Research Article

Psychological Science
22(11) 1442–1451

DOI: 10.1177/0956797611419521
http://pss.sagepub.com

# The Brain Basis of the Phonological Deficit in Dyslexia Is Independent of IQ

**Hiroko Tanaka[1,2], Jessica M. Black[1,3,4], Charles Hulme[5], Leanne M. Stanley[1,2], Shelli R. Kesler[1,3], Susan Whitfield-Gabrieli[6,7], Allan L. Reiss[1,3], John D. E. Gabrieli[6,7], and Fumiko Hoeft[1,3]**

[1]Center for Interdisciplinary Brain Sciences Research, Stanford University School of Medicine; [2]Pacific Graduate School of Psychology, Palo Alto University; [3]Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine; [4]Graduate School of Social Work, Boston College; [5]Department of Psychology, York University; [6]Department of Brain and Cognitive Sciences, Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology; and [7]McGovern Institute for Brain Research, Massachusetts Institute of Technology

## Abstract

Although the role of IQ in developmental dyslexia remains ambiguous, the dominant clinical and research approaches rely on a definition of dyslexia that requires reading skill to be significantly below the level expected given an individual's IQ. In the study reported here, we used functional MRI (fMRI) to examine whether differences in brain activation during phonological processing that are characteristic of dyslexia were similar or dissimilar in children with poor reading ability who had high IQ scores (discrepant readers) and in children with poor reading ability who had low IQ scores (nondiscrepant readers). In two independent samples including a total of 131 children, using univariate and multivariate pattern analyses, we found that discrepant and nondiscrepant poor readers exhibited similar patterns of reduced activation in brain areas such as left parietotemporal and occipitotemporal regions. These results converge with behavioral evidence indicating that, regardless of IQ, poor readers have similar kinds of reading difficulties in relation to phonological processing.

## Keywords

academic achievement, aptitude measures, dyslexia, neuroimaging, reading

Received 4/20/11; Revision accepted 6/3/11

Learning to read may be the most critical skill that children acquire in early education because reading ability is necessary to access much of the educational curriculum. In the United States, 5% to 10% of children in first through fifth grades are reported to have developmental dyslexia (Shaywitz, Escobar, Shaywitz, Fletcher, & Makugh, 1992). A critical issue with important educational, clinical, and theoretical implications is how to define dyslexia. Defining dyslexia as reading achievement below what would be expected given the individual's IQ has been central to considering dyslexia as a specific learning difficulty. Although the 2004 reauthorization of the Individuals with Disabilities Education Act mandates that states can no longer require school districts to use IQ tests to identify individuals with learning disabilities (Fletcher, Lyon, Fuchs, & Barnes, 2006), the majority of U.S. schools and school psychologists still rely on the discrepancy between reading achievement and IQ to define dyslexia (Machek & Nelson, 2007). The discrepancy standard posits that reading difficulties in the presence of intact general intellectual ability may arise from different causes and require different forms of treatment than reading difficulties accompanied by lower intellectual ability.

In contrast to the assumptions behind discrepancy-based definitions, results from a number of behavioral studies indicate that the underlying phonological deficit is similar in both discrepant poor readers (low reading score and higher IQ score, a combination that supports the discrepancy model) and nondiscrepant poor readers (low reading score and low IQ score, a combination that supports the alternative low-achievement model; Fletcher et al., 1994; O'Malley, Francis, Foorman, Fletcher, & Swank, 2002; Stanovich & Siegel, 1994; Stuebing et al., 2002; Tunmer & Greaney, 2010). Further, both kinds of poor readers respond similarly to structured, phonetically based remedial reading programs designed to ameliorate phonological deficits (Stuebing et al., 2002; Vellutino, Scanlon, Small, & Fanuele, 2006). Moreover, longitudinal analyses indicate that there is a decoupling between reading and IQ in poor readers, and in particular, those who do not compensate for poor reading over time (Ferrer, Shaywitz, Holahan, Marchione, & Shaywitz, 2010). These findings suggest that

**Corresponding Author:**
Fumiko Hoeft, 401 Quarry Rd., MC5795, Stanford, CA 94305
E-mail: fumiko@stanford.edu

the underlying brain basis of reading failure is similar in all children with low reading scores whether or not those scores are discrepant from performance on IQ-type measures of broader intellectual ability (Stanovich, 2005).

In the study reported here, we used functional MRI (fMRI) to examine whether poor reading associated with impaired phonological processing of printed words involved similar or dissimilar brain processes in two independent samples of matched groups of children who were either discrepant poor readers (poor reading scores and normal IQ estimates) or nondiscrepant poor readers (poor reading scores and low IQ estimates). We compared the brain activation patterns of both groups to the patterns of typically developing children. To the extent that a neurophysiological difference underlies impaired phonological processing of printed words in both children with high IQ scores and children with low IQ scores, we expected the two groups of poor readers to exhibit similar patterns of brain-activation differences relative to typically developing readers. Alternatively, different patterns of brain activation in the two groups would suggest a different neurophysiological basis for impaired phonological processing that is related to IQ.

Studies of functional brain differences in dyslexia have frequently reported reduced left-hemisphere activations in a neural circuit implicated in reading and language, including inferior frontal, parietotemporal, and occipitotemporal regions (Hoeft et al., 2006, 2007; Maisog, Einbinder, Flowers, Turkeltaub, & Eden, 2008; Richlan, Kronbichler, & Wimmer, 2009). Studies of dyslexia in general have overwhelmingly used a discrepancy criterion for inclusion of participants with dyslexia. Because we used the same definition of dyslexia as these studies did, we expected to find similar results, namely, reduced activation in these brain regions in the poor readers with high IQs. It was unknown, however, whether poor readers with low IQs would also show a similar pattern of reduced activation in these brain regions.

