruction for children with reading difficulties varying in severity (Torgesen et al., 2006). Criteria for inclusion in the project were a score at or below the $30^{th}$ percentile on the combination of the sight word efficiency and phonological decoding subtests of the Test of Word Reading Efficiency (Torgesen et al., 1999) during its initial administration, and a score at or above the $5^{th}$ percentile on the Peabody Picture Vocabulary Test (Dunn and Dunn, 1997). The good readers (designated as average to above average by their teachers) were recruited from the same schools.

### Remedial Instruction

The main goal of the neuroimaging was to determine whether there was a relation between reading improvement and changes in white matter (regardless of the focus of the various remedial instruction programs). The poor readers were randomly assigned to either a control condition that did not include remedial instruction or to one of four remedial reading programs: Corrective Reading ($n = 9$), Wilson Learning System ($n = 9$), Spell Read Phonological Auditory Training ($n = 10$), and Failure Free Reading ($n = 7$). All of these programs provided systematic and explicit instruction in word-level decoding skills. Failure Free Reading focuses on developing recognition of words by sight, whereas the other three programs emphasize phonemic decoding. Additional detail about the specific instructional approaches and how they were implemented can be found elsewhere (Meyler et al., 2008; Torgesen et al., 2006). The instruction was delivered 5 days per week for 50 min a day to groups of one to three students at a time, for a period of 6 months, providing a total of approximately 100 hr of intensive reading instruction.

### Diffusion Tensor Imaging

Diffusion data were acquired on a 3T Siemens Allegra Scanner at the Brain Imaging Research Center of Carnegie Mellon and the University of Pittsburgh. A diffusion-weighted, double spin-echo, echo-planar imaging sequence was used to reduce effects of eddy currents, with TR = 4400 ms, TE = 74 ms, bandwidth = 2298 Hz/Voxel, FOV = 200 mm, and matrix size = 128 × 128. Thirty-six 3-mm-thick slices were imaged (no slice gap) with no diffusion weighting ($b = 0$ s/mm$^2$) and with diffusion-weighting gradients applied in six orthogonal directions ($b = 850$ s/mm$^2$). Twelve images of each slice by gradient direction (and $b = 0$) combination were acquired and averaged to produce the final diffusion imaging dataset for each participant. The FMRIB Diffusion Toolkit (v. 2.0, part of the FMRIB Software Library, http://www.fmrib.ox.ac.uk/fsl/) was used for motion and eddy current correction prior to analysis.

### Data Analyses

Diffusion tensors and scalar diffusion parameter maps were calculated for each participant in native space using standard algorithms (Basser and Pierpaoli, 1996; Basser et al., 1994). For voxel-wise analyses, the diffusion tensor data were reduced to FA maps for each participant. For normalization of the DTI data to a standard space, a custom template was created from the T2-weighted b0 scans of all participants. SPM2 (Wellcome Department of Cognitive Neurology, London, UK) was used to first normalize each participant's b0 volume to the Montreal Neurological Institute (MNI) EPI template using an affine transformation and 12 iterations of the default SPM2 nonlinear normalization algorithm. These normalized T2-weighted images were then averaged across all participants in both reading ability groups to produce a new template customized for the ages and reading abilities of the sample. Each participant's original, native-space b0 volume was then normalized to this new template using the same algorithm, and the transformation parameters for this normalization were applied to the participant's FA map and the maps for axial diffusivity ($\lambda 1$) and radial diffusivity ($\lambda 2 + \lambda 3/2$). For most of the analyses, the normalized maps for the three DTI scalar measures were spatially smoothed with an 8 mm FWHM Gaussian filter to accommodate imprecision of the normalization procedure, to improve signal to noise ratio, and to satisfy assumptions of Gaussian random field theory. Each participant's DTI data were masked on the basis of their individual FA map at a threshold of 0.2 in order to restrict the analyses to white matter.

Analyses of standardized test scores were carried out in SAS (v. 9.1) software using mixed-effects analyses of variance (ANOVAs) (PROC MIXED) and paired or two-sample t tests, with corrections for multiple comparisons made by using a false discovery rate (Benjamini and Hochberg, 1995) of 5%, where appropriate (PROC MULTTEST). Voxel-wise statistical analyses of FA were carried out in SPM2 using the general linear model. Random-effects contrasts of FA data were carried out using Group (Good Reader Controls, Poor Reader Controls, and Remediated Poor Readers) as a between-subject variable and instructional Time (Pre versus Post) as a within-subject variable. Reliable simple effects of Time within groups and Group within each time are reported for clusters of voxels exceeding a voxel-level threshold of $p < 0.005$ (uncorrected) and a cluster size threshold of $p < 0.05$, corrected for multiple corrections in the context of random Gaussian field theory as implemented in SPM2. Additional random effects multiple regression analyses were carried out within the preremediation and postremediation phases of the experiment and for postremediation minus preremediation difference images, with age and raw reading scores entered as continuous independent predictor variables. Voxel-wise nonparametric tests reported in the Supplemental Results and Discussion were carried out using the Randomize (v.1.2) tool included in version 4.1 of the FMRIB Software Library (FSL, http://www.fmrib.ox.ac.uk/fsl/) with 5000 permutations and default neighborhood connectivity parameters for the threshold-free cluster enhancement option for multiple comparison correction. FA and axial and radial diffusivity were also analyzed by extracting the scalar values from each subject for each voxel showing reliable effects on FA, and the values averaged across voxels were submitted to mixed-effects ANOVAs and multiple regression analyses (SAS PROC MIXED) and to stepwise hierarchical regression analyses (SAS PROC REG). Stepwise regressions were conducted using the default options of the REG procedure for variable entry and removal ($p < 0.15$ for both).

To produce an averaged diffusion tensor dataset for each group at each phase of the study, a 12 parameter affine transformation was computed between the b0 scan for each participant and the b0 template created above. This affine transformation was then applied separately to each component of the participant's diffusion tensor dataset, the spatially transformed components were recombined for each subject, and the eigenvectors of the resulting tensor data were reoriented using the preservation of principal directions (PPD) method (Alexander et al., 2001) as implemented in the Camino software package (Cook et al., 2006). The individual components were then averaged across participants within each group at each phase, and the resulting averaged and reoriented components were recombined to produce a group-averaged diffusion tensor data set. Deterministic streamline fiber tracking of group-averaged diffusion tensor data was carried out using a modified version of the FACT algorithm (Mori et al., 1999) as implemented in Camino, using a curvature threshold of 70 degrees and a liberal anisotropy threshold of 0.05 to allow estimated fibers to penetrate gray matter in order to better characterize the possible cortical and subcortical regions connected by the estimated fibers. Tractography was seeded using the cluster showing a group by time interaction for the FA data at the preremediation phase shown in Figure 2A.

## SUPPLEMENTAL DATA

Supplemental Data include four figures, two tables, and Supplemental Results and Discussion, and can be found with this article online at http://www.cell.com/neuron/supplemental/S0896-6273(09)00847-2.

