

# Neural systems predicting long-term outcome in dyslexia

Fumiko Hoeft<sup>a,b,1</sup>, Bruce D. McCandliss<sup>c</sup>, Jessica M. Black<sup>a,d</sup>, Alexander Gantman<sup>a</sup>, Nahal Zakerani<sup>a</sup>, Charles Hulme<sup>e</sup>, Heikki Lyytinen<sup>f</sup>, Susan Whitfield-Gabrieli<sup>g</sup>, Gary H. Glover<sup>h</sup>, Allan L. Reiss<sup>a,b,h</sup>, and John D. E. Gabrieli<sup>h</sup>

<sup>a</sup>Center for Interdisciplinary Brain Sciences Research, and <sup>b</sup>Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA 94129; <sup>c</sup>Department of Psychology and Human Development, Vanderbilt University, Nashville, TN 37203; <sup>d</sup>Graduate School of Social Work, Boston College, Chestnut Hill, MA 02467; <sup>e</sup>Department of Psychology, University of York, York YO10 5DD, United Kingdom; <sup>f</sup>Department of Psychology, University of Jyväskylä, 40351 Jyväskylä, Finland; <sup>g</sup>Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139; and <sup>h</sup>Department of Radiology, Stanford University School of Medicine, Stanford, CA 94305

Edited by Marcus E. Raichle, Washington University of St. Louis, St. Louis, MO, and approved November 2, 2010 (received for review June 24, 2010)

**Individuals with developmental dyslexia vary in their ability to improve reading skills, but the brain basis for improvement remains largely unknown. We performed a prospective, longitudinal study over 2.5 y in children with dyslexia ( $n = 25$ ) or without dyslexia ( $n = 20$ ) to discover whether initial behavioral or brain measures, including functional MRI (fMRI) and diffusion tensor imaging (DTI), can predict future long-term reading gains in dyslexia. No behavioral measure, including widely used and standardized reading and language tests, reliably predicted future reading gains in dyslexia. Greater right prefrontal activation during a reading task that demanded phonological awareness and right superior longitudinal fasciculus (including arcuate fasciculus) white-matter organization significantly predicted future reading gains in dyslexia. Multivariate pattern analysis (MVPA) of these two brain measures, using linear support vector machine (SVM) and cross-validation, predicted significantly above chance (72% accuracy) which particular child would or would not improve reading skills (behavioral measures were at chance). MVPA of whole-brain activation pattern during phonological processing predicted which children with dyslexia would improve reading skills 2.5 y later with >90% accuracy. These findings identify right prefrontal brain mechanisms that may be critical for reading improvement in dyslexia and that may differ from typical reading development. Brain measures that predict future behavioral outcomes (neuroprognosis) may be more accurate, in some cases, than available behavioral measures.**

inferior frontal gyrus | prediction | compensation | fractional anisotropy | rhyming

**D**evelopmental dyslexia, which occurs in 5–17% of children, is a persistent difficulty in learning to read that is not explained by sensory deficits, cognitive deficits, lack of motivation, or lack of adequate reading instruction (1). Approximately one-fifth of individuals with developmental dyslexia manage to compensate for their underlying learning difficulties and develop adequate reading skills by the time they reach adulthood (2), but the mechanisms by which this compensation occurs remain largely unknown. Improved reading observed in developmental dyslexia is rarely complete, but instead refers to a level of reading superior to clinical cutoff scores that closes the gap between poor reader and typical readers, and that allows children to read adequately for purposes of learning. Many factors likely influence whether dyslexic children make substantial progress in reading, including access to educational resources and interventions, and neuropsychological and behavioral characteristics (reviewed in refs. 3 and 4), such as whether children have multiple deficits (e.g., in both rapid naming and phonological processing; ref. 5). A number of studies have examined neuropsychological, behavioral, and demographic predictors of developing dyslexia (e.g., refs. 6–8) and short-term response to intervention (RTI) (3, 4), but there is little evidence about long-term compensation toward adulthood. Here, we asked whether neuroimaging indices of brain function, measured by functional magnetic resonance imaging (fMRI), or brain structural connectivity, measured by

diffusion tensor imaging (DTI), predict long-term reading improvement in children with dyslexia.

Functional neuroimaging studies of dyslexia have focused mainly on identifying the neural correlates of dyslexia but have also identified neural systems that could mediate successful remediation. Multiple imaging studies have reported hypoactivation in children and adults with dyslexia during reading-related tasks, especially those that demand phonological analysis of print, in left parietotemporal and occipitotemporal regions (9–17). Many studies have also reported hyperactivation in dyslexia, most often left and right inferior frontal gyri (IFG) (10, 12, 18–24). Such hyperactivation may reflect compensatory processes engaged by individuals with dyslexia attempting to overcome dysfunctions in left posterior cortical areas subserving phonological and orthographic processes. Supporting this possibility is the finding that ventral IFG activation increased with age in children with dyslexia, who may be developing compensatory abilities, but not in typical readers (19). These findings raise the possibility that children with dyslexia make progress in reading by relying on an atypical involvement of IFG regions for reading. This possibility can be evaluated prospectively by testing the hypothesis that greater involvement of the IFG in reading predicts future long-term gains in reading for children with dyslexia.

We conducted a prospective, longitudinal study to examine whether functional (fMRI) or structural (DTI) brain measures could predict reading improvement in dyslexia over a 2.5-y period and how such brain-based measures compared with conventional behavioral measures. At the start of the study, we assessed reading on standardized tests of reading and language and performed neuroimaging in 25 children with dyslexia and 20 children without dyslexia (Table S1). During fMRI, participants performed a printed-word rhyme judgment task designed to invoke the phonological analysis of orthographic input that is thought to be a core deficit in dyslexia (1). Then, we reassessed reading on standardized tests 2.5 y later and asked what brain or standardized reading measures taken at baseline predicted how much a child's reading skills would improve by using change in a standard measure for single-word reading accuracy, using as the outcome measure the change in Woodcock Reading Mastery Test Revised (WRMT) Word Identification Subtest standard score per year. In addition, we examined typically reading children to ask whether the results were specific to dyslexia or more generally related to gains in reading ability over time.