## Method

### *Participants*

We collected data at two sites, Carnegie Mellon University (CMU) and Stanford University (Table 1; see also Table S1 in the Supplemental Material available online). Fifty-seven participants ages 8.5 to 12.6 years ($M$ = 10.3 years, $SD$ = 1.1) were drawn from a larger study at CMU of third- and fifth-grade typical and poor readers from public schools surrounding Pittsburgh in Allegheny County, Pennsylvania (Torgesen et al., 2006). Seventy-four typical and poor readers ages 7.7 to 16.9 years ($M$ = 13.4 years, $SD$ = 2.5) were recruited from the

**Table 1.** Descriptive Statistics for the Samples in the Study

| Sample and measure | Typical readers | Discrepant poor readers | Nondiscrepant poor readers |
|---|---|---|---|
| Carnegie Mellon University | | | |
| Age (years) | 10.0 (1.0)$_a$ | 10.3 (1.0)$_{a,b}$ | 10.9 (1.1)$_b$ |
| PPVT score | 114.2 (10.6)$_a$ | 103.8 (10.5)$_b$ | 82.6 (5.2)$_c$ |
| WRMT score: Word Identification subtest | 109.6 (12.3)$_a$ | 81.7 (9.5)$_b$ | 84.3 (5.5)$_b$ |
| Discrepancy (PPVT score – WRMT Word Identification subtest score) | 4.6 (10.3)$_a$ | 22.1 (17.6)$_b$ | −1.7 (7.1)$_a$ |
| WRMT score: Word Attack subtest | 114.6 (13.7)$_a$ | 88.6 (9.4)$_b$ | 89.1 (8.6)$_b$ |
| WRMT score: Passage Comprehension subtest | 112.8 (10.3)$_a$ | 87.8 (14.3)$_b$ | 87.2 (11.3)$_b$ |
| Task performance (% correct) | 95.2 (6.7)$_a$ | 71.9 (18.2)$_b$ | 73.7 (19.9)$_b$ |
| Stanford University | | | |
| Age (years) | 12.7 (3.0)$_a$ | 14.1 (1.8)$_a$ | 14.0 (1.6)$_a$ |
| PPVT score | 116.4 (13.8)$_a$ | 99.2 (7.9)$_b$ | 80.2 (8.4)$_c$ |
| WRMT score: Word Identification subtest | 112.1 (11.3)$_a$ | 82.5 (6.5)$_b$ | 79.8 (7.7)$_b$ |
| Discrepancy (PPVT score – WRMT Word Identification subtest score) | 4.3 (14.2)$_a$ | 16.7 (8.2)$_b$ | 0.5 (9.2)$_a$ |
| WRMT score: Word Attack subtest | 109.9 (10.8)$_a$ | 87.3 (6.3)$_b$ | 89.0 (9.2)$_b$ |
| WRMT score: Passage Comprehension subtest | 113.9 (8.5)$_a$ | 90.1 (9.8)$_b$ | 79.5 (8.5)$_c$ |
| Task performance (% correct) | 94.9 (6.8)$_a$ | 81.7 (13.1)$_b$ | 77.5 (11.6)$_b$ |

Note: The table presents means with standard deviations in parentheses, and standardized scores are reported for all tests. The Carnegie Mellon University sample included 26 typical readers, 16 discrepant poor readers, and 15 nondiscrepant poor readers, and the Stanford University sample included 36 typical readers, 18 discrepant poor readers, and 20 nondiscrepant poor readers. Within each row, values with the same subscript are not significantly different ($p$ < .05). PPVT = Peabody Picture Vocabulary Test (Dunn & Dunn, 1997); WRMT = Woodcock Reading Mastery Tests–Revised/Normative Update (Woodcock, 1998).

Stanford–San Francisco Bay Area. Participants were all native English speakers. Exclusion criteria for both studies were the diagnosis of a neurological or psychiatric disorder (e.g., sensory disorders, attention-deficit/hyperactivity disorder), the use of psychotropic medication, the presence of contraindication to MRI (e.g., metal in the subject's body), or a combination of any of these factors. This study was approved by the institutional review boards at both sites. Written informed consent and assent forms were collected from parents and their children, respectively.

## Group assignment

Participants were assigned to one of three groups on the basis of their performance on the Word Identification (WID) subtest of the Woodcock Reading Mastery Tests–Revised/Normative Update (Woodcock, 1998), which is a single-word reading measure, and the third edition of the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 1997), which was used to estimate IQ. These measures have been used in several previous studies to assess reading ability and IQ (Hoeft et al., 2006; Hurford, Schauf, Bunce, Blaich, & Moore, 1994) and in behavioral studies to examine the discrepancy model and the low-achievement model (Stuebing et al., 2002). PPVT scores are highly correlated with full-scale IQ scores from other measures ($r$ = .90; Dunn & Dunn, 1997). Children were classified as having low reading achievement if they scored equal to or less than the 25th percentile (standardized score ≤ 90; O'Malley et al., 2002; Stanovich & Siegel, 1994) on the WID. Children were classified as having low IQ if they had an estimated IQ of equal to or less than the 25th percentile (standardized score ≤ 90). Using these criteria, three groups were identified: (a) typical readers (i.e., children with typical reading ability and IQ; $n$ = 26 and $n$ = 36 for the CMU and Stanford samples, respectively), (b) poor readers with typical IQ (discrepant poor readers; $n$ = 16 and $n$ = 18, respectively), and (c) poor readers with low IQ (nondiscrepant poor readers; $n$ = 15 and $n$ = 20, respectively).

There were no significant differences in demographic variables among the three groups except that the typical readers were significantly younger (< 1 year) than the nondiscrepant poor readers in the CMU sample only ($p$ = .04). Socioeconomic status (SES) as measured by parental education was also not significantly different between the two groups (CMU sample: $p$ = .16; Stanford sample: $p$ = .31).