## ACKNOWLEDGMENTS

This research was supported by grants from the R.K. Mellon Foundation, the National Institute of Mental Health (Grant MH029617), and the William and Flora Hewlett Foundation. Participants were recruited through the Power4Kids program, which is a public-private partnership including the Haan Foundation for Children; Institute of Education Sciences, U.S. Department of Education; Heinz Endowments; Smith Richardson Foundation; W.K. Kellogg Foundation; Grable Foundation; Rockefeller Foundation; Ambrose Monell Foundation; Raymond Foundation; and Barksdale Reading Institute. We thank Anne Meyler for help with data collection and collation, Kwan-Jin Jung and Vladimir Cherkassky for technical assistance, Cindy Haan and Joe Torgesen for leadership of the Power4Kids program, and Donna Durno, Rosanne Javorsky, and the Allegheny Intermediate Unit for their central coordinating efforts throughout the project.

Accepted: October 13, 2009
Published: December 9, 2009

## REFERENCES

Alexander, D.C., Pierpaoli, C., Basser, P.J., and Gee, J.C. (2001). Spatial transformations of diffusion tensor magnetic resonance images. IEEE Trans. Med. Imaging *20*, 1131–1139.

Aylward, E.H., Richards, T.L., Berninger, V.W., Nagy, W.E., Field, K.M., Grimme, A.C., Richards, A.L., Thomson, J.B., and Cramer, S.C. (2003). Instructional treatment associated with changes in brain activation in children with dyslexia. Neurology *61*, 212–219.

Basser, P.J., and Pierpaoli, C. (1996). Microstructural and physiological features of tissue elucidated by quantitative-diffusion-tensor MRI. J. Magn. Reson. B. *111*, 209–219.

Basser, P.J., Mattiello, J., and LeBihan, D. (1994). Estimation of the effect self-diffusion tensor from the NMR spin echo. J. Magn. Reson. B. *103*, 247–254.

Beaulieu, C. (2002). The basis of anisotropic water diffusion in the nervous system - a technical review. NMR Biomed. *15*, 435–455.

Beaulieu, C., Plewes, C., Paulson, L.A., Roy, D., Snook, L., Concha, L., and Phillips, L. (2005). Imaging brain connectivity in children with diverse reading ability. Neuroimage *25*, 1266–1271.

Ben-Shachar, M., Dougherty, R.F., and Wandell, B.A. (2007). White matter pathways in reading. Curr. Opin. Neurobiol. *17*, 258–270.

Bengtsson, S.L., Nagy, Z., Skare, S., Forsman, L., Forssberg, H., and Ullen, F. (2005). Extensive piano practicing has regionally specific effects on white matter development. Nat. Neurosci. *8*, 1148–1150.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. J. R. Stat. Soc. Series B Stat. Methodol. *57*, 289–300.

Bitan, T., Burman, D.D., Chou, T., Dong, L., Cone, N.E., Cao, F., Bigio, J.D., and Booth, J.R. (2007). The interaction between orthographic and phonological information in children: An fMRI study. Hum. Brain Mapp. *28*, 880–891.

Cook, P.A., Bai, Y., Nedjati-Gilani, S., Seunarine, K.K., Hall, M.G., Parker, G.J., and Alexander, D.C. (2006). Camino: Open-Source diffusion-MRI econstruction and processing. Proc. Int. Soc. Magn. Reson. Med. *14*, 2759.

Demerens, C., Stankoff, B., Logak, M., Anglade, P., Allinquant, B., Couraud, F., Zalc, B., and Lubetzki, C. (1996). Induction of myelination in the central nervous system by electrical activity. Proc. Natl. Acad. Sci. USA *93*, 9887–9892.

Deutsch, G.K., Dougherty, R.F., Bammer, R., Siok, W.T., Gabrieli, J.D.E., and Wandell, B. (2005). Children's reading performance is correlated with white matter structure measured by diffusion tensor imaging. Cortex *41*, 354–363.

Dougherty, R.F., Ben-Shachar, M., Deutsch, G.K., Hernandez, A., Fox, G., and Wandell, B.A. (2007). Temporal-callosal pathway diffusivity predicts phonological skills in children. Proc. Natl. Acad. Sci. USA *104*, 8556–8561.

Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U., and May, A. (2004). Changes in grey matter induced by training. Nature *427*, 311–312.

Draganski, B., Gaser, C., Kempermann, G., Kuhn, H.G., Winkler, J., Buchel, C., and May, A. (2006). Temporal and spatial dynamics of brain structure changes during extensive learning. J. Neurosci. *26*, 6314–6317.

Dunn, L.M., and Dunn, L.M. (1997). Peabody Picture Vocabulary Test-Revised (Circle Pines, MN: American Guidance Service).

Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C.J., Wedeen, V.J., and Sporns, O. (2008). Mapping the structural core of human cortex. PLoS Biol. *6*, e159.

Hampson, M., Tokoglu, F., Sun, Z., Schafer, R.J., Skudlarski, P., Gore, J.C., and Constable, R.T. (2006). Connectivity-behavior analysis reveals that functional connectivity between left BA39 and Broca's area varies with reading ability. Neuroimage *31*, 513–519.

Hoeft, F., Hernandez, A., McMillon, G., Taylor-Hill, H., Martindale, J.L., Meyler, A., Keller, T.A., Siok, W.T., Deutsch, G.K., Just, M.A., et al. (2006). Neural basis of dyslexia: a comparison between dyslexic and nondyslexic children equated for reading ability. J. Neurosci. *26*, 10700–10708.

Hoeft, F., Meyler, A., Hernandez, A., Juel, C., Taylor-Hill, H., Martindale, J.L., McMillon, G., Kolchugina, G., Black, J.M., Faizi, A., et al. (2007). Functional and morphometric dissociation between dyslexia and reading ability. Proc. Natl. Acad. Sci. USA *104*, 4234–4239.

Horwitz, B., Rumsey, J.M., and Donohue, B.C. (1998). Functional connectivity of the angular gyrus in normal reading and dyslexia. Proc. Natl. Acad. Sci. USA *95*, 8939–8944.

Ishibashi, T., Dakin, K.A., Stevens, B., Lee, P.R., Kozlov, S.V., Stewart, C.L., and Fields, R.D. (2006). Astrocytes promote myelination in response to electrical impulses. Neuron *49*, 823–832.

Jones, D.K., Symms, M.R., Cercignani, M., and Howard, R.J. (2005). The effect of filter size on VBM analyses of DT-MRI data. Neuroimage *26*, 546–554.

Just, M.A., Cherkassky, V.L., Keller, T.A., Kana, R.K., and Minshew, N.J. (2007). Functional and anatomical cortical underconnectivity in autism: Evidence from an fMRI study of an executive function task and corpus callosum morphometry. Cereb. Cortex *17*, 951–961.

Klingberg, T., Hedehus, M., Temple, E., Saltz, T., Gabrieli, J.D.E., Moseley, M.E., and Poldrack, R.A. (2000). Microstructure of temporo-parietal white matter as a basis for reading ability: Evidence from diffusion tensor magnetic resonance imaging. Neuron *25*, 493–500.

Lee, J.E., Chung, M.K., Lazar, M., DuBray, M.B., Kim, J., Bigler, E.D., Lainhart, J.E., and Alexander, A.L. (2009). A study of diffusion tensor imaging by tissue-specific, smoothing-compensated voxel-based analysis. NeuroImage *44*, 870–883.