Author contributions: F.H., B.D.M., and J.D.E.G. designed research; F.H., J.M.B., A.G., N.Z., S.W.-G., and J.D.E.G. performed research; F.H. contributed new reagents/analytic tools; F.H., B.D.M., J.M.B., A.G., N.Z., S.W.-G., G.H.G., A.L.R., and J.D.E.G. analyzed data; and F.H., C.H., H.L., S.W.-G., G.H.G., A.L.R., and J.D.E.G. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

<sup>1</sup>To whom correspondence should be addressed. E-mail: fumiko@stanford.edu.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1008950108/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1008950108/-DCSupplemental).

We performed univariate and multivariate pattern analyses (MVPA). For univariate fMRI analyses, we examined potential compensatory neural systems in two regions of interest (ROIs), the left and right inferior frontal regions, because previous imaging studies reported age-related increase in brain activation in dyslexia or increase with successful intervention (19, 25). For DTI, left and right superior longitudinal fasciculus [SLF, including arcuate fasciculus (AF)] were chosen on an a priori basis because of its known left-hemisphere role in language processing and reading, its connection to the ventrolateral prefrontal and parietotemporal language systems (26, 27), and its disruption in dyslexia (28, 29) (note, however, ref. 30).

MVPA of fMRI data involves whole-brain pattern classification with, for example, a support vector machine (SVM) aimed at decoding information in the pattern of activation across all voxels that may distinguish between two classes, namely those children with dyslexia who exhibited substantial gains in reading vs. those children who did not over the next 2.5 y. MVPA often appears more sensitive to group or condition contrasts than traditional univariate analyses (31, 32). Because such analyses can over-fit particular datasets, we used a leave-one-out method in which each individual's activation values are analyzed by the pattern classification defined by the other individuals, so that the pattern classification generalizes to each new individual (33). We also adopted a recursive feature elimination (RFE) procedure (34), where features (e.g., voxels) that contributed little would be recursively eliminated until the optimal pattern that gives maximal performance is obtained. We then asked whether MVPA of fMRI was better than the MVPA of behavioral measures or MVPA of results from univariate imaging analyses in predicting long-term outcome.

## Results

**Demographics and Behavioral Results.** Demographic information and behavioral results are reported for all participants (Table S1 and Fig. 1) and for age-matched groups (Table S2, also SI Results). None of the 17 standardized behavioral scores from Time 1 (or composite scores) significantly predicted improvement of reading scores (slope-Word Identification, slope-WID[ss]) ([ss], standard score) in the dyslexic group even at a lenient threshold of  $P = 0.05$  uncorrected for multiple comparisons with univariate analyses. The results were similar if scores on a test of passage comprehension (Woodcock Reading Mastery Test Passage Comprehension Subtest, WRMT PC[ss]) or the composite reading measure [slope-Reading, principal component of all reading measures which included WRMT, Test of Word Reading Efficiency (TOWRE) and Gray Oral Reading Test (GORT) scores] were used as the outcome measure instead of single-word reading accuracy. These measures correlated with each other; change over time on the WRMT WID [ss] (i.e., Slope-WID[ss]) correlated positively with change over time on the WRMT PC[ss] ( $r = 0.76, P < 0.001$ ) and on the composite reading score ( $r = 0.68, P < 0.001$ ). Of the three outcome measures (slope-WID[ss], slope-PC[ss], slope-Reading) and independent measures (17 Time 1 behavioral measures and, thus, 51 possible correlations), the only measures that showed a predictive correlation was the GORT Comprehension [ss] with slope-PC[ss] ( $r = 0.43, P = 0.03$  uncorrected) (Table S3).

**Univariate fMRI Results.** Most subjects (i.e., those who met the current inclusion criteria for typical and poor readers, see Table S1 for the numbers) were included in a prior cross-sectional whole-brain comparison of typically developing and dyslexic children, and between-group comparisons can be found there (17). Here, we examined a priori regions of interest (ROIs) the left and right IFG areas that are candidates for compensatory activation in dyslexia. In the dyslexic group, single-word reading improvement over 2.5 y (slope-WID[ss]) correlated positively with Time 1 right IFG activation (rhyme vs. rest) [right inferior operculum, Brodmann Area 44, Talairach coordinates  $x = 57, y = 8, z = 18, T$  value = 4.34,  $P = 0.04$  small volume correction (SVC) and Bonferroni corrected for two ROIs, mean

cluster  $r = 0.66$ ; Fig. 2A and B]. The right IFG region was the only cluster that showed positive correlation with gains in single-word reading when the whole brain was examined by using a more lenient threshold of  $P = 0.001$  uncorrected. No region showed a significant negative correlation. Activation in right IFG did not show a significant correlation across individuals in the dyslexic group with Time 1 single-word reading skills ( $r = -0.17, P = 0.43$ ). Activation in right IFG ( $x = 44, y = 31, z = -2, t = 3.25, P = 0.05$  SVC,  $r = 0.51$ ) and left IFG ( $x = -30, y = 33, z = -3, t = 3.15, P = 0.05$  SVC,  $r = 0.47$ ) correlated positively with age in dyslexia. Controls showed no significant correlations between baseline brain activations and future gains in reading skill in the two ROIs or elsewhere even when whole-brain was examined at a lenient threshold of  $P = 0.001$  uncorrected (Fig. 2A and B).

**Univariate DTI Results.** With DTI, single-word reading improvement over 2.5 y (slope-WID[ss]) correlated positively with Time 1 fractional anisotropy (FA) values in the right SLF in the dyslexic group ( $n = 21$ , four individuals were missing DTI data;  $r = 0.52, P = 0.03$  Bonferroni correction for the two left and right SLF ROIs; Fig. 2C). FA values of right SLF correlated positively with fMRI brain activation in right IFG ( $r = 0.51, P = 0.04$ ) of the dyslexic group (Fig. 2D). In the control group, there was no correlation between single-word reading improvement and Time 1 FA values in the right SLF ( $n = 17$ , three individuals were missing DTI data,  $r = -0.07, P = 0.80$ ) (Fig. 2C) or between right SLF FA values and right IFG activation ( $r = -0.004, P = 0.99$ ) (Fig. 2D). In the left SLF, single-word reading improvement over 2.5 y did not correlate positively with Time 1 FA values in the dyslexic group ( $r = 0.31, P = 0.17$ ) or in the control group ( $r = 0.18, P = 0.50$ ). Left SLF FA values correlated positively with Time 1 WRMT WID[ss] in the control group ( $r = 0.57, P = 0.027$ ), but not in the dyslexic group ( $r = 0.30, P = 0.21$ ). Age- (and handedness-) matched individuals showed similar effects (Fig. 2B–D).