## fMRI task design

To assess brain activation associated with awareness of the phonology of printed words, we administered a block-design word-rhyme task, with alternating rhyme and rest conditions, to participants in the scanner (Hoeft et al., 2006, 2007). During the rhyme condition, participants judged whether or not two visually presented words rhymed (e.g., "bait," "gate") or did not (e.g., "price," "miss"), and they responded with a right- or left-handed button press, respectively. Word pairs were selected so that the visual appearance of the last letters of the two words could not be used to determine whether the words rhymed. Stimuli were balanced for frequency of occurrence, number of letters, and syllables between rhyme and nonrhyme trials and across blocks (Zeno, Ivens, Millard, & Duvvuri, 1995; for the full list of stimuli, see the supplemental material in Hoeft et al., 2006).

Each 6-s trial consisted of a 4-s presentation of two words followed by a 2-s fixation cross. Each rhyme block began with a 2-s cue period followed by five trials (32 s total). During rest blocks, participants saw a fixation cross on the screen for either 16 s (CMU sample) or 15 s (Stanford sample). The entire scan lasted 234 s (CMU sample) or 223 s (Stanford sample), including two practice trials at the beginning, and consisted of four rhyme blocks and five rest blocks. Image acquisition and preprocessing methods are detailed in Supplementary Method in the Supplemental Material.

## fMRI univariate analyses

We performed univariate modeling of fMRI data to compare regional brain activation among the three groups. We identified six regions of interest based on previous neuroimaging reports of dyslexia (Brunswick, McCrory, Price, Frith, & Frith, 1999; Hoeft et al., 2007; Kronbichler et al., 2006; Maisog et al., 2008; Paulesu et al., 2001; Richlan et al., 2009; Rumsey et al., 1997; Shaywitz et al., 1998). These regions were combined using the Automated Talairach Atlas Label (Tzourio-Mazoyer et al., 2002) in the Wake Forest University PickAtlas toolbox (Maldjian, Laurienti, Kraft, & Burdette, 2003) to form one mask comprised of bilateral inferior frontal (pars triangularis, pars opercularis), parietotemporal (inferior parietal lobule, or IPL), and occipitotemporal (fusiform gyrus, or FG) regions. Conjunction analyses were performed (following the conjunction null method in Nichols, Brett, Andersson, Wager, & Poline, 2005) with a random-effects model (Friston, Holmes, Price, Buchel, & Worsley, 1999) using the rhyme versus rest contrast images to identify brain regions that showed significantly greater activation for typical readers than for both discrepant and nondiscrepant poor readers. A voxel-wise statistical threshold ($p$) of .05 (false discovery rate after small-volume correction) was applied. Because there was a significant difference in age between typical readers and nondiscrepant poor readers in the CMU data set, we also repeated the analyses with age as a covariate.

## fMRI multivoxel pattern analyses

Multivoxel pattern analysis (MVPA) was used to examine all voxels; pattern-classification algorithms were used to identify naturally occurring groupings of children. We performed leave-one-out linear support vector machine (SVM) analyses (Regularization Parameter C = 1) using in-house MATLAB-based (The MathWorks, Natick MA) tools, which adopt LIBSVM software (a library for support vector machines;

Chang & Lin, 2011) and have been used successfully in several prior studies (Gothelf et al., 2011; Hoeft et al., 2008; Hoeft, McCandliss, et al., 2011; Hoeft, Walter, et al., 2011; Marzelli, Hoeft, Hong, & Reiss, 2011).

First, we constructed a class vector constituting either +1 or –1 and assigned each child from the three groups to one of two class labels (either Group +1 or Group –1) depending on the analysis. Next, we converted contrast images (greater activation during rhyme than during rest) into an $N$-by-$V$ matrix, where $N$ is the number of subjects and $V$ is the number of voxels (2 × 2 × 2 mm), and normalized all matrix elements so that the mean was 0 and the standard deviation was 1. Numbers of features were reduced by using a gray-matter mask to include only gray-matter voxels, by performing principal component analyses and then transforming the matrix to principal components, and by performing recursive feature elimination iteratively, removing 30% of worst-discriminating features at a time until performance started deteriorating (De Martino et al., 2008).

Classification accuracy (whether children who actually belonged to Groups +1 and –1 were classified correctly relative to the total number of subjects), sensitivity (whether the total proportion of children who belonged to Group +1 were correctly classified), specificity (whether the proportion of children who actually belonged to Group –1 were correctly classified), and positive predictive value were calculated for each classification. Further, distance from the classifier dividing each group was also measured, such that the further the child was from the hyperplane dividing Groups +1 and –1, the more confident the classifier was of that child's group membership (e.g., a large positive distance measure of a child who had a group label of +1 indicated that the brain-based classifier was highly confident that the child was in the +1 group). This distance was plotted for each participant from each classifier (i.e., the division separating typical readers from discrepant poor readers, and the division separating typical readers from nondiscrepant poor readers) to examine the overall classification pattern for all participants.

All procedures were performed by keeping training data to construct the classifier and test data independent using leave-one-out cross-validation to avoid overfitting and allow generalization of the models (classifiers). Significance was determined using permutation analysis by randomly reassigning class labels 2,000 times ($p < .05$). Brain maps were constructed by transforming features (principal components) that remained during recursive feature elimination back into voxel space. For visualization purposes, we used a significance level of .05 in these permutation analyses to show brain regions that carried significant positive and negative weights.

Using this approach, we performed a series of classification analyses to examine the similarities between the two groups of poor readers and the differences between these two groups and typical readers. First, we performed SVM analyses that paralleled our univariate analyses. Using data from both the Stanford and CMU samples, we examined whether differences in brain-activation patterns could be used to discriminate between typical readers and the combined group of poor readers (discrepant and nondiscrepant), and between the two groups of poor readers.