Meyler, A., Keller, T.A., Cherkassky, V.L., Lee, D., Hoeft, F., Whitfield-Gabrieli, S., Gabrieli, J.D., and Just, M.A. (2007). Brain activation during sentence comprehension among good and poor readers. Cereb. Cortex *17*, 2780–2787.

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

Mori, S., Crain, B.J., Chacko, V.P., and van Zijl, P.C. (1999). Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. Ann. Neurol. *45*, 265–269.

Niogi, S.N., and McCandliss, B.D. (2006). Left lateralized white matter microstructure accounts for individual differences in reading ability and disability. Neuropsychologia *44*, 2178–2188.

Oakes, T.R., Fox, A.S., Johnstone, T., Chung, M.K., Kalin, N., and Davidson, R.J. (2007). Integrating VBM into the general linear model with voxelwise anatomical covariates. NeuroImage *34*, 500–508.

Odegard, T.N., Farris, E.A., Ring, J., McColl, R., and Black, J. (2009). Brain connectivity in non-reading impaired children and children diagnosed with developmental dyslexia. Neuropsychologia *47*, 1972–1977.

Pugh, K.R., Mencl, W.E., Shaywitz, B.A., Shaywitz, S.E., Fulbright, R.K., Constable, R.T., Skudlarski, P., Marchione, K.E., Jenner, A.R., Fletcher, J.M., et al. (2000). The angular gyrus in developmental dyslexia: Task-specific differences in functional connectivity within posterior cortex. Psychol. Sci. *11*, 51–56.

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

Richards, T., Stevenson, J., Crouch, J., Johnson, L.C., Maravilla, K., Stock, P., Abbott, R., and Berninger, V. (2008). Tract-based spatial statistics of diffusion tensor imaging in adults with dyslexia. AJNR Am. J. Neuroradiol. *29*, 1134–1139.

Rollins, N.K., Vachha, B., Srinivasan, P., Chia, J., Pickering, J., Hughes, C.W., and Gimi, B. (2009). Simple developmental dyslexia in children: alterations in diffusion- tensor metrics of white matter tracts at 3 T. Radiology *251*, 882–891.

Shaywitz, B.A., Shaywitz, S.E., Pugh, K.R., Mencl, W.E., Fulbright, R.K., Skudlarski, P., Constable, R.T., Marchione, K.E., Fletcher, J.M., Lyon, G.R., and Gore, J.C. (2002). Disruption of posterior brain systems for reading in children with developmental dyslexia. Biol. Psychiatry *52*, 101–110.

Shaywitz, B.A., Shaywitz, S.E., Blachman, B.A., Pugh, K.R., Fulbright, R.K., Skudlarski, P., Mencl, W.E., Constable, R.T., Holahan, J.M., Marchione, K.E., et al. (2004). Development of left occipitotemporal systems for skilled reading in children after a phonologically-based intervention. Biol. Psychiatry *55*, 926–933.

Simos, P.G., Breier, J.I., Fletcher, J.M., Bergman, E., and Papanicolaou, A.C. (2000a). Cerebral mechanisms involved in word reading in dyslexic children: A magnetic source imaging approach. Cereb. Cortex *10*, 809–816.

Simos, P.G., Breier, J.I., Fletcher, J.M., Foorman, B.R., Bergman, E., Fishbeck, K., and Papanicolaou, A.C. (2000b). Brain activation profiles in dyslexic children during non-word reading: a magnetic source imaging study. Neurosci. Lett. *290*, 61–65.

Simos, P.G., Fletcher, J.M., Bergman, E., Berier, J.I., Foorman, B.R., Castillo, E.M., Davis, R.N., Fitzgerald, M., and Papanicolaou, A.C. (2002). Dyslexia-specific brain activation profile becomes normal following successful remedial training. Neurology *58*, 1203–1213.

Smith, S.M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T.E., Mackay, C.E., Watkins, K.E., Ciccarelli, O., Cader, M.Z., Matthews, P.M., and Behrens, T.E. (2006). Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. Neuroimage *31*, 1487–1505.

Song, S.-K., Sun, S.-W., Ramsbottom, M.J., Chang, C., Russell, J., and Cross, A.H. (2002). Dysmyelination revealed through MRI as increased radial (but unchanged axial) diffusion of water. Neuroimage *17*, 1429–1436.

Song, S.K., Yoshino, J., Le, T.Q., Lin, S.J., Sun, S.W., Cross, A.H., and Armstrong, R.C. (2005). Demyelination increases radial diffusivity in corpus callosum of mouse brain. Neuroimage *26*, 132–140.

Stevens, B., Porta, S., Haak, L.L., Gallo, V., and Fields, R.D. (2002). Adenosine: A neuron-glial transmitter promoting myelination in the CNS in response to action potentials. Neuron *36*, 855–868.

Sweet, L.H., Paskavitz, J.F., Haley, A.P., Gunstad, J.J., Mulligan, R.C., Nyalakanti, P.K., and Cohen, R.A. (2008). Imaging phonological similarity effects on verbal working memory. Neuropsychologia *46*, 1114–1123.

Temple, E., Deutsch, G.K., Poldrack, R.A., Miller, S.L., Tallal, P., Merzenich, M.M., and Gabrieli, J.D.E. (2003). Neural deficits in children with dyslexia ameliorated by behavioral remediation: Evidence from functional MRI. Proc. Natl. Acad. Sci. USA *100*, 2860–2865.

Torgesen, J.K., Wagner, R.K., and Rashotte, C.A. (1999). Test of Word Reading Efficiency (TOWRE). (Austin, TX: Pro-ed).

Torgesen, J.K., Myers, D., Schirm, A.S.E., Vartivarian, S., Mansfield, W., Stancavage, F., Durno, D., Javorsky, R., and Haan, C. (2006). Closing the reading gap: first year findings from a randomized trial of four reading interventions for striving readers. Volume II: National Assessment of Title I: Interim Report to Congress, Institute of Education Sciences.

Wheeler-Kingshott, C.A.M., and Cercignani, M. (2009). About "axial" and "radial" diffusivities. Magn. Reson. Med. *61*, 1255–1260.

**Neuron, Volume 64**

**Supplemental Data for:**

# Altering cortical connectivity: Remediation-induced changes in the white matter of poor readers

Timothy A. Keller[1] & Marcel Adam Just[1]

[1]*Center for Cognitive Brain Imaging, Department of Psychology, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA.*

