Effects of Stimulus Difficulty and Repetition on Printed Word Identification: An fMRI Comparison of Nonimpaired and Reading-disabled Adolescent Cohorts

Kenneth R. Pugh[1,2], Stephen J. Frost[1], Rebecca Sandak[1], Nicole Landi[1],
Jay G. Rueckl[1,3], R. Todd Constable[2], Mark S. Seidenberg[1,4],
Robert K. Fulbright[2], Leonard Katz[1,3], and W. Einar Mencl[1]

Abstract

■ Functional neuroimaging studies indicate that a primary marker of specific reading disability (RD) is reduced activation of left hemisphere (LH) posterior regions during performance of reading tasks. However, the severity of this disruption, and the extent to which these LH systems might be available for reading under any circumstances, is unclear at present. Experiment 1 examined the cortical effects of stimulus manipulations (frequency, imageability, consistency) that have known facilitative effects on reading performance for both nonimpaired (NI) and RD readers. Experiment 2 examined stimulus repetition, another facilitative variable, in an additional sample of adolescent NI and RD readers. For NI readers, factors that made words easier to process were associated with relatively reduced activation. For RD readers, facilitative factors resulted in increased activation in these same reading-related sites, suggesting that the LH reading circuitry in adolescent RD is poorly trained but not wholly disrupted. ■

INTRODUCTION

Converging evidence from functional neuroimaging studies indicates that a primary neurobiological marker of specific reading disability (RD) is reduced activation of left hemisphere (LH) posterior regions relative to activation levels for nonimpaired (NI) readers during tasks that make demands on language and printed word processing. Together with the failure to reliably engage LH temporo-parietal and occipito-temporal regions, RD readers tend also to show heightened activation of right hemisphere (RH) posterior and bilateral frontal regions (see Sarkari et al., 2002; Pugh, Mencl, Jenner, et al., 2000 for reviews); the tendency to hyperengage these regions may serve to compensate for deficient linguistic processing in the LH. Although this RD profile appears to be reasonably stable across different ages, tasks, and languages (Paulesu et al., 2001), the question still remains as to the severity of this LH disruption (Pugh, Mencl, Shaywitz, et al., 2000). Evidence from recent intervention studies suggests that compromised LH systems appear to be responsive to intensive training in young

RD populations (Shaywitz et al., 2004; Temple et al., 2003; Simos et al., 2002). That is, many LH regions that are critically involved in reading and are not activated during reading tasks in young RD readers prior to an intervention period show increased activation after intervention. However, the extent to which these LH systems are available for reading in older children whose reading difficulties have persisted is less studied. Recently, Cao, Bitan, Chou, Burman, and Booth (2006) contrasted NI and RD children in a rhyming task with easy versus hard trials. The often-observed diminished activation of key LH regions in RD was obtained only on hard trials, suggesting that stimulus difficulty is an important variable in making group contrasts. Hoeft et al. (2007) used reading age (RA) and chronological age (CA) controls to assess performance effects on group contrasts, and concluded that hyperactivation in frontal areas in RD is experience and effort-related whereas hypoactivation in RD at LH posterior regions (particularly LH temporo-parietal sites) is more fundamental to the syndrome (indeed, structural imaging reinforced this conclusion; with reliably reduced gray matter volume in temporo-parietal areas in RD). The current experiments were designed to examine learning and difficulty effects in NI and RD more directly by focusing on learning differences as a window on latent functionality

---

[1]Haskins Laboratories, New Haven, CT, [2]Yale University School of Medicine, New Haven, CT, [3]University of Connecticut, [4]University of Wisconsin

in RD. Addressing this issue is of high priority for predicting likely neurobiological and behavioral outcomes of systematic reading intervention and remediation in older children. As a first step, it is critical that we begin to examine whether adolescent RD readers demonstrate reading-related functionality in these LH systems under any conditions.

The current study examines the cortical effects of stimulus manipulations that have known facilitative effects on word reading latencies and accuracy in adolescent RD readers. In Experiment 1, we focus specifically on the ways in which stimulus familiarity (frequency) and semantic features (imageability) modulate processing of words that vary with regard to complexity of orthographic-to-phonological mappings (consistency). This experiment provides a window into how top–down, semantic, facilitative features are able to affect the altered neural circuitry for reading in RD readers. In Experiment 2, we examine stimulus repetition (a highly salient facilitative variable in RD performance) to characterize on-line learning processes in RD readers. By examining on-line learning, we can begin to move beyond first-generation questions of where in the brain activation levels differ in these groups in general, to a more systems-level account of the ways in which readers with very different skill levels modulate activation patterns in the context of learning. We predict that this more dynamic approach will provide a better model for clinical contrast than one that simply looks for regional differences in a more static manner.