Second, we performed an additional MVPA to address the possibility that failure to find significant differences between the two groups of poor readers would reflect only null or negative findings. We performed pattern classification between one group of poor readers (e.g., discrepant poor readers) and typical readers and applied the resulting classifiers to the other group of poor readers (e.g., nondiscrepant poor readers). If a classifier developed to discriminate typical readers from one group of poor readers also significantly classified the other group of poor readers as poor readers, then this would constitute positive evidence that the two groups of poor readers are described by a common pattern of brain activation. In order to test this hypothesis, we first performed leave-one-out SVM (leaving one child from either group out at a time, hence producing as many classifiers as there were participants) to train the model to distinguish between discrepant poor readers and typical readers, and we used these data as training data. Then data from nondiscrepant poor readers were used as test data for each classifier to examine how likely each nondiscrepant poor reader was to be classified as a discrepant poor reader. Similarly, we examined how likely discrepant poor readers were to be classified as nondiscrepant poor readers rather than typical readers when the classifier derived from nondiscrepant poor readers versus typical readers was applied to discrepant poor readers. These analyses were performed for both samples.

## Results

### Behavioral analysis

Typical readers, compared with the two groups of poor readers, showed significantly higher reading-related scores and more accurate performance on the rhyme-judgment task, but there were no significant differences between the two groups of poor readers on these measures (Table 1; see also Table S1). IQ scores of the typical readers were significantly higher than IQ scores of the discrepant poor readers, who had significantly higher IQ scores than the nondiscrepant poor readers. The discrepant poor readers showed a significant difference between reading ability and IQ scores, whereas the nondiscrepant poor readers did not.

### fMRI univariate analyses

In the CMU sample, both discrepant and nondiscrepant poor readers exhibited significantly lower activations relative to typical readers in left IPL (Talairach coordinates: $x = −32$, $y = −47$, $z = 41$; $Z = 4.05$, $p = .02$, corrected) and in left FG (Talairach coordinates: $x = −44$, $y = −53$, $z = −14$; $Z = 4.05$, $p = .02$, corrected; see Fig. 1). The two groups of poor readers did not

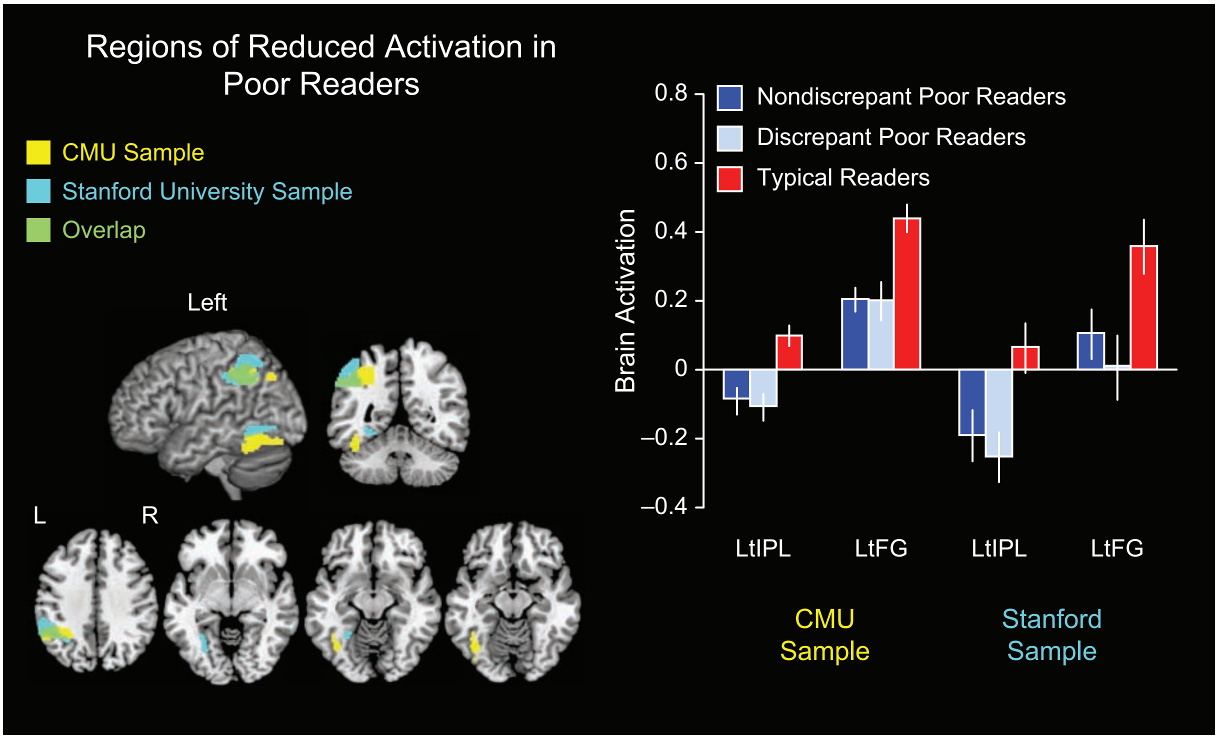


**Fig. 1.** Results of analyses examining increased brain activation during the word-rhyme task (relative to rest). The brain images show regions where this activation was reduced in the two groups of poor readers combined, relative to typical readers. Results are shown for both the Carnegie Mellon University (CMU) sample and the Stanford University sample. The graphs show parameter estimates for left inferior parietal lobule (LtIPL) and left inferior fusiform gyrus (LtFG) in the three groups of readers for both the CMU and the Stanford samples. Error bars represent standard errors of the mean. L = left; R = right.

exhibit significant differences in activation from one another. Adding age as a covariate to the analysis in the CMU sample did not change the results (see Fig. S1 in the Supplemental Material).

In the Stanford University sample, both groups of poor readers also exhibited significantly lower activations relative to typical readers in left IPL (Talairach coordinates: $x = -55$, $y = -42$, $z = 48$; $Z = 3.82$, $p = .046$, corrected) and in left FG (Talairach coordinates: $x = -26$, $y = -70$, $z = 0$; $Z = 3.47$, $p = .046$, corrected; Fig. 1), and, again, the two groups of poor readers did not significantly differ from one another.