**MVPA Results.** For MVPA, children with dyslexia were divided into two groups based on whether they were above or below the median for improvement in single-word reading skills over the next 2.5 y relative to their age (Table S1 and Fig. 1). The median was a 1.65[ss] increase in WRMT WID per year in dyslexia. Children who were above the median exhibited significant improvements in both single-word reading (slope-WID, mean = 3.56[ss/y], SD = 1.42,  $t_{(12)} = 9.03, P < 0.001$ ) and reading comprehension (slope-Passage Comprehension, slope-PC; mean = 3.72[ss/y], SD = 2.44,  $t_{(12)} = 5.50, P < 0.001$ ); children who were below the median did not exhibit significant improvement on either reading measure (slope-WID: mean =  $-0.41$ [ss/y], SD = 1.42,  $t_{(11)} = 0.94, P = 0.36$ ; slope-PC: mean = 0.47[ss/y], SD = 1.42,  $t_{(11)} = 0.80, P = 0.44$ ).

Whole-brain MVPA was performed in the dyslexic group with voxel intensities of whole-brain contrast images (rhyme > rest) as features to examine whether whole-brain activation patterns could discriminate between those who gained more versus less in reading ability (results of comparable univariate analysis are in SI Results). Classification accuracy using MVPA between those who improved more versus less was 92% (Fig. 2A), which was significantly better than chance ( $P < 0.001$ ), and other MVPA results including those of behavioral measures ( $P < 0.001$ ; Fig. S1) of univariate fMRI analysis, i.e., right IFG activation ( $P < 0.001$ ); of univariate DTI analysis, i.e., right SLF white matter integrity ( $P < 0.001$ ); or of a combination of fMRI and DTI univariate results ( $P = 0.045$ ) (SI Results). Distance of each point (i.e., subject) from the hyperplane (hyperplane instead of a line since there are more than two features) that optimally bisects between those who improved more in reading versus those who improved less when using whole brain fMRI activation patterns showed significant positive correlation with gains in reading ability (slope-WID[ss];  $r = 0.73, P = 0.00004$ ; Fig. 3B). Voxels contributing to the classification included regions in right IFG

(which overlapped with the cluster from the univariate fMRI analysis) and left prefrontal cortex with positive weights (positive weights indicate that there were more activation as a pattern in those who showed reading gain) and left parietotemporal region with negative weights (Fig. 3 C and D and Table S4). Adding behavioral measures to fMRI and DTI measures or to whole-brain patterns did not significantly improve classification accuracy ( $p > 0.1$ ) and instead reduced accuracy to 48% and 80%, respectively. Results were similar when RFE was not used to reduce features. Results were also similar when children were divided into groups who improved (positive slope-WID values) versus those who did not improve (zero or negative slope-WID values) were compared.

## Discussion

This study revealed that variation in brain function and structure predicted long-term reading improvement in children with dyslexia. Children with dyslexia who at baseline showed greater activation in the right IFG during a rhyme-judgment task and stronger white matter integrity in the right SLF (including AF) showed greater reading improvement over the next 2.5 y. This brain mechanism appears to be specific to dyslexia rather than reflecting growth in reading ability more generally, because typical readers did not show this pattern. Further, patterns of brain activation across the whole brain predicted future reading gains in children with dyslexia significantly better than the combination of right IFG activation and right SLF white matter integrity, which, in turn, performed better than using either of these brain measures alone. The recruitment of right frontal regions in children with dyslexia may provide a mechanism for enduring improvement that promotes relatively successful reading development. The power of the brain measures to predict future reading gains in children with dyslexia (>90% accuracy) may be contrasted with the finding that none of 17 widely used standardizes measures of reading and language predicted reading gains in these children.

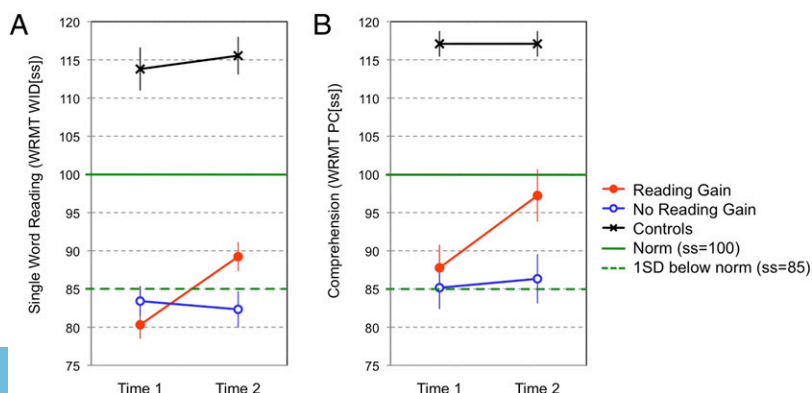
In regards to reading pathways, it appears that dyslexic readers who showed gains in reading did so by depending on a right-hemisphere pathway, in contrast to the left-hemisphere pathway that characterizes typical reading. These results are in agreement with the apparent role of right IFG in compensating for problems in learning to read (19). Further, the right IFG activation correlated positively with age in the dyslexic group, consistent with findings by Shaywitz et al. (19), suggesting that this region plays an increasing role over time in dyslexic children. Indeed, whereas typical reading development appears to involve a shift from right-hemisphere to left-hemisphere involvement (35), gains in reading in dyslexia appear to involve greater dependence on right-hemisphere pathways. Neuroimaging studies examining the neural correlates of effective interventions have reported both a growth of activation in posterior left-hemisphere regions that are typically hypoactive in dyslexia ("normalization"), and also a growth of activation in right-hemisphere regions, including right IFG ("compensation") (24, 36–39).