**SUPPLEMENTAL FIGURES**

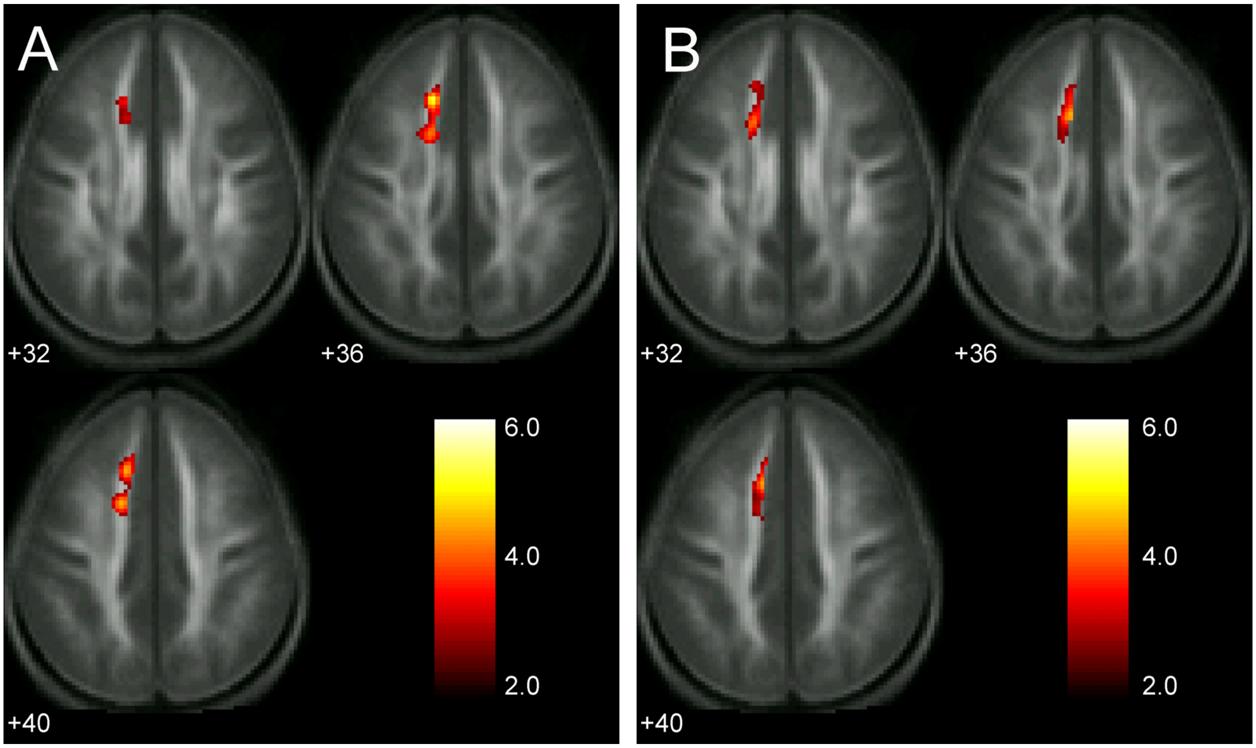


**Figure S1**. Results of parametric voxel-wise analyses conducted on unsmoothed FA data. (A) Regions where the treated poor reader group showed an increase in FA between the pre-remediation and post-remediation scans. For the unsmoothed data the remediation effect is found in two separate clusters in the same region as that reported for the smoothed data in Figure 1A (peak $t(34) = 4.86$, at Montreal Neurological Institute (MNI) coordinates -14 26 36, spatial extent = 18 voxels, $p = .08$, corrected, and peak $t(34) = 4.04$ at MNI coordinates -16 6 38, spatial extent = 23 voxels, $p < .05$, corrected). (B) Region showing a significant difference in FA between good readers and all poor readers at the initial scan (peak $t(70) = 4.15$, at MNI coordinates -12 18 38, spatial extent = 40 voxels, $p < .05$, corrected). Statistical maps are overlaid on normalized FA images averaged across all participants in both scans. The MNI z-coordinate is shown at the bottom left of each axial slice. Color scale represents t-values. Reported p values for t-tests are corrected for multiple comparisons across all of white matter based on cluster extent in the context of Gaussian Random Field theory.

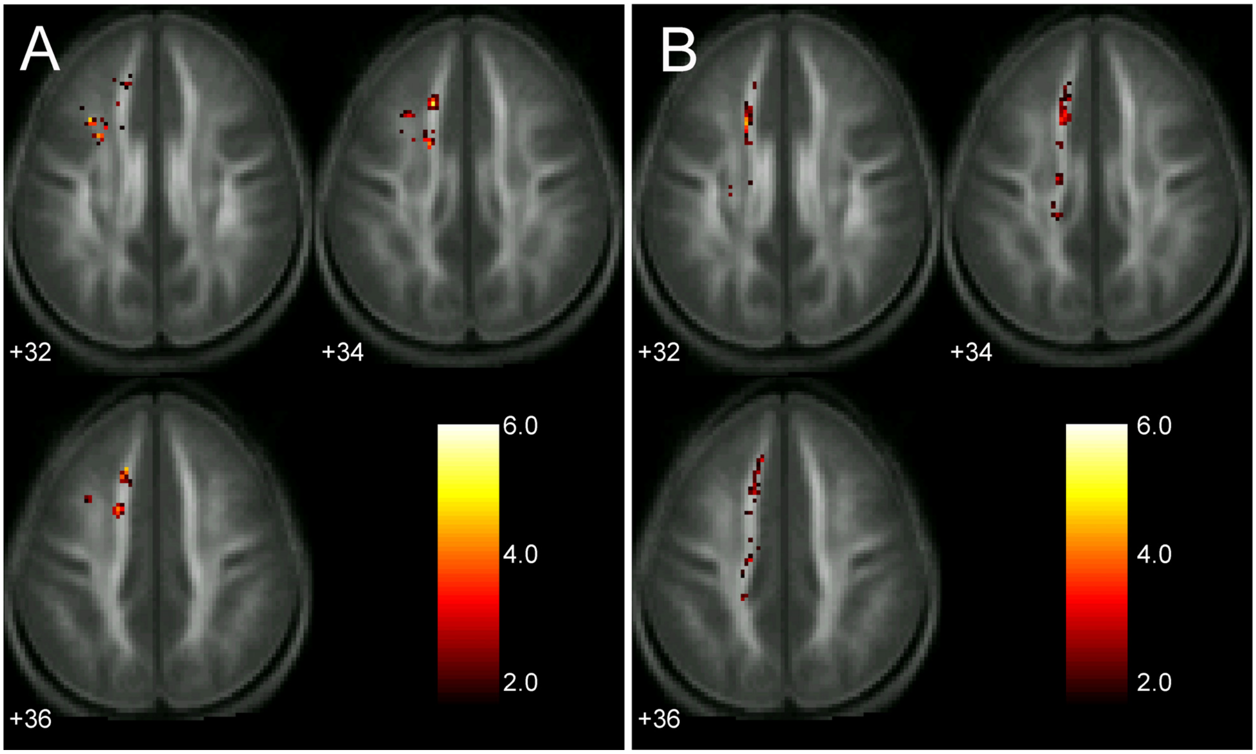


**Figure S2**. Results of non-parametric voxel-wise analyses conducted on unsmoothed FA data. (A) Voxels where the treated poor reader group showed an increase in FA between the pre-remediation and post-remediation scans. The largest increase is found in the same region as that reported for the smoothed data in Figure 1A (peak paired-$t(34) = 4.59$, at Montreal Neurological Institute (MNI) coordinates -12 26 36, $p < .05$, corrected. (B) Voxels showing a significant difference in FA between good readers and all poor readers at the initial scan (peak two-sample $t(70) = 4.15$, at MNI coordinates -12 18 38, $p < .05$, corrected). The MNI z-coordinate is shown at the bottom left of each axial slice. Color scale represents t-values.  Reported p values for t-tests are corrected for multiple comparisons across all of white matter based on cluster extent using the Threshold-Free Cluster Enhancement method (Smith & Nichols, 2009).