## *fMRI MVPAs*

Table 2 shows classification accuracy for four sets of analyses. In the CMU sample, typical and poor readers (discrepant and nondiscrepant combined) were discriminated significantly from one another with an accuracy of 78.9% (sensitivity = 83.9%, specificity = 73.1%; $p < .001$). Brain regions that contributed to the classification of typical and poor readers included left IPL, FG, IFG, caudate, insula, and middle temporal gyrus (Fig. 2a). Discrimination between the two groups of poor readers was not reliably above chance (accuracy = 64.5%, $p = .16$). Analysis of the Stanford University sample yielded similar results with a discrimination accuracy of 79.7% (sensitivity = 76.3%, specificity = 83.3%; $p < .001$) between typical and poor readers, and 44.7% ($p > .1$) between the two groups of poor readers.

**Table 2.** Multivariate Pattern Classification Results

| Classification and group | Classification accuracy (%) |
|---|---|
| Typical vs. poor readers | |
| Carnegie Mellon University | 78.90* |
| Stanford University | 79.70* |
| Discrepant vs. nondiscrepant poor readers | |
| Carnegie Mellon University | 64.50 |
| Stanford University | 44.70 |
| Nondiscrepant poor readers classified as discrepant poor readers | |
| Carnegie Mellon University | 81.30* |
| Stanford University | 98.00* |
| Discrepant poor readers classified as nondiscrepant poor readers | |
| Carnegie Mellon University | 86.70* |
| Stanford University | 71.30* |

Note: Asterisks indicate that accuracy was significantly better than chance ($p < .001$).

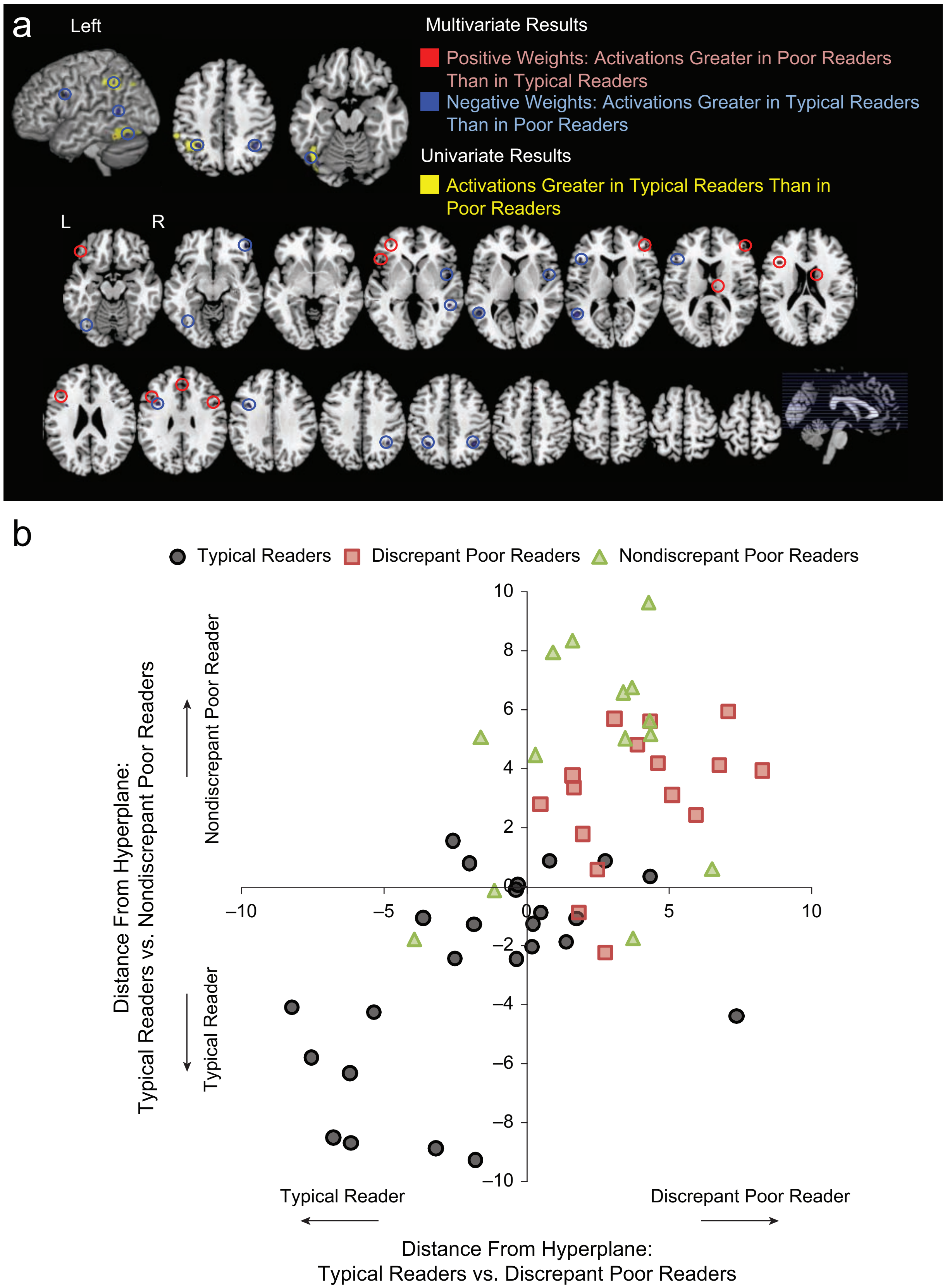


**Fig. 2.** Results from the Carnegie Mellon University sample. The brain images (a) show activation differences between typical readers and the two groups of poor readers combined (results are shown for univariate analyses and for multivoxel pattern analyses, or MVPAs). For display purposes, we show MVPA results using a statistical threshold ($p$) of .05 (2,000 permutations) to show brain regions that carried significant positive and negative weights. The scatter plot (b), which is based on MVPA, shows each participant's distance from the hyperplanes discriminating typical readers from IQ-discrepant poor readers ($x$-axis) and discriminating typical readers from non-IQ-discrepant poor readers ($y$-axis). Typical readers with positive values and discrepant and nondiscrepant poor readers with negative values are those who were misclassified. L = left; R = right.