In one study, right-hemisphere plasticity was not apparent after one year of remediation, but was apparent a year later, as if enduring gains in reading involved the right prefrontal cortex (24). It remains to be learned whether the apparent importance of right IFG in reading gains in dyslexia reflects plasticity in that region, plasticity in other regions, or plasticity in a larger network including right IFG.

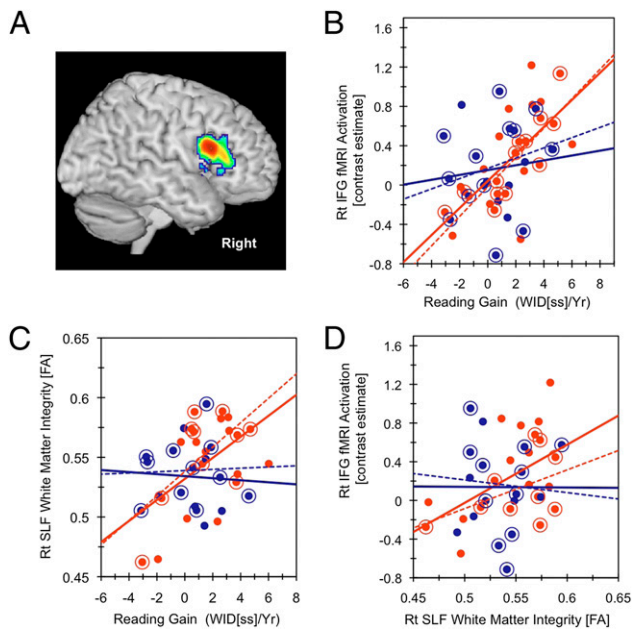
A retrospective study addressed the neural substrates of compensation by comparing adolescents with dyslexia who were compensated readers versus persistent poor readers (i.e., failed to compensate) (40). Activation in the right superior frontal gyrus (SFG) for a phonological task was greater in compensated compared with persistently poor readers (in the present prospective study, the greater activation for readers who would progress to compensation versus readers who would remain persistently poor readers occurred in another right frontal gyrus, the right IFG). Further, in the retrospective study, activation in right IFG was greater in both compensated and persistently poor readers than to typical readers. Thus, both studies point to the atypical importance of right prefrontal cortex in dyslexia in older children and young adults for reading and for differential progress in reading.

The integrity of white matter in the right SLF, as indexed by FA values in DTI, predicted future reading gains in dyslexic children. This white-matter pathway includes fibers that connect ventrolateral prefrontal (Broca's area in the left hemisphere) and parietal and temporal regions (Wernicke's area in the left hemisphere) either through direct or indirect pathways (27) and, thus, may play important roles in language and reading. The high correlation between right prefrontal activation and SLF integrity suggests that these measures of right frontal function and structure, respectively, may be two facets of a common developmental mechanism in right prefrontal cortex in dyslexia. Studies contrasting white-matter connectivity between typical and dyslexic readers have often reported reduced FA in dyslexic readers in or near the left SLF (28, 29). Several studies have also reported a positive correlation between reading skill and left-hemisphere white-matter pathways in typically developing children with a wide-range of reading ability (26, 28, 29, 41, 42), and we also found such a correlation in left SLF in typical readers. Thus, although specific findings vary across studies, the integrity of left-hemisphere white-matter pathways as measured by FA are associated with better reading skill in both typical and dyslexic children, and here we found that right-hemisphere SLF integrity is associated with gains in reading in dyslexic children (but unrelated to reading gains in typical children).

The ability to predict future reading gains in dyslexia with neuroimaging evidence is noteworthy in several regards. First, widely used and standardized measures of reading and language failed to predict future reading gains in dyslexia, although it is possible that there are behavioral measures not included here that may predict outcome (e.g., ref. 43). It is also unknown whether the fMRI contrast in the present study was optimal for making predictions. The contrast between phonological judgments for printed words and rest is a blunt contrast that includes many reading



**Fig. 1.** Change in reading skills. Change in single-word reading skills {WRMT Word Identification standard score (WID[ss]); A} and reading comprehension skills {WRMT Passage Comprehension subtest standard score (PC[ss]); B} over time for the dyslexic group that showed more (Reading Gain group,  $n = 13$ ) or less improvement (No Reading Gain group,  $n = 12$ ) in reading ability (slope-WID [ss]) over 2.5 y based on median split, and the control group.