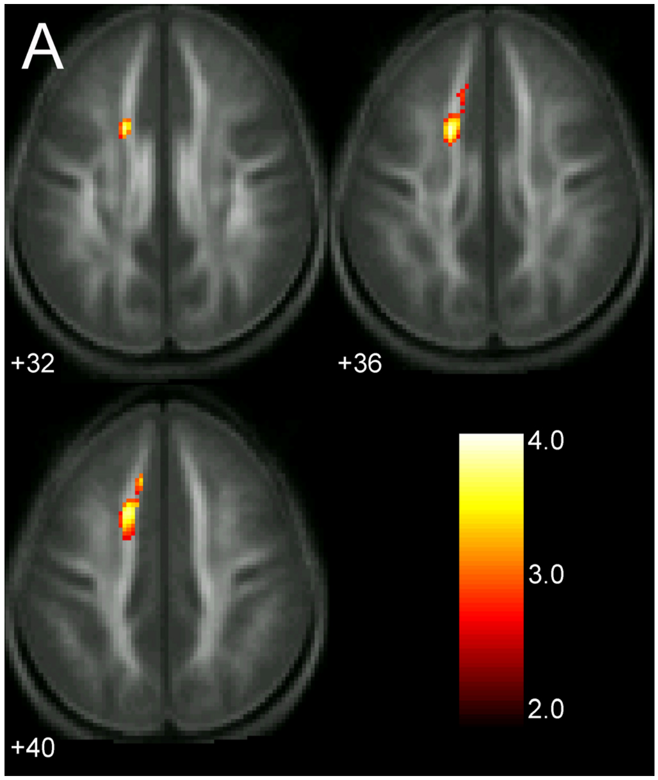


**Figure S3.** Area showing a positive relationship across the entire sample of chidren between raw WRMT Word Attack scores and FA (controlling for age) at the pre-remediation scan (peak t(69) = 3.89 at MNI coordinates -16 12 34, p < .005 corrected for multiple comparisons). The MNI z-coordinate is shown at the bottom left of each axial slice. Color scale represents t-values.

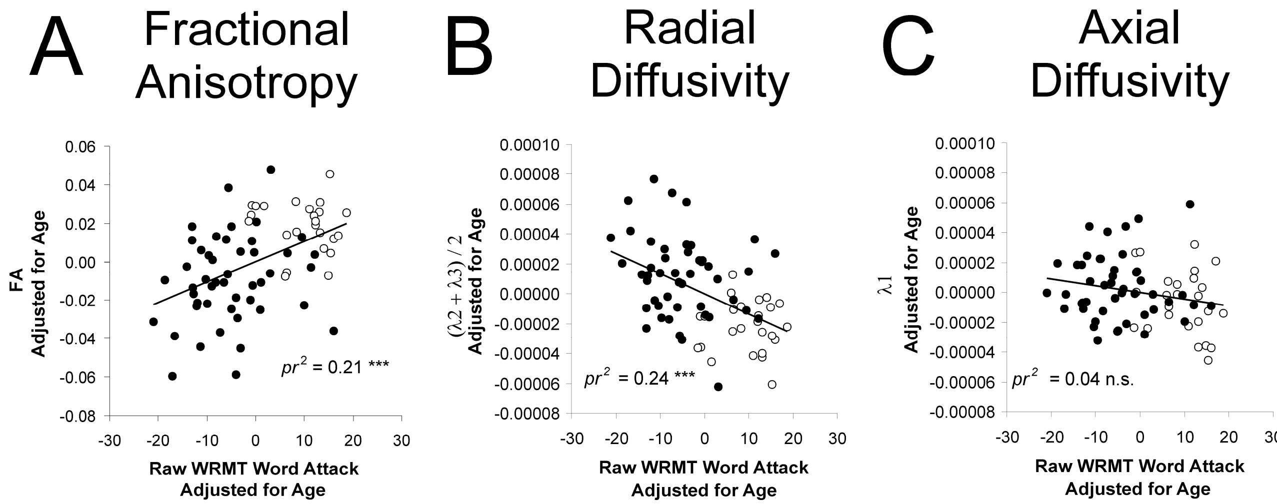


**Figure S4.** Continuous relationships between phonological decoding skill (age-standardized WRMT-R Word Attack) and three DTI measures at the pre-remediation scan across the region showing a reliable group difference in FA. Good readers are shown as open circles and poor readers as filled circles. $pr^2$-values are the squared partial correlation coefficients from a regression model including age as an additional covariate. (A) Fractional anisotropy is reliably positively related to phonological decoding ability. (B) Radial diffusivity is reliably negatively related to phonological decoding ability. (C) Axial diffusivity shows only a modest negative relationship to reading ability. *** $p < .0001$.

## SUPPLEMENTAL TABLES

### Table S1. Reading ability and other scores collected at the time of the pre-remediation scan.

| | Group | | | | | | Group Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| Measure | PR | | PC | | GC | | PR - PC | PR - GC | PC - GC |
| N (n female) | 35 (8) | | 12 (6) | | 25 (8) | | | | |
| | *Mean* | *SD* | *Mean* | *SD* | *Mean* | *SD* | *t(45)* | *t(58)* | *t(35)* |
| Age (years) | 10.0 | (1.1) | 10.2 | (1.2) | 9.8 | (1.0) | 0.71 | 0.81 | 1.41* |
| Grade-standardized PPVT | 96.5 | (13.2) | 97.5 | (12.9) | 110.8 | (12.8) | 0.04 | 4.43** | 3.47** |
| Raw PPVT | 120.9 | (19.0) | 128.2 | (20.2) | 140.5 | (20.2) | 1.00 | 4.17** | 2.20** |
| Age-standardized WRMT-R Word Attack | 92.2 | (11.0) | 95.7 | (7.0) | 114.7 | (11.6) | -1.02 | 7.63*** | 5.23*** |
| Raw WRMT-R Word Attack | 18.0 | (8.6) | 22.5 | (6.8) | 33.9 | (6.2) | -1.66 | 7.94*** | 5.11*** |
| Age-standardized WRMT-R Word ID | 90.5 | (10.4) | 88.8 | (9.3) | 108.4 | (8.2) | 0.53 | 7.13*** | 6.52*** |
| Raw WRMT-R Word ID | 54.2 | (11.3) | 53.0 | (17.2) | 73.4 | (7.8) | 0.28 | 7.28*** | 5.00*** |
| Age-standardized WRMT-R Passage Comprehension | 94.3 | (11.2) | 96.8 | (10.5) | 111.8 | (10.4) | -0.68 | 6.15*** | 4.10** |
| Raw WRMT-R Passage Comprehension | 31.7 | (7.1) | 34.3 | (7.1) | 42.2 | (6.8) | -1.07 | 5.76*** | 3.30** |
| Age-standardized WRMT-R Basic Skills Cluster | 90.6 | (10.2) | 89.7 | (8.6) | 114.3 | (12.3) | 0.29 | 8.14*** | 6.24*** |
| Age-standardized WRMT-R Total Reading Cluster | 91.4 | (9.7) | 90.3 | (9.8) | 111.8 | (9.4) | 0.35 | 8.14*** | 6.45*** |
| Grade-standardized TOWRE Sight Word Efficiency | 86.7 | (9.1) | 87.5 | (5.7) | 110.3 | (11.1) | -0.30 | 9.08*** | 6.70*** |
| Raw TOWRE Sight Word Efficiency | 49.3 | (9.9) | 51.1 | (6.0) | 70.8 | (8.0) | -0.58 | 8.96*** | 7.52*** |
| Grade-standardized TOWRE Phonological Decoding Efficiency | 83.0 | (10.3) | 84.0 | (6.8) | 107.9 | (12.9) | -0.31 | 8.33*** | 6.00*** |
| Raw TOWRE Phonological Decoding Efficiency | 15.8 | (8.7) | 17.0 | (7.6) | 35.8 | (9.4) | -0.42 | 8.47*** | 6.03*** |
| Grade Standardized TOWRE Composite | 81.7 | (10.8) | 82.9 | (6.3) | 111.0 | (13.2) | -0.36 | 9.42*** | 6.94*** |
| Raw TOWRE Composite | 65.1 | (17.1) | 68.1 | (11.9) | 106.6 | (15.5) | -0.55 | 9.60*** | 7.57*** |
| Grade-standardized GRADE | 96.3 | (11.1) | 99.2 | (12.7) | 114.6 | (7.0) | -0.74 | 7.26*** | 4.79*** |
| Raw GRADE | 13.8 | (4.8) | 15.1 | (5.7) | 24.4 | (3.1) | -0.78 | 9.85*** | 6.52*** |
| Grade-standardized AIMS WEB | 91.3 | (7.2) | 91.5 | (6.5) | 114.4 | (11.6) | -0.08 | 9.49*** | 6.34*** |
| Grade-standardized WJIII Spelling | 90.9 | (11.7) | 92.4 | (8.7) | 113.7 | (11.3) | -0.4 | 7.53*** | 5.72*** |
| Raw WJIII Spelling | 27.8 | (5.7) | 28.6 | (3.7) | 37.4 | (5.1) | -0.44 | 6.76*** | 5.36*** |
| Grade-standardized WJIII Calculation | 97.4 | (9.0) | 96.5 | (6.8) | 118.6 | (11.0) | 0.33 | 8.21*** | 6.39*** |
| Raw WJIII Calculation | 15.6 | (3.3) | 15.7 | (3.0) | 21.4 | (4.5) | -0.04 | 5.73*** | 3.99** |