To show a positive similarity between the two groups of poor readers, we applied the classification model that discriminated between typical readers and one group of poor readers to classify the other group of poor readers. In the CMU sample, the classification model that discriminated typical readers from discrepant poor readers significantly classified nondiscrepant poor readers as discrepant poor readers with 81.3% accuracy ($p < .001$, permutation based). The classification model that discriminated typical readers from nondiscrepant poor readers also classified discrepant poor readers as nondiscrepant poor readers with a significant accuracy of 86.7% ($p < .001$). Identical analyses of the Stanford data showed that discrepant poor readers were classified as nondiscrepant poor readers with 98.0% accuracy, and nondiscrepant poor readers were classified as discrepant poor readers with 71.8% accuracy (both $p$s $< .001$).

To show the similarity between the two groups of poor readers, we calculated each subject's distance from two hyperplanes (Fig. 2b). A hyperplane is a multidimensional plane that optimally divides two groups. Distance from the hyperplane indicates how much that particular individual (vector) is like the group on the same side of the hyperplane. In this study, the two groups of poor readers were both contrasted with typical readers, and the resulting hyperplanes were overlaid orthogonally to produce four separate classification quadrants. Generally, typical readers were clustered in the same lower left quadrant, and the two groups of poor readers were clustered within the same upper right quadrant. These groupings indicate that the two groups of poor readers show similar brain-activation patterns.

## Discussion

Overall, atypical brain function for the phonological processing of printed words was highly similar in two carefully matched groups of poor readers who had IQ estimates that were either discrepant (higher IQ) or nondiscrepant (equivalent IQ) with their poor reading scores. Although typical univariate analyses and pattern-classification methods of brain activation reliably distinguished patterns of brain activation between typical and poor readers, there were no reliable functional brain differences between the two types of poor readers. The shared reductions of activation occurred in two left-hemisphere brain regions that often exhibit reduced activation in dyslexia: the left FG, which is thought to be important for specialized visual analysis of print, and the left IPL, which may be important for relating print to sound. Altogether, multiple independent analyses of brain activations in two independent samples converge on the conclusion that the brain regions implicated in weakness in phonological awareness, which is thought to be the main deficit in dyslexia, are similar in poor readers irrespective of their IQ scores.

There was great similarity in patterns of brain activation between poor readers with discrepant IQs and poor readers with nondiscrepant IQs, but a concern about such similarity is that it is based on the absence of statistical differences between the groups of poor readers. Three findings, however, mitigate this concern. First, the two poor reading groups in the two samples showed reliable differences relative to typical readers: The former groups showed reduced activations in left IPL and FG regions, so the study had suitable and replicable statistical power to reveal brain-activation differences. Second, the two groups in the two samples could reliably be discriminated from typical readers in MVPAs. None of these analyses, however, could reliably differentiate the two groups of poor readers from each other. Third, and perhaps most compelling, is the positive evidence in both samples that the statistical models from the MVPAs that reliably discriminated one group of poor readers from typical readers classified the other group of poor readers as poor readers rather than as typical readers.

Our findings are consistent with substantial, converging evidence, including a prior neuroimaging study (Temple et al., 2001), and this consistency suggests that the relationship between IQ and the phonological-awareness deficit underlying dyslexia may be quite weak (O'Malley et al., 2002; Stuebing et al., 2002; Tunmer & Greaney, 2010). Thus, the validity of the discrepancy definition of dyslexia is called into question, despite several lines of reasoning that have seemed to support that definition. From a research perspective, behavioral difficulties are often most easily analyzed in the context of strong dissociations in which a single disability is isolated among many spared abilities such that the disability does not seem secondary to other deficits. From a clinical or educational perspective, remediation seems most targeted and effective when it addresses an isolated disability. Further, in unselected populations, there is a correlation between IQ and reading ability (in the .3 to .6 range; Hulme & Snowling, 2009), suggesting some link between the broad cognitive abilities assayed by IQ-type measures and reading ability, although it appears that dyslexia may break this link (Ferrer et al., 2010). Finally, it seems likely that children or adults with broad and severe cognitive deficits, well below those of the nondiscrepant children in the study reported here, would fail to read as a secondary consequence of their cognitive disabilities and therefore would not benefit from interventions focused on reading skills per se. In sum, there are a number of complex factors that have encouraged the scientific and educational communities to rely on the apparent straightforwardness of the discrepancy criterion.

The discrepancy criterion, however, seems to lack validity and reliability in predicting the course of reading failure, the response to remedial intervention (Stuebing et al., 2002; Vellutino et al., 2006), or the brain dysfunctions that underlie dyslexia. Although the discrepancy criterion may be intuitively appealing, its strict application would deprive nondiscrepant children of the educational interventions that could promote their advancement in reading ability. Further, the exclusion of nondiscrepant children from research studies examining the genetic, neural, and psychological bases of dyslexia would slow progress in truly understanding dyslexia by arbitrarily excluding many children from such research studies. Expanding the definition of dyslexia to include children with nondiscrepant IQ scores increases the dissociation of reading from broad cognitive abilities by suggesting that the impairment in phonological awareness leading to dyslexia can occur across a broad range of IQ abilities.

There are several limitations to this study, including the following. First, IQ was estimated using the PPVT, which does not require reading, is a strong indicator of general verbal ability, correlates highly with full-scale IQ ($r$ = .90; Dunn & Dunn, 1997), and has been used in many studies investigating the effect of IQ in reading outcome (Stuebing et al., 2002). Previous studies have demonstrated that the choice of IQ measure (e.g., verbal vs. nonverbal) has not improved the validity of the IQ-discrepancy model (Fletcher et al., 1994; O'Malley et al., 2002; Stanovich & Siegel, 1994; Stuebing et al., 2002). Verbal IQ tests, such as vocabulary, tend to show the most resistance to effects of neurobiologic alterations (Brown et al., 2011), but they are influenced by differences in environmental enrichment and opportunity (Stern, 2009). The groups in the study reported here were matched for SES (indeed, the control participants in the CMU sample were classmates of the poor readers), so it is unlikely that SES differences accounted for any findings. Alternate IQ tests, however, might yield different results.