EXPERIMENT 1

One of the most often used indices of the influence of phonology on printed word naming is the spelling-to-sound consistency effect. This refers to the finding that identification is faster and more accurate for words that have consistent (1:1) correspondences between the orthographic body and phonological rime (e.g., –ill only corresponds to /Il/ as in *pill*, *mill*) than for inconsistent words that have multiple body–rime correspondences (e.g., –int corresponds to /Int/ as in *mint* and /aInt/ as in *pint*). Behavioral studies have shown that frequency modulates this effect such that consistency effects are most robust for low-frequency words (Jared, McRae, & Seidenberg, 1990; Seidenberg, Waters, Barnes, & Tanenhaus, 1984). Further work by Strain, Patterson, and Seidenberg (1995, 2002) demonstrated that the typically obtained interaction of consistency and frequency during printed word naming is modulated by a semantic variable, imageability. Consistency effects were observed primarily on words that are both low in frequency and imageability; consistency effects for high-imageable, low-frequency words were either attenuated or not significant across experiments. These findings reveal that semantics can attenuate the difficulties

associated with reading words that have inconsistent orthographic-to-phonological mappings.

Research on phonology and lexico-semantics in RD readers has shown that they are particularly challenged by spelling-to-sound inconsistent words and particularly benefited by frequency and imageability. That is, RD readers demonstrate amplified consistency effects relative to NI readers (Bruck, 1992) and greater advantages than skilled readers for high-frequency words relative to low-frequency words in both accuracy and latency (Shaywitz et al., 2003). Moreover, with regard to the top–down influence of semantics (Strain et al., 1995, 2002), poor readers show even greater benefit from imageability than skilled readers on the processing of difficult-to-decode inconsistent words (Strain & Herdman, 1999). Thus, for all reading levels, performance on difficult-to-decode inconsistent words is facilitated when tokens are high frequency and/or highly imageable, but this effect is amplified for poor readers. In summary, these semantic factors can, at least to some degree, offset problematic phonological assembly skills in RD readers. Identification of the neurobiological correlates of this type of top–down modulation in RD will allow us to address changing activation profiles as demands on core processes are systematically manipulated, in order to assess potential functionality in LH reading-related systems in RD.

Experiment 1 examines go/no-go naming responses for words while manipulating frequency, consistency, and imageability. Go/no-go naming (in which participants are required to name the token if it is a real English word but not if it is a pseudoword) was chosen because this overt naming paradigm strongly engages phonological processing, while also accentuating the influences of lexico-semantics (Frost et al., 2005).

We predict that our behavioral findings will follow those of previous studies: Consistency effects will be greater for RD relative to NI readers, and frequency and/or imageability will facilitate naming latencies and accuracy on difficult-to-decode inconsistent words for all participants, but with a much larger effect in RD. With respect to anticipated brain activation patterns, for NI readers, previous research indicates that stimuli that are easier to process should be associated with reduced blood oxygenation level-dependent (BOLD) signal at reading-related regions in the LH, reflecting increased processing efficiency within these regions (Frost et al., 2005; Katz et al., 2005; Sandak, Mencl, Frost, & Pugh, 2004; Sandak, Mencl, Frost, Rueckl, et al., 2004; Poldrack & Gabrieli, 2001). For RD readers in this age range, the predictions are less clear. As we stated earlier, adolescent RD, for whom altered reading circuitry has been established, may exhibit complete dysfunctionality in LH systems for reading. If so, we would anticipate that words that are easier to process will be associated only with modulated activation of the compensatory RH and frontal circuitry. However, if LH posterior regions are

less stable but not wholly dysfunctional, we would predict that words that are easier to process should reveal greater engagement of reading-related LH posterior regions for RD readers. That is, whereas inconsistent words in general should be associated with diminished LH posterior responses in RD, the facilitative influences of frequency and imageability might be associated with increased LH posterior responsiveness (and possibly with reduced reliance on RH posterior and bilateral anterior compensatory circuits).