**Notes.** PR - poor readers who would receive remediation, PC - poor reader controls, GC - good reader controls. * $P < .05$, ** $P < .005$, *** $P < .0001$.

**Table S2. Raw reading ability scores collected at the time of the post-remediation scan, and differences from pre-remediation scan.**

| Measure | Group | | | | Change in Scores (Time 2 - Time 1) | | | |
|---|---|---|---|---|---|---|---|---|
| | PR | | PC | | PR | | PC | |
| | *Mean* | *SD* | *Mean* | *SD* | *Diff* | *t(34)* | *Diff* | *t(11)* |
| WRMT-R Word Attack | 25.1 | (8.2) | 24.3 | (3.1) | 7.1 | 6.39*** | 1.8 | 1.16 |
| WRMT-R Word ID | 61.0 | (10.3) | 59.5 | (11.5) | 6.8 | 4.35*** | 6.5 | 1.14 |
| WRMT-R Passage Comprehension | 36.8 | (11.5) | 34.3 | (3.5) | 5.1 | 2.46* | 0.0 | 0 |
| TOWRE Sight Word Efficiency | 56.5 | (10.7) | 59.6 | (8.7) | 7.2 | 5.76*** | 8.5 | 4.13** |
| TOWRE Phonemic Decoding Efficiency | 24.5 | (17.1) | 24.8 | (5.6) | 8.7 | 3.65** | 7.8 | 4.15** |
| TOWRE Composite | 80.9 | (19.9) | 84.4 | (11.1) | 15.8 | 7.71*** | 16.3 | 5.18** |
| AIMSWEB Total Errors | 10.3 | (7.0) | 13.3 | (8.9) | -6.4 | -3.91** | -1.8 | -0.77 |
| AIMSWEB Total Correct | 271.9 | (82.6) | 273.1 | (55.5) | 58.7 | 8.91*** | 54.2 | 5.54** |
| GRADE | 18.1 | (9.1) | 16.0 | (6.8) | 4.3 | 3.18** | 0.9 | 0.96 |
| WJIII Spelling | 30.8 | (6.9) | 31.1 | (2.5) | 3.0 | 3.70** | 2.5 | 3.68** |
| WJIII Calculation | 19.0 | (3.8) | 19.1 | (3.5) | 3.3 | 6.55*** | 3.4 | 5.40** |

**Notes.** PR - poor readers who would receive remediation, PC - poor reader controls. * $p < .05$, ** $p < .005$, *** $p < .0001$. Data were not collected for good readers at this time.

## SUPPLEMENTAL RESULTS AND DISCUSSION

**Equated reading and other abilities among poor readers at the pre-remediation scan.** The randomized assignment of poor readers to treatments succeeded in equating poor readers who would receive remediation (PR) with the poor reader control group (PC) who wouldn't, with respect to age and all test scores acquired prior to the initial scan and to the remediation, as shown by the two-sample t-tests in Table S1. Additional one-way ANOVAs on each measure comparing the five groups of poor readers (four instruction program groups and one group that received no remedial instruction) indicated that there were no reliable differences among any of the groups of poor readers for any of the measures prior to the treatment. The good reader control group (GC) was slightly younger than each of the two groups of poor readers, and scored significantly higher on every measure of vocabulary knowledge, reading ability, and spelling and calculation ability, collected at the initial scan.

**Similarity of remediation effects on reading scores among poor readers who received slightly different treatments.** Although it was not a focus of the neuroimaging study due to the small sample size per treatment, we also assessed whether there were differential effects of the four remedial programs on reading ability with two-factor ANOVAs (4 remedial treatment groups by 2 times of scan (pre-/post-remediation)) conducted on of the all measures in Table S1. Eighteen of the 22 ability measures showed a reliable main effect of time after correction for the number of tests considered, with an increase following the instruction. There were no measures that showed a main effect of instruction group and no measures that showed an interaction between instruction group and time of testing after correction for multiple tests. Thus, there was very substantial evidence that reading ability improved among poor readers who received some form of intensive remediation program, but no indication that the different instruction types were differentially effective in improving any specific reading skill.

**Remediation effects on reading scores.** Comparisons of the subscores of the reading measures of all remediated poor readers with control poor readers before and after the remediation period indicated that the remedial instruction improved the phonological decoding skills of the poor readers who received the instruction. Table S2 presents the raw scores on all reading ability measures collapsed across the four groups of poor readers that received some form of intensive remedial reading instruction, and compares them with the scores of the control group of poor readers who received no remedial instruction. Because the poor reader control group had continued to receive their normal reading curricula in the classroom during the course of the study, some improvement in raw scores among this group would be expected and indeed there were a number of measures on which the poor reader controls did improve. A 2 (treatment vs. control poor readers) by 2 (time) ANOVA conducted for raw scores on the WRMT-R Word Attack subtest (measuring the ability to decode pronounceable non-words) showed a reliable group by time interaction ($F_{1, 45} = 6.18$, $p < .05$), resulting from an improvement in scores between the two tests among poor readers who received remedial instruction, but no change in scores among poor readers who did not receive the remedial instruction (see Table S2). No other measure from any of the tests showed a reliable group by time interaction. Considered together, these results suggest that the primary effect of the remedial instruction was an improvement in phonological decoding as measured by the Word Attack subtest.