A second limitation of this study is that the functional brain differences were found by comparison of a task that demanded phonological awareness of the sounds of printed words relative to a rest condition, and other functional contrasts may reveal additional differences between poor readers. For example, the fMRI task used here does not distinguish between the initial underlying cause and the additional consequence of reduced reading experience in poor readers. Although the poor-reader groups were not different in terms of brain activation in reading-related regions, they might show differences using other imaging techniques, such as volumetric or diffusion-tensor imaging studies, that may reflect the effect of IQ. Additionally, although the entire sample utilized in this study was quite large, sample sizes in each reading group may have been insufficient to detect effects. Continued studies using larger groups, more comprehensive cognitive-behavioral assessment, and complementary neuroimaging methods are required.

In summary, it has so far proven remarkably difficult to assign precise psychological characterizations to neurobehavioral disorders, perhaps because of the many difficulties that occur when defining current diagnostic categories. It has been a hope that biological measures, such as genetics and neuroimaging, would provide new insights into these disorders that would, in turn, help restructure diagnostic categorizations into more precise and validated taxonomies. In the study reported here, convergent psychological, educational, and neurobiological evidence suggests that the long-standing and widely applied diagnosis of dyslexia by IQ discrepancy is not supportable. The evidence indicates that any child with a reading difficulty, regardless of his or her general level of cognitive abilities (IQ), should be encouraged to seek reading intervention.

## Acknowledgments

Hiroko Tanaka and Jessica M. Black contributed equally to this research. Portions of this work were previously included in an oral presentation at the annual meeting of the Society for the Scientific Study of Reading, June 2009, in Boston, MA, and as a poster presentation at the annual convention of the Association for Psychological Science, May 2009, in San Francisco, CA.

## Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

## Funding

This work was supported by grants from the William and Flora Hewlett Foundation, the Richard King Mellon Foundation, and the Ellison Medical Foundation to J. D. E. G., and by grants from the National Institute of Child Health and Human Development (Grant K23HD054720) and the Dyslexia Foundation, as well as a Lucile Packard Foundation for Children's Health, Spectrum Child Health & Clinical and Translational Science Award and the National Alliance for Research in Schizophrenia and Depression Young Investigator Award, to F. H.

## Supplemental Material

Additional supporting information may be found at http://pss.sagepub.com/content/by/supplemental-data

## References

Brown, L. J., Ferner, H. S., Robertson, J., Mills, N. L., Pessotto, R., Deary, I. J., & MacLullich, A. M. (2011). Differential effects of delirium on fluid and crystallized cognitive abilities. *Archives of Gerontology and Geriatrics*, *52*, 153–158. doi:10.1016/j.archger.2010.03.005

Brunswick, N., McCrory, E., Price, C. J., Frith, C. D., & Frith, U. (1999). Explicit and implicit processing of words and pseudowords by adult developmental dyslexics. *Brain*, *122*, 1901–1917.

Chang, C.-C., & Lin, C.-J. (2011). LIBSVM—A library for support vector machines (Version 3.1) [Computer software]. Retrieved from http://www.csie.ntu.edu.tw/~cjlin/libsvm/

De Martino, F., Valente, G., Staeren, N., Ashburner, J., Goebel, R., & Formisano, E. (2008). Combining multivariate voxel selection and support vector machines for mapping and classification of fMRI spatial patterns. *NeuroImage*, *43*, 44–58.

Dunn, L. M., & Dunn, D. M. (1997). *Peabody Picture Vocabulary Test* (3rd ed.). Circle Pines, MN: American Guidance Service.

Ferrer, E., Shaywitz, B. A., Holahan, J. M., Marchione, K., & Shaywitz, S. E. (2010). Uncoupling of reading and IQ over time: Empirical evidence for a definition of dyslexia. *Psychological Science*, *21*, 93–101. doi:10.1177/0956797609354084

Fletcher, J. M., Lyon, G. R., Fuchs, L. S., & Barnes, M. (2006). *Learning disabilities: From identification to intervention*. New York, NY: Guilford Press.

Fletcher, J. M., Shaywitz, S. E., Shankweiler, D. P., Katz, L., Liberman, I. Y., Stuebing, K. K., . . . Shaywitz, B. A. (1994). Cognitive profiles of reading disability: Comparisons of discrepancy and low achievement definitions. *Journal of Educational Psychology*, *86*, 6–23.

Friston, K. J., Holmes, A. P., Price, C. J., Buchel, C., & Worsley, K. J. (1999). Multisubject fMRI studies and conjunction analyses. *NeuroImage*, *10*, 385–396.

Gothelf, D., Hoeft, F., Ueno, T., Sugiura, L., Lee, A. D., Thompson, P., & Reiss, A. L. (2011). Developmental changes in multivariate neuroanatomical patterns that predict risk for psychosis in 22q11.2 deletion syndrome. *Journal of Psychiatric Research*, *45*, 322–331. doi:10.1016/j.jpsychires.2010.07.008

Hoeft, F., Hernandez, A., McMillon, G., Taylor-Hill, H., Martindale, J. L., Meyler, A., . . . Gabrieli, J. (2006). Neural basis of dyslexia: A comparison between dyslexic and nondyslexic children equated for reading ability. *Journal of Neuroscience*, *26*, 10700–10708.

Hoeft, F., Lightbody, A. A., Hazlett, H. C., Patnaik, S., Piven, J., & Reiss, A. L. (2008). Morphometric spatial patterns differentiating boys with fragile X syndrome, typically developing boys, and developmentally delayed boys aged 1 to 3 years. *Archives of General Psychiatry*, *65*, 1087–1097.

Hoeft, F., McCandliss, B. D., Black, J. M., Gantman, A., Zakerani, N., Hulme, C., . . . Gabrieli, J. D. E. (2011). Neural systems predicting long-term outcome in dyslexia. *Proceedings of the National Academy of Sciences, USA*, *108*, 361–366.