Methods

Participants

Forty-four native English speakers (27 males, 17 females) ranging from 11.0 years in age to 19.0 years participated in exchange for payment: 24 NI and 20 RD. NI readers had averaged standard scores >100 (mean = 116) on a composite of the following three assessments: (1) Woodcock Johnson basic reading (mean = 115), (2) Woodcock Johnson spelling (mean = 124), and (3) TOWRE total word reading (mean = 108) measures. Twenty met our criteria for RD of averaged standard scores <90 on a composite of these reading tests [Woodcock Johnson basic reading (mean = 90), Woodcock Johnson spelling (mean = 87), and TOWRE total word reading (mean = 75) measures], and/or averaged standard score <100 with RD history. Groups did not differ on age (RD = 14.94; NI = 15.32, $t < 1$) or WASI performance IQ (RD = 105; NI = 109, $t < 1$). All reported normal or corrected-to-normal vision and no history of known neurological impairments. The experiment was conducted with the understanding and the written consent of each participant and all procedures were approved by the Yale University Institutional Review Board.

Stimuli and Design

Word stimuli represented the crossing of frequency (low, high), imageability (low, high), and spelling-to-sound consistency (consistent, inconsistent) with 40 words per condition, yielding a total of 320 "go" trials. Nonword trials were made up of 80 pseudowords matched on factors including length, bigram frequency, and initial phoneme, which served as "no-go" trials. Because we were unable to obtain naming latencies during functional scanning, behavioral data were collected in a separate session. In order to assess the behavioral profiles for each group, we asked participants to return to the lab on a day subsequent to functional magnetic resonance imaging (fMRI) scanning and perform the go/no-go task with the same materials in order to record behavioral data. Thirty (16 NI, 14 RD) subjects returned for the out-of-scanner session (fMRI analyses include the full cohort of 44).

Procedure

For the fMRI session, a go/no-go naming paradigm was employed in a block design. Each 20-sec experimental block consisted of five 4-sec trials in which participants were presented with a letter string for 1 sec that either made a real English word or not and were instructed to name it aloud if it was a real English word ("go trial") and to do nothing if it was not a word ("no-go trial"). The proportion of go and no-go trials was equivalent across conditions. During baseline blocks, participants passively viewed displays of intermixed hash marks and asterisks.

fMRI Image Acquisition and Analysis

Functional imaging runs consisted of eight 20-sec experimental blocks (one for each stimulus condition) of four word trials and one pseudoword trial, and five 20-sec baseline blocks. A total of 1300 full-brain functional images were acquired across 10 imaging runs; 100 images for each experimental condition and 500 images in the baseline condition. Each subject received the same pseudorandom order of runs. The order of activation block types was counterbalanced across runs.

Functional imaging was performed on GE Signa 1.5-Tesla and Siemens Sonata 1.5-Tesla MR systems. Participants' heads were immobilized within a circularly polarized head coil using a neck support, foam wedges, and a restraining band drawn tightly around the forehead. Prior to functional imaging, 20 axial–oblique anatomic images (TE = 11 msec; TR = 500 msec; FOV = 200 mm; 6 mm slice thickness, no gap; $256 \times 256 \times 2$ NEX) were prescribed parallel to the intercommissural line based on sagittal localizer images (TE = 11; TR = 600 msec; FOV = 240 mm; 23 slices, 5 mm slice thickness, no gap; $256 \times 256 \times 1$ NEX). Activation images were collected using single-shot, gradient-echo echo-planar acquisitions (flip angle = 80°; TE = 50 msec; TR = 2000 msec; FOV = 200 mm; 6 mm slice thickness, no gap; $64 \times 64 \times 1$ NEX) at the same 20 slice locations used for anatomic images.

Functional images were first sinc-interpolated to correct for slice acquisition time, corrected for motion (Friston et al., 1995), and spatially smoothed with a Gaussian filter of size 3.125 mm full width at half maximum. For each subject, an affine transformation to the standardized space defined by the Montreal Neurological Institute (MNI) was obtained using Bio-ImageSuite (Papademetris, Jackowski, Schultz, Staib, & Duncan, 2003; www.bioimagesuite.org), mapping between the subject-space T1 anatomic and the MNI space "Colin" brain (available at www.bic.mni.mcgill.ca). Prior to across-subjects analysis, this transformation was applied to the single-subject activation maps, with trilinear interpolation, into 2-mm isotropic MNI space. For each subject and voxel, linear regression was used to compare

the mean signal during each experimental condition to the baseline condition, and these differences were converted to standardized activation values by dividing them by the square root of the error mean square for the model. Across subjects, these values were entered into a mixed-model or repeated measures analysis of variance (ANOVA; Holmes & Friston, 1998; Woods, 1996; Kirk, 1982) with planned comparisons for main effects, frequency, imageability, and consistency, and their interactions, conducted on a voxelwise basis.