**Equated FA among poor readers at the pre-remediation scan.** A whole-brain, voxel-wise ANOVA confirmed that there were no significant differences in FA among the five groups of poor readers at the pre-remediation scan. This analysis revealed no clusters that reliably differed among the poor reader groups after correction for multiple comparisons. This was true even when the analysis was restricted to the cluster that showed a reliable difference in FA between good and poor readers at the pre-remediation scan. In addition, an analysis of the FA data averaged across voxels within this same volume of interest also failed to show any reliable differences among the five groups of poor readers at the pre-remediation scan (Failure Free

Mean = 0.35, SEM = 0.009; Spell Read Mean = 0.36; SEM = 0.007; Corrective Reading Mean = 0.36, SEM = 0.008; Wilson Reading Mean = 0.36, SEM = 0.008; Control Group Mean = 0.37; SEM = 0.007; $F_{1, 42} = 1.30$, $p = .28$).

**Similarity in FA changes among poor readers who received slightly different treatments.** To assess whether there were differences in the effects of the four instructional programs on the change in FA before and after remediation, the FA data were submitted to a 4 (group) by 2 (time) whole-brain, voxel-wise mixed ANOVA. The resulting F-maps for the test of the interaction were thresholded at $p < .005$ with a liberal cluster-size threshold of 50 voxels. No clusters survived this threshold. Additional voxel-wise, pair-wise contrasts conducted between each of the groups for the change in FA across the two scans found no clusters in any contrast that reliably differed between any of the pairs of groups at $p < .005$ for t-value and $p < .05$ using a cluster-size threshold to correct for multiple comparisons. Averaged FA data from each participant in each scan extracted from the volume of interest showing a reliable difference between good and poor readers at the pre-remediation scan were also submitted to a 4 (group) by 2 (time) ANOVA. Although there was a large main effect of time in this analysis ($F_{1, 31} = 12.48$, $p < .005$) resulting from an increase in FA across the groups, there was no main effect of remedial treatment group ($F_{3, 31} = 0.25$, $p = .86$) and no group by time interaction ($F_{3, 31} = 1.37$, $p = .27$). In sum, the analyses of both the behavioral data and the FA data provide no evidence of differences among the various remedial instruction programs. All the analyses presented in the main article and the subsequent analyses presented here therefore collapse across this factor.

**Effect of spatial smoothing on the voxel-wise analyses of FA.** The voxel-based methods used for comparing DTI data across groups have been questioned because of the dependence of the parametric statistical tests on the size of the spatial filtering kernel that is used to satisfy assumptions of Random Gaussian Fields Theory (Jones et al. 2005). To assess whether the choice of filter size used here influenced the ability to detect group FA differences and remediation effects among poor readers in left temporo-parietal white matter, the main voxel-wise analyses of the FA data reported in the main article were repeated without any spatial filtering of the data. Comparison of this outcome (shown in Figure S1) with Figure 1 in the main article indicates that although the sensitivity for detecting effects of group, time, and the interaction was enhanced by spatial filtering, the main conclusions of the study are not altered.

**Non-parametric tests of changes and group differences in FA.** The purpose of smoothing in the method used in our study is not only to enhance the signal to noise ratio in the data, but also to ensure that the assumptions underlying Gaussian random field theory, on which the correction for multiple comparisons are based, are met. It has been shown that as the smoothing filter width is decreased, these assumptions are increasingly violated for FA data (Jones et al. 2005). We therefore also explored alternative non-parametric, permutation-based methods for thresholding and correcting for multiple comparisons (Nichols and Holmes, 2002) in the analyses of the unsmoothed FA data. The two main contrasts of interest, change in FA among the PR group (paired t-test) and the good reader - poor reader group difference in FA at the pre-remediation scan (two-sample t-test) were tested in voxel-wise analyses by submitting the unsmoothed FA data to 5,000 permutations to generate null distributions of the statistics, and the threshold-free cluster enhancement method (TFCE, Smith & Nichols, 2009) was used to correct for multiple comparisons using the default neighborhood connectivity parameters. The resulting t-maps, presented in Figure S2, were thresholded at $p < .05$ after TFCE multiple comparison correction. A comparison between Figure S2 with Figure 1 in the main article indicates that our conclusions regarding both the longitudinal change in FA in the left anterior centrum semiovale among the poor readers that received intensive reading remediation, and the group difference in FA in this same region prior to the instruction phase, were both supported using these alternative, nonparametric methods.

**Analyses restricted to an *a priori* left temporo-parietal region of interest.** The voxel-wise analyses presented in the main article found no reliable differences in FA as a function of reading ability at the pre-remediation scan in left temporo-parietal white matter, in contrast to a number of previous studies (Beaulieu et al. 2005; Deutsch et al., 2005; Klingberg et al. 2000; Niogi & McCandliss, 2006; Odegard et al. 2009; Richards et al., 2008; Rollins et al. 2009). We also conducted voxel-wise analyses on both the smoothed and unsmoothed data restricted to a region of interest defined by a sphere of 1-cm radius centered at the mean of the MNI coordinates (-28 -22 26) reported as showing an FA relationship to reading ability in three previous studies (Beaulieu et al. 2005; Deutsch et al., 2005; Klingberg et al. 2000). No areas of reliably reduced FA among poor readers were found within this a priori ROI following a small volume correction for multiple comparisons, and no areas were reliably correlated with any reading ability measures in the area in voxel-wise analyses. However, an ROI-based analysis of the same region did suggest some consistency with earlier reports of reduced FA in this region. A one-tailed, two-sample t-test conducted for mean smoothed FA across the entire ROI indicated that poor readers had marginally lower FA in the ROI than good readers at the initial scan (Poor Reader Mean = 0.436, SEM = 0.003; Good Reader Mean = 0.442, SEM = 0.003; $t(70) = 1.89$, $p = .06$). A similar test on the unsmoothed data indicated reliably lower mean FA across the entire ROI for the poor readers at the initial scan (Poor Reader Mean = 0.449, SEM = 0.003; Good Reader Mean = 0.458, SEM = 0.003; $t(70) = 2.27$, $p < .05$), consistent with the previous findings. Additional analyses of mean radial and axial diffusivity in this a priori region of interest suggested non-significant trends toward higher radial diffusivity among poor readers at the initial scan in both the smoothed data (Poor Reader Mean = $6.24 \times 10^{-4}$ $mm^2/s$, SEM = $4.01 \times 10^{-6}$; Good Reader Mean = $6.16 \times 10^{-4}$ $mm^2/s$, SEM = $4.90 \times 10^{-6}$; $t(70) = 1.56$, $p = .12$) and the unsmoothed data (Poor Reader Mean = $5.82 \times 10^{-4}$ $mm^2/s$, SEM = $3.40 \times 10^{-6}$; Good Reader Mean = $5.74 \times 10^{-4}$ $mm^2/s$, SEM = $4.04 \times 10^{-6}$; $t(70) = 1.82$, $p = .07$), but there were no differences in axial diffusivity in either analysis (both p's > .25). Despite the indication of cross-sectional group differences in this area, additional analyses of longitudinal changes in these three DTI measures across the same a priori region of interest revealed no significant changes between scans among any of the groups (all p's > .15 for paired t-tests) and no reliable group x time interactions (all p's > .15 for interaction F-tests for FA, radial diffusivity, and axial diffusivity).