Hoeft, F., Meyler, A., Hernandez, A., Juel, C., Taylor-Hill, H., Martindale, J. L., . . . Gabrieli, J. D. E. (2007). Functional and morphometric brain dissociation between dyslexia and reading ability. *Proceedings of the National Academy of Sciences, USA*, *104*, 4234–4239.

Hoeft, F., Walter, E., Lightbody, A. A., Hazlett, H. C., Chang, C., Piven, J., & Reiss, A. L. (2011). Neuroanatomical differences in toddler boys with fragile X syndrome and idiopathic autism. *Archives of General Psychiatry*, *68*, 295–305.

Hulme, C., & Snowling, M. J. (2009). *Developmental disorders of language learning and cognition*. West Sussex, England: Wiley-Blackwell.

Hurford, D. P., Schauf, J. D., Bunce, L., Blaich, T., & Moore, K. (1994). Early identification of children at risk for reading disabilities. *Journal of Learning Disabilities*, *27*, 371–382.

Kronbichler, M., Hutzler, F., Staffen, W., Mair, A., Ladurner, G., & Wimmer, H. (2006). Evidence for a dysfunction of left posterior reading areas in German dyslexic readers. *Neuropsychologia*, *44*, 1822–1832.

Machek, G. R., & Nelson, J. M. (2007). How should reading disabilities be operationalized? A survey of practicing school psychologists. *Learning Disabilities Research & Practice*, *22*, 147–157.

Maisog, J. M., Einbinder, E. R., Flowers, D. L., Turkeltaub, P. E., & Eden, G. F. (2008). A meta-analysis of functional neuroimaging studies of dyslexia. *Annals of the New York Academy of Sciences*, *1145*, 237–259.

Maldjian, J. A., Laurienti, P. J., Kraft, R. A., & Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage*, *19*, 1233–1239.

Marzelli, M. J., Hoeft, F., Hong, D. S., & Reiss, A. L. (2011). Neuroanatomical spatial patterns in Turner syndrome. *NeuroImage*, *55*, 439–447.

Nichols, T., Brett, M., Andersson, J., Wager, T., & Poline, J. B. (2005). Valid conjunction inference with the minimum statistic. *NeuroImage*, *25*, 653–660.

O'Malley, K. J., Francis, D. J., Foorman, B. R., Fletcher, J. M., & Swank, P. R. (2002). Growth in precursor and reading-related skills: Do low-achieving and IQ-discrepant readers develop differently? *Learning Disabilities Research & Practice*, *17*, 19–34.

Paulesu, E., Demonet, J. F., Fazio, F., McCrory, E., Chanoine, V., Brunswick, N., . . . Frith, U. (2001). Dyslexia: Cultural diversity and biological unity. *Science*, *291*, 2165–2167.

Richlan, F., Kronbichler, M., & Wimmer, H. (2009). Functional abnormalities in the dyslexic brain: A quantitative meta-analysis of neuroimaging studies. *Human Brain Mapping*, *30*, 3299–3308.

Rumsey, J. M., Nace, K., Donohue, B., Wise, D., Maisog, J. M., & Andreason, P. (1997). A positron emission tomographic study of impaired word recognition and phonological processing in dyslexic men. *Archives of Neurology*, *54*, 562–573.

Shaywitz, S. E., Escobar, M. D., Shaywitz, B. A., Fletcher, J. M., & Makugh, R. (1992). Evidence that dyslexia may represent the lower tail of a normal distribution of reading ability. *New England Journal of Medicine*, *326*, 145–150.

Shaywitz, S. E., Shaywitz, B. A., Pugh, K. R., Fulbright, R. K., Constable, R. T., Mencl, W. E., . . . Gore, J. C. (1998). Functional disruption in the organization of the brain for reading in dyslexia. *Proceedings of the National Academy of Sciences, USA*, *95*, 2636–2641.

Stanovich, K. E. (2005). The future of a mistake: Will discrepancy measurement continue to make the learning disabilities field a pseudoscience? *Learning Disability Quarterly*, *28*, 103–106.

Stanovich, K. E., & Siegel, L. S. (1994). Phenotypic performance profile of children with reading disabilities: A regression-based test of the phonological-core variable-difference model. *Journal of Educational Psychology*, *86*, 24–53.

Stern, Y. (2009). Cognitive reserve. *Neuropsychologia*, *47*, 2015–2028.

Stuebing, K. K., Fletcher, J. M., LeDoux, J. M., Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2002). Validity of IQ-discrepancy classifications of reading disabilities: A meta-analysis. *American Educational Research Journal*, *39*, 469–518.

Temple, E., Poldrack, R. A., Salidis, J., Deutsch, G. K., Tallal, P., Merzenich, M. M., & Gabrieli, J. D. (2001). Disrupted neural responses to phonological and orthographic processing in dyslexic children: An fMRI study. *NeuroReport*, *12*, 299–307.

Torgesen, J., Myers, D., Schirm, A., Stuart, E., Vartivarian, S., Mansfield, W., . . . Haan, C. (2006). *National Assessment of Title I: Interim Report to Congress: Vol. II. Closing the reading* *gap: First year findings from a randomized trial of four reading interventions for striving readers*. Retrieved from http://www2.ed.gov/rschstat/eval/disadv/title1interimreport/index.html

Tunmer, W., & Greaney, K. (2010). Defining dyslexia. *Journal of Learning Disabilities*, *43*, 229–243.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., . . . Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, *15*, 273–289.

Vellutino, F. R., Scanlon, D. M., Small, S., & Fanuele, D. P. (2006). Response to intervention as a vehicle for distinguishing between children with and without reading disabilities: Evidence for the role of kindergarten and first-grade interventions. *Journal of Learning Disabilities*, *39*, 157–169.

Woodcock, R. W. (1998). *Woodcock Reading Mastery Tests–Revised/normative update*. Circle Pines, MN: American Guidance Service.

Zeno, S. M., Ivens, S. H., Millard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide*. New York, NY: Touchstone Applied Science Associates.