Region-of-interest (ROI) coordinates were defined by peak activation sites within the Group by Stimulus difficulty interaction analysis (see below). From this analysis, we identified seven LH regions that (1) have been previously implicated in reading (cf., Price, 2000; Pugh, Mencl, Jenner, et al., 2000; Pugh, Mencl, Shaywitz, et al., 2000; Posner, Abdullaev, McCandliss, & Sereno, 1999) and (2) showed a significant Group by Stimulus difficulty interaction. Table 1 presents the MNI coordinates and significance levels for peak activation as well as the volume for the following regions: LH fusiform/occipito-temporal (OT), middle temporal gyrus (MTG), thalamus (THAL), superior temporal gyrus (STG), insula (INS), inferior frontal gyrus (IFG), and supramarginal gyrus (SMG). RH regions were examined based on previous studies implicating RD differences and the presence of a Group by Stimulus difficulty interaction (Pugh, Mencl, Jenner, et al., 2000; Pugh, Mencl, Shaywitz, et al., 2000).

Results

Behavioral Analysis

Separate $2 \times 2 \times 2 \times 2$ mixed-factors ANOVA were conducted on latencies to correct responses and on errors. Frequency (low/high), imageability (low/high), and consistency (consistent/inconsistent) served as within-subjects factors and reader group (NI/RD) served as a between-subjects factor. Overall, naming latencies were slower [$F(1, 28) = 13.05, p < .01$] and less accurate [$F(1, 28) = 8.15, p < .01$] for RD relative to NI participants.

Table 1. Regions of Interest (ROIs) Showing Reader Group × Stimulus Difficulty Interaction

Region	x	y	z	p	Volume (mm^3)
L. Occipito-temporal/fusiform	−48	−46	−18	.0031	184
L. Middle temporal	−44	−48	−6	.0004	304
L. Thalamus	−18	−30	−2	.0056	248
L. Superior temporal	−44	−44	10	.0002	384
L. Insula	−38	0	4	.0003	408
L. Inferior frontal	−42	6	24	.0007	664
L. Supramarginal	−50	−34	32	.0001	320

The Group by Imageability interaction was marginal for naming latency [$F(1, 28) = 3.46, p = .07$] and significant for accuracy [$F(1, 28) = 6.91, p < .05$], revealing heightened effects of imageability (faster and more accurate responses for high imageable relative to low imageable words) in RD readers. There were also reliable Group by Frequency interactions on both naming latency [$F(1, 28) = 12.27, p < .01$] and accuracy [$F(1, 28) = 6.44, p < .05$], indicating heightened effects of frequency (faster and more accurate responses for high-frequency relative to low-frequency words) in RD. These interactions were further qualified by a three-way interaction of Group × Frequency × Imageability on accuracy [$F(1, 28) = 4.81, p < .05$], such that frequency effects were greater for low-imageable relative to high-imageable words and this difference was greater for RD readers (3%) than for NI readers (4%). A reliable four-way interaction of Group × Frequency × Imageability × Consistency was also observed [$F(1, 28) = 7.01, p < .05$, driven by a maximum drop in proportion correct to 0.86 for RD on low-frequency, low-imageable, inconsistent words, as anticipated from previous research].

Given that both frequency and imageability had a heightened facilitative influence on naming inconsistent words for RD participants, a targeted stimulus difficulty analysis of extreme conditions contrasting low-frequency/low-imageable/inconsistent words (henceforth LF–LI–INC) with high-frequency/high-imageable/inconsistent words (HF–HI–INC) was performed in order to compare the reader groups on difficult-to-decode words that differ systematically with regard to familiarity and semantic features. This extreme contrast allows us to directly examine purported top–down effects on problematic decoding in RD with maximal power. A group interaction was obtained for both accuracy [$F(1, 28) = 5.87, p < .025$] and latencies [$F(1, 28) = 8.73, p < .01$] (shown in Figure 1A and B, respectively). As predicted, whereas both groups were faster and more accurate on HF–HI–INC than on LF–LI–INC words, this advantage was amplified for RD participants.

fMRI Analysis

Naming engaged a broad bi-hemispheric circuitry in general, and overall, activation during the go/no-go task (collapsed across stimulus type) was reliably higher in a large number of regions for NI, relative to RD, participants, as seen