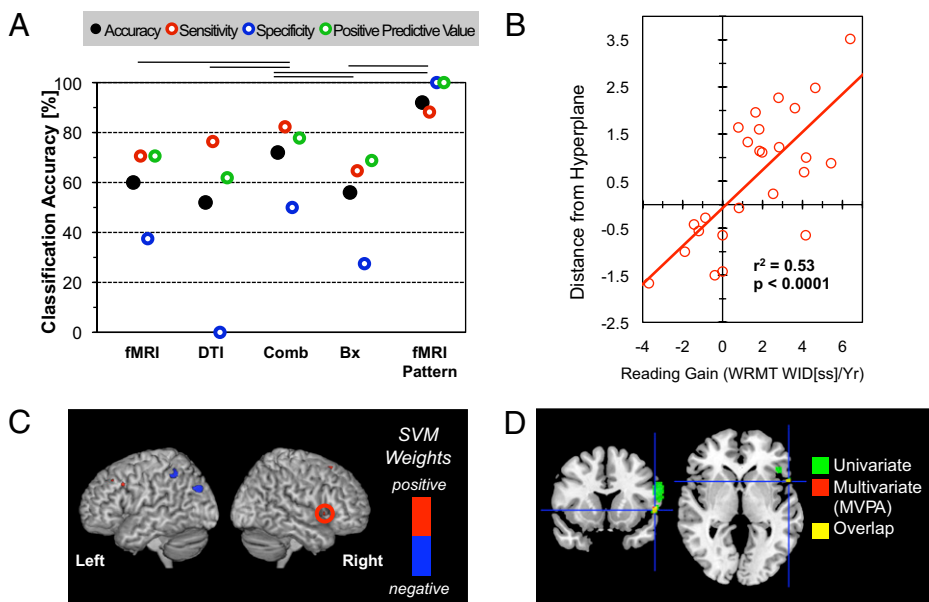
**Fig. 2.** fMRI and DTI predictors of reading gains in dyslexia. (A) Association between brain activation (rhyme > rest) and future reading improvement. This statistical map shows a region in right (Rt) inferior frontal gyrus (IFG) region where significant positive correlation was found with reading gains (increase in WRMT WID[ss] as a function of time) 2.5 y later in the group with dyslexia. No other significant correlations were found either in the group with dyslexia or controls. (B) Association between right IFG activation and future reading improvement. Individual contrast estimates from entire right inferior frontal cluster in Fig. 1A showed significant positive correlation with reading gains in the group with dyslexia ( $n = 25$ ; red circles; mean cluster  $r^2 = 0.46$ ), but not in the control group ( $n = 20$ ; blue circles; mean cluster  $r^2 = 0.016$ ). Subgroups of dyslexic and control participants matched for age are shown with larger hollow ring; regression line for age-matched subgroups are shown in dotted lines (B–D). (C) Association between white matter integrity and future reading improvement. Individual FA values of the Rt superior longitudinal fasciculus (SLF) showed significant positive correlation between gains in reading in the dyslexic group ( $r^2 = 0.27$ ), but not in the control group ( $r^2 = 0.004$ ). Dyslexia group ( $n = 21$ , 4 individuals were missing DTI data), red circles and regression line in red. Control group ( $n = 17$ , 3 individuals were missing DTI data), blue circles and regression line in blue. (D) Association between DTI white matter integrity and fMRI activation. FA values of Rt SLF and Rt IFG contrast estimates show significant positive correlation in the dyslexia group ( $r^2 = 0.26$ ), but not in the control group ( $r^2 = 0.00002$ ). Dyslexic group ( $n = 21$ ), red circles and regression line in red. Control group ( $n = 17$ ), blue circles and regression line in blue.

processes, and it is unknown whether more precise contrasts that isolate specific reading skills (e.g., orthographic or phonological processes) would increase or decrease prediction accuracy. Second, the combination of just two measures, right IFG activation and right SLF FA values, predicted reading gains over 2.5 y with well above-chance accuracy (72%) and significantly better than using either measure alone in children with dyslexia. The brain regions examined were selected on the basis of the prior literature, so these findings may generalize to the dyslexic population. Third, multivariate patterns of brain activation over the whole gray matter predicted reading gains with >90% accuracy. Classification accuracy of reading gains occurred in an independent sample (test-set) rather than the training-set used to construct a classifier, so that the model (classifier) is likely to generalize, at least more reliably than typical regression analyses where cross-validation is not performed. The fact that there was a significant association between the distance from the hyperplane of the classifier that bisects the two dyslexic groups and change in reading scores further supported the validity of the classifier and reduced the likelihood that the results are due to artifacts.

This study provides evidence that identifying neuroimaging patterns can predict future gains in reading in dyslexia better than behavior, but with several caveats. First, we studied children with well-documented and persistent dyslexia around 14 y of age at the beginning of the study. Future studies involving younger children can determine whether such brain measures predict reading progress in younger prereaders or early readers. Second, we followed these children for 2.5 y, and future studies can examine even longer-term outcomes. Third, our comparison between children with and without dyslexia included group differences on age, IQ, and fMRI task performance. We controlled for these factors in subsidiary analyses that produced similar findings (*SI Results*). Importantly, there were no such differences between the children with dyslexia who improved versus those who did not improve, so these factors could not have influenced predictions of reading gains within the dyslexic children. Fourth, although we made efforts to minimize the effect of regression to the mean by choosing different measures for identifying those with dyslexia versus predicting reading outcome, there is always a possibility of regression to the mean. In this context, this interpretation would suggest that the initial 17 behavioral measures had overestimated randomly the severity of dyslexia in the group who appeared to advance over the next 2.5 y. Overall this interpretation seems unlikely, because dyslexia is not a fluctuating disorder, and scores were highly correlated across many tests. If, however, the results were driven by regression to the mean, the finding would still be important because the brain measures then would still be more accurate for identifying which children were initially misdiagnosed than any behavioral or other known measure. Fifth, some children with dyslexia received various types of remediation between our two measurement points (e.g., tutoring, commercial reading programs), but receiving remediation did not affect how much a child improved in their reading abilities (*SI Results*). Children were not randomly assigned to treatment, however, so these findings may be related to biases in regards to which children received interventions. Future studies could control for intervention effects and include large samples where the effect of intervention could be examined separately. Sixth, although the magnitude of right IFG correlated with reading gains over the next 2.5 y, the role of right IFG in improved reading or persistent poor reading at the end of the 2.5 y was not directly measured and is therefore unknown, although there is evidence that right prefrontal cortex continues to play an atypically important role in phonological awareness in poor readers into young adulthood (40). Finally, although we performed cross-validation analysis to maximize the possibility that the models will generalize to new independent samples, further studies will be needed to determine whether these findings extend to a broad range of children with dyslexia by applying this classifier to a new independent sample defined similarly and using the many different selection criteria available.

Brain imaging studies of dyslexia have shown some predictive power, although prior studies did not identify specific brain mechanisms as in the present study. Evoked-response potentials (ERP) measured with newborns from families with a risk for reading difficulty have correlated with language and reading scores years later (44). ERPs in prereading children with familial risk added significantly to behavioral measures in predicting reading outcome in prereading children, but only the ERP measure predicted reading ability 5 y later (45). Gains in phonological decoding across a school year in poor readers were predicted similarly by behavioral and MRI measures, but the combination of behavioral and brain measures predicted gains significantly better than either measure alone (46).