Thus, although there is some limited evidence that our sample of poor readers had reduced FA in the same area of left temporo-parietal white matter as reported in previous studies (Beaulieu et al. 2005, Klingberg et al. 2000; Deutsch et al. 2005), and that this lower FA was due to higher radial diffusivity in the area, there was no evidence of remediation-related changes in the microstructure of the white matter in this region.

**Cross-sectional relationships at the pre-remediation scan among DTI measures, components of reading ability, and age.** Good and poor readers were defined on the basis of total scores on the TOWRE prior to the remedial treatment, and thus the group difference in reading ability related to the reduction in FA involves measures of reading of both real words (which can be accessed via a direct orthographic route) and non-words (which can be produced only through a phonological route). Although these measures are highly correlated with each other and not independent (real words can also be read via a phonological decoding), it is nevertheless interesting to ask whether these different measures are differentially associated with FA across the entire range of reading ability. To investigate these relationships, the analyses used the scores from two subtests from the WRMT-R (Word Attack, measuring non-word reading, and Word ID, measuring real-word reading) as covariates, chosen because they were slightly less correlated with each other in our sample than other candidate measures and because they provide consistency with previous studies of the relationship between reading ability and FA (Beaulieu et al., 2005; Klingberg et al., 2000; Niogi and McCandliss, 2006). Pre-remediation FA data from all children were submitted to a voxel-wise multiple regression analysis with age, raw Word Attack score, and raw Word ID score, as independent variables.

The whole-brain, voxel-wise multiple regression analysis revealed no areas large enough to survive correction for multiple comparisons that were independently related to either Word Attack or Word ID scores (controlling for other variables in the model). Because the two reading ability measures were highly correlated across the entire sample of children ($r = .68$, $p < .0001$), we also tested separate models including only one or the other of the reading ability measures.

When only Word Attack scores and age were entered as independent variables, a large cluster showing a positive relationship between raw Word Attack scores and FA (controlling for age), was found in the same region of the anterior centrum semiovale that had shown a group difference in FA between good and poor readers (208 voxels with peak $t(69) = 3.89$ at MNI coordinates -16 12 34). This cluster is shown in Figure S3, and it survives correction for multiple comparisons when the region of interest is restricted to the area showing a group difference in FA at the pre-remediation scan ($p < .005$ corrected at the cluster level and $p < .009$ corrected at the voxel level). In contrast, when only Word ID scores and age were entered as independent variables, no areas of white matter were found to be reliably related to Word ID after controlling for age, even when the region of interest was restricted to the region showing a group difference in FA at the pre-remediation scan. These exploratory analyses suggest that the difference in FA between good and poor readers in the left anterior centrum semiovale at the first scan was due primarily to differences in phonological decoding ability rather than sight-word reading ability.

Consistent with previous studies of the development of white matter in this age range, there was a positive effect of age (controlling for both WRMT Word Attack and Word ID scores) on FA values in both hemispheres in this cross-sectional analysis. A large cluster of voxels with a peak near the right putamen (peak $t(68) = 5.39$, MNI coordinates 16 6 -8, spatial extent = 2894 voxels) extended inferiorly into the right anterior corona radiata and uncinate fasciculus and through the cerebral peduncle into the left hemisphere, eventually reaching the left putamen. Notably, this large cluster included areas consistent with the location of the right and left superior longitudinal fasciculi, indicating that FA in these important tracts was continuing to increase with development in this age range. There were no regions showing a reliable negative relationship between age and FA values after controlling for raw WRMT scores.

To further explore the relationships between phonological decoding and sight-word reading abilities with DTI measures, region of interest-based multiple regression analyses were carried out on the average FA, radial diffusivity, and axial diffusivity across all voxels in the cluster which showed a main effect of group at the pre-remediation scan. The full model regressing FA on age, raw WRMT-R Word ID scores, and raw WRMT-R Word Attack scores was significant ($F_{3, 68} = 6.73$, adjusted multiple $R^2 = .19$, $p < .005$). There was no significant effect of age or raw Word ID scores, but there was a significant effect of raw Word Attack scores for this region (partial $r^2 = .10$, $t(68) = 2.75$, $p < .01$). When Word ID scores were dropped from the model, Word Attack scores accounted for 21% of the variance in FA, as shown in Figure S4A. When Word Attack scores were dropped from the model, Word ID scores explained 13% of the variance in FA. These results again indicate that the relationship between reading ability and FA in this region is primarily due to the phonological decoding aspects of reading ability.

A similar analysis assessed the relation between the two components of reading ability and radial diffusivity in the volume of interest showing a group difference at the pre-remediation scan. This analysis also indicated that the full model was significant ($F_{3, 68} = 7.44$, adjusted multiple $R^2 = .21$, $P < .002$). As in the analysis of FA, only the partial regression coefficient for raw Word Attack scores was significant for the analysis of radial diffusivity in the region (partial $r^2 = .10$, $t(68) = -2.82$, $p < .01$). Dropping Word ID scores from the model resulted in raw Word Attack scores accounting for 24% of the variance in radial diffusivity after adjusting for age, as shown in Figure S4B. Dropping Word Attack scores from the model resulted in raw Word ID scores accounting for 15% of the variance in radial diffusivity after adjusting for age. With axial diffusivity as the dependent variable, however, the full regression model was not significant ($F_{3, 68} = 1.76$, adjusted multiple $R^2 = .03$, $p = .16$), and none of the partial regression coefficients were reliable. For a model including only age and WRMT Word Attack, the latter variable accounted for only 4% of the variance in axial diffusivity (Figure S4C). For a model including only age and WRMT Word ID, the latter variable accounted for only 7% of the variance.

These cross-sectional analyses of diffusion measures acquired prior to remediation suggest that the region of the left anterior centrum semiovale showing a reduction in FA among poor readers may involve pathways of particular importance in phonological decoding and that the positive relationship between phonological decoding ability and FA is primarily due to a negative relationship between this ability and radial diffusivity.

**SUPPLEMENTAL REFERENCES**

Howe, K.B., & Shinn, M.M. (2002) Standard Reading Assessment Passages for Use in General Outcome Assessment: A Manual Describing Development and Technical Features. (Eden Prairie: MN: Edformation, Inc.)

S.M. Smith, M. Jenkinson, M.W. Woolrich, C.F. Beckmann, T.E.J. Behrens, H. Johansen-Berg, P.R. Bannister, M. De Luca, I. Drobnjak, D.E. Flitney, R. Niazy, J. Saunders, J. Vickers, Y. Zhang, N. De Stefano, J.M. Brady, and P.M. Matthews. (2004) Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage *23(S1),* 208-219.

Smith, S.M. and Nichols, T.E. (2009). Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. Neuroimage *44*, 83-98.

Williams, K.T. (2001). Group Reading and Diagnostic Evaluation. (Circle Pines, MN: American Guidance Service).

Woodcock, R.W. (1998). Woodcock Reading Mastery Tests-Revised NU (WRMT-R/NU). (Circle Pines, MN: American Guidance Service).

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