Two educational implications arise from these findings. First, it appears that gains in reading for dyslexic children reflect, at least in part, different neural mechanisms and pathways than those that support gains in reading for typically developing children. This finding encourages consideration of intervention approaches that capitalize on alternative reading strategies in addition to current interventions that build on typical reading instruction. Indeed, those children in this sample who received typical interventions progressed no more than those who had not



**Fig. 3.** Multivariate pattern classification of reading gains in dyslexia. (A) Classification accuracy using fMRI right inferior frontal gyrus (IFG) activation only (fMRI), DTI right superior longitudinal classification (SLF) only (DTI), combination of the two (Comb), patterns of behavior (Bx), and fMRI whole-brain pattern (fMRI Pattern). Black horizontal lines above the graph show comparisons of interest that are significant (e.g., fMRI pattern showed significantly greater accuracy than Comb). (B) Association between distance from hyperplane for the whole-brain fMRI pattern classifier and gains in reading ability. (C) Brain activation patterns that discriminate between children with dyslexia who gained in reading ability vs. those who did not. Voxels in red are positive weights (those that as a pattern showed greater activation in children who improved compared with those who did not) and in blue are negative weights. (D) Overlap between results from fMRI univariate regression analysis (green) and multivariate analysis (red) in the right IFG activation is displayed (yellow).

received such interventions. Second, although fMRI is typically viewed as a research tool that has little practical implication for an individual with dyslexia, the present findings, as well as prior ERP studies (e.g., ref. 45), suggest that brain measures may already better predict long-term outcomes in reading development than available behavioral measures. The MVPA findings in the present study, that predicted which particular children would or would not advance in reading over the next 2.5 y, also suggest that such brain measures have promise for identifying individual children's future reading trajectory.

In general terms, these findings suggest that brain imaging may play a valuable role in neuroprognosis, the use of brain measures to predict future reductions, or exacerbations of symptoms in clinical disorders. Although much has been learned about the brain basis of dyslexia and other neurobehavioral disorders, it has been unclear as to how such brain measurement can enhance the treatment of those disorders. Brain imaging, however, has shown value in predicting recovery from depression 8 mo later (47), relapse in methamphetamine dependence 1 y later (48), and onset of psychosis in at-risk individuals (49). Further, imaging studies have predicted response to treatment for depression (50, 51), anxiety (52), and cognitive-behavioral therapy in schizophrenia (53). In many of these cases, the brain measures were superior in predicting recovery relative to conventional tests or ratings scales, as was the case in this study. Altogether these findings motivate further research into discovering whether neuroprognosis is useful for identifying treatments that are more or less likely to lead to recovery for an individual and for discovering neurobiological variation within diagnostic categories that are important for treatment selection and effectiveness.

## Methods

**Participants.** Participants were 45 healthy native English-speaking children with ( $n = 25$ , 21 with DTI data) and without ( $n = 20$ , 17 with DTI data) dyslexia (Table S1, SI Results, and SI Methods).

**Criteria for Dyslexia and Behavioral Assessment.** Reading ability and intelligence were assessed by using standardized measures at baseline (Time 1) and  $30 \pm 2$  mo later (Time 2) (Table S1 and S2). Criteria for dyslexia were met if a child's behavioral performance fell within the following limits on age-adjusted tests: performance on a composite score of timed decoding and single word reading within the lower 25th percentile, i.e., TOWRE Phonemic Decoding Efficiency (PDE[ss]) and Sight Word Efficiency (SWE[ss])  $\leq 90$ , and estimated performance IQ within one SD of the norm, i.e., Wechsler Abbreviated Scale of Intelligence Matrix Reasoning Subtest (WASI MR[ss])  $\geq 85$ . Different educators and research groups define dyslexia differently, hence our

current findings will need to be expanded to children with dyslexia defined in different ways. Gains in reading were assessed by change in age-adjusted scores of single-word reading from Time 1 to Time 2 as a function of time (slope-WID[ss]), i.e.,  $(\text{Time 2} - \text{Time 1})_{\text{WID[ss]}} / (\text{years elapsed between Time 1 and Time 2})$ .

**Task Design, fMRI, and DTI Image Acquisition and Processing Procedures.** A written word rhyme task was used as the fMRI task at Time 1 (SI Methods).

**fMRI Data Analyses.** Each subject's data were high pass filtered at 97 s and analyzed by using a fixed effects model comparing task vs. rest. The two ROIs were defined as bilateral inferior frontal (pars triangularis, pars opercularis, and precentral gyrus) regions by using the Automated Talairach Atlas Label (AAL) ([www.cyceron.fr/web/aal\\_anatomical\\_automatic\\_labeling.html](http://www.cyceron.fr/web/aal_anatomical_automatic_labeling.html)) in the WFU PickAtlas toolbox (<http://fmri.wfubmc.edu/cms/software#PickAtlas>). Group analysis was performed with a random effects model (54) by using the rhyme vs. rest contrast images for each dyslexic and control group independently. Primary analyses were restricted to the two ROIs of interest. Whole brain analyses performed as control analyses, not restricting to these ROIs, are noted in the results wherever applicable. Simple regression analyses were performed for the dyslexic and control groups independently to examine associations between brain activation at Time 1 and gains in reading over 2.5 y, i.e., WRMT WID[ss] increase per year. To examine whether activation in the two ROIs also correlated with baseline single-word reading skills, correlation analyses with Time 1 WRMT WID[ss] and brain activation were also performed for each group separately. Further, age was correlated with activation in these two ROIs to examine (cross-sectional) developmental effects.

A statistical threshold of  $P = 0.05$  for voxel-height (uncorrected) was used, which then was thresholded at a voxel-wise threshold of  $P = 0.05$  for family-wise error (FWE) after small-volume correction (SVC). Further, this threshold was Bonferroni corrected for the number of ROIs tested (i.e., 2 ROIs, right and left inferior frontal region, so that the  $P$  values were doubled).

**DTI Data Analyses.** We used the Reproducible Objective Quantification Scheme (ROQS), which is a semiautomated ROI approach in native space that segments white matter structures based on user-selected seed pixels (55). Selection and analysis was implemented with software written in Interactive Data Language v6.0 (IDL; Research Systems). Left and right SLF (including AF) were chosen on an a priori basis because of its known role in language processing and reading and its connection to the ventrolateral prefrontal and parietotemporal language systems (26, 27) and disruption in dyslexia (refs. 28 and 29 but note 30). The investigator placing seeds was blind to subjects' diagnoses. Mean FA values were calculated for each subject and each tract. Associations between FA values and reading gains (slope-WID[ss]) were examined for the dyslexic and control groups. Associations between brain activation determined above and FA values were examined by extracted fMRI contrast estimates (linear combination of beta estimates) and by fMRI whole-brain regression analyses by using the FA values as covariates of interest.

**MVPA.** We performed leave-one-out linear SVM analysis (regularization parameter  $C = 1$ ) by using in-house tools based on Matlab (Mathworks) (56). Classification accuracy (accuracy of which children with dyslexia showed improvement/positive change were classified as “+1” and those who did not improve/negative change were classified as “-1” relative to the total group), sensitivity, specificity, and positive predictive value were calculated for each classification. Further, distance measure derived from the classifier (e.g., a large positive distance measure of a child who improved in reading over time will indicate that the brain-based classifier was highly confident that the child improved) was correlated with change in reading ability (slope-WID[ss]) to support the validity of the classifier. In the dyslexic group, this was compared with classification performance based on using the results from combining all behavioral measures (subtests and total scores) and univariate analyses (i.e., SVM with either the mean level of right IFG activation as the only feature, right SLF FA values as the only feature, or the combination of the two). Significance

was determined by using permutation analysis by randomly reassigning class labels 2,000 times ( $P < 0.05$ ). See *SI Methods* for details.

**ACKNOWLEDGMENTS.** We thank Dana Wittenberg, Candy Ho, and Joshua Heitzmann for assistance with follow-up data collection. This study was supported by National Institute of Child Health and Human Development Grant HD054720 (to F.H.), the Stanford University Lucile Packard Children's Hospital Child Health Research Program (F.H.), National Center for Research Resources Grant P41RR009874 (to G.H.G.), the William and Flora Hewlett Foundation (J.D.E.G.), and the Richard King Mellon Foundation (J.D.E.G.). Recruitment and characterization of the participants with dyslexia at baseline was supported by a grant from BrightStar. F.H. was also supported by a National Alliance for Research on Schizophrenia and Depression Young Investigator Award. J.D.E.G. was supported by the Ellison Medical Foundation, Massachusetts Institute of Technology Class of 1976 Funds for Dyslexia Research, and Bard and Julie Richmond through the Martin Richmond Memorial Fund.

- Shaywitz SE (1998) Dyslexia. *N Engl J Med* 338:307–312.
- Lyytinen H, Erskine J, Aro M, Richardson U (2006) Reading and reading disorders. *Handbook of Language Development*, eds Hoff E, Shatz M (Blackwell Publishers, Oxford).
- Al Otaiba S, Fuchs D (2002) Characteristics of children who are unresponsive to early literacy intervention. *Remedial Spec Educ* 23:300–317.
- Nelson JR, Benner GJ, Gonzalez J (2003) Learner characteristics that influence the treatment effectiveness of early literacy interventions: A meta-analytic review. *Learn Disabil Res Pract* 18:255–267.
- Wolf M, Bowers PG (1999) The double-deficit hypothesis for the developmental dyslexias. *J Educ Psychol* 91:415–438.
- Catts HW, Fey ME, Zhang X, Tomblin JB (2001) Estimating the risk of future reading difficulties in kindergarten children: A research-based model and its clinical implementation. *Lang Speech Hear Serv Sch* 32:38–50.
- Scarborough HS (1989) Prediction of reading disability from familial and individual differences. *J Educ Psychol* 81:101–108.
- Snowling M, Bishop DVM, Stothard SE (2000) Is preschool language impairment a risk factor for dyslexia in adolescence? *J Child Psychol Psychiatry* 41:587–600.
- Cao F, Bitan T, Booth JR (2008) Effective brain connectivity in children with reading difficulties during phonological processing. *Brain Lang* 107:91–101.
- Shaywitz SE, et al. (1998) Functional disruption in the organization of the brain for reading in dyslexia. *Proc Natl Acad Sci USA* 95:2636–2641.
- Rumsey JM, et al. (1997) A positron emission tomographic study of impaired word recognition and phonological processing in dyslexic men. *Arch Neurol* 54:562–573.
- Brunswick N, McCrory E, Price CJ, Frith CD, Frith U (1999) Explicit and implicit processing of words and pseudowords by adult developmental dyslexics: A search for Wernicke's Wortschatz? *Brain* 122:1901–1917.
- Kronbichler M, et al. (2006) Evidence for a dysfunction of left posterior reading areas in German dyslexic readers. *Neuropsychologia* 44:1822–1832.
- Schulz E, et al. (2008) Impaired semantic processing during sentence reading in children with dyslexia: Combined fMRI and ERP evidence. *Neuroimage* 41:153–168.
- McCrory EJ, Mechelli A, Frith U, Price CJ (2005) More than words: A common neural basis for reading and naming deficits in developmental dyslexia? *Brain* 128:261–267.
- Paulesu E, et al. (2001) Dyslexia: Cultural diversity and biological unity. *Science* 291: 2165–2167.
- Hoefl F, et al. (2007) Functional and morphometric brain dissociation between dyslexia and reading ability. *Proc Natl Acad Sci USA* 104:4234–4239.
- Rumsey JM, et al. (1992) Failure to activate the left temporoparietal cortex in dyslexia. An oxygen 15 positron emission tomographic study. *Arch Neurol* 49:527–534.
- Shaywitz BA, et al. (2002) Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biol Psychiatry* 52:101–110.
- Salmelin R, Service E, Kiesilä P, Uutela K, Salonen O (1996) Impaired visual word processing in dyslexia revealed with magnetoencephalography. *Ann Neurol* 40:157–162.
- Richards TL, et al. (2002) Reproducibility of proton MR spectroscopic imaging (PEPSI): Comparison of dyslexic and normal-reading children and effects of treatment on brain lactate levels during language tasks. *AJNR Am J Neuroradiol* 23:1678–1685.
- Georgiewa P, et al. (2002) Phonological processing in dyslexic children: A study combining functional imaging and event related potentials. *Neurosci Lett* 318:5–8.
- Milne RD, Syngieniotis A, Jackson G, Corballis MC (2002) Mixed lateralization of phonological assembly in developmental dyslexia. *Neurocase* 8:205–209.
- Shaywitz BA, et al. (2004) Development of left occipitotemporal systems for skilled reading in children after a phonologically-based intervention. *Biol Psychiatry* 55:926–933.
- Shaywitz BA, et al. (2007) Age-related changes in reading systems of dyslexic children. *Ann Neurol* 61:363–370.
- Beaulieu C, et al. (2005) Imaging brain connectivity in children with diverse reading ability. *Neuroimage* 25:1266–1271.
- Catani M, Jones DK, ffytche DH (2005) Perisylvian language networks of the human brain. *Ann Neurol* 57:8–16.
- Deutsch GK, et al. (2005) Children's reading performance is correlated with white matter structure measured by diffusion tensor imaging. *Cortex* 41:354–363.
- Klingberg T, et al. (2000) Microstructure of temporo-parietal white matter as a basis for reading ability: Evidence from diffusion tensor magnetic resonance imaging. *Neuron* 25:493–500.
- Ben-Shachar M, Dougherty RF, Deutsch GK, Wandell BA (2007) Differential sensitivity to words and shapes in ventral occipito-temporal cortex. *Cereb Cortex* 17:1604–1611.
- Haynes JD, Rees G (2006) Decoding mental states from brain activity in humans. *Nat Rev Neurosci* 7:523–534.
- Norman KA, Polyn SM, Detre GJ, Haxby JV (2006) Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. *Trends Cogn Sci* 10:424–430.
- Pereira F, Mitchell T, Botvinick M (2009) Machine learning classifiers and fMRI: A tutorial overview. *Neuroimage* 45 (1, Suppl)S199–S209.
- De Martino F, et al. (2008) Combining multivariate voxel selection and support vector machines for mapping and classification of fMRI spatial patterns. *Neuroimage* 43:44–58.
- Turkeltaub PE, Gareau L, Flowers DL, Zeffiro TA, Eden GF (2003) Development of neural mechanisms for reading. *Nat Neurosci* 6:767–773.
- Aylward EH, et al. (2003) Instructional treatment associated with changes in brain activation in children with dyslexia. *Neurology* 61:212–219.
- Eden GF, et al. (2004) Neural changes following remediation in adult developmental dyslexia. *Neuron* 44:411–422.
- Meyler A, Keller TA, Cherkassky VL, Gabrieli JD, Just MA (2008) Modifying the brain activation of poor readers during sentence comprehension with extended remedial instruction: A longitudinal study of neuroplasticity. *Neuropsychologia* 46:2580–2592.
- Temple E, et al. (2003) Neural deficits in children with dyslexia ameliorated by behavioral remediation: Evidence from functional MRI. *Proc Natl Acad Sci USA* 100: 2860–2865.
- Shaywitz SE, et al. (2003) Neural systems for compensation and persistence: Young adult outcome of childhood reading disability. *Biol Psychiatry* 54:25–33.
- Nagy Z, Westerberg H, Klingberg T (2004) Maturation of white matter is associated with the development of cognitive functions during childhood. *J Cogn Neurosci* 16: 1227–1233.
- Niogi SN, McCandliss BD (2006) Left lateralized white matter microstructure accounts for individual differences in reading ability and disability. *Neuropsychologia* 44:2178–2188.
- Hatcher PJ, Hulme C (1999) Phonemes, rhymes, and intelligence as predictors of children's responsiveness to remedial reading instruction: Evidence from a longitudinal intervention study. *J Exp Child Psychol* 72:130–153.
- Molfese VJ, Molfese DL, Modgline AA (2001) Newborn and preschool predictors of second-grade reading scores: An evaluation of categorical and continuous scores. *J Learn Disabil* 34:545–554.
- Maurer U, et al. (2009) Neurophysiology in preschool improves behavioral prediction of reading ability throughout primary school. *Biol Psychiatry* 66:341–348.
- Hoefl F, et al. (2007) Prediction of children's reading skills using behavioral, functional, and structural neuroimaging measures. *Behav Neurosci* 121:602–613.
- Canli T, et al. (2005) Amygdala reactivity to emotional faces predicts improvement in major depression. *Neuroreport* 16:1267–1270.
- Paulus MP, Tapert SF, Schuckit MA (2005) Neural activation patterns of methamphetamine-dependent subjects during decision making predict relapse. *Arch Gen Psychiatry* 62:761–768.
- Koutsouleris N, et al. (2009) Use of neuroanatomical pattern classification to identify subjects in at-risk mental states of psychosis and predict disease transition. *Arch Gen Psychiatry* 66:700–712.
- Davidson RJ, Irwin W, Anderle MJ, Kalin NH (2003) The neural substrates of affective processing in depressed patients treated with venlafaxine. *Am J Psychiatry* 160:64–75.
- Fu CH, et al. (2004) Attenuation of the neural response to sad faces in major depression by antidepressant treatment: A prospective, event-related functional magnetic resonance imaging study. *Arch Gen Psychiatry* 61:877–889.
- Whalen PJ, et al. (2008) A functional magnetic resonance imaging predictor of treatment response to venlafaxine in generalized anxiety disorder. *Biol Psychiatry* 63:858–863.
- Kumar V, et al. (2009) Dorsolateral prefrontal cortex activity predicts responsiveness to cognitive-behavioral therapy in schizophrenia. *Biol Psychiatry* 66:594–602.
- Friston KJ, Holmes AP, Price CJ, Büchel C, Worsley KJ (1999) Multisubject fMRI studies and conjunction analyses. *Neuroimage* 10:385–396.
- Niogi SN, Mukherjee P, McCandliss BD (2007) Diffusion tensor imaging segmentation of white matter structures using a Reproducible Objective Quantification Scheme (ROQS). *Neuroimage* 35:166–174.
- Hoefl F, et al. (2008) Morphometric spatial patterns differentiating boys with fragile X syndrome, typically developing boys, and developmentally delayed boys aged 1 to 3 years. *Arch Gen Psychiatry* 65:1087–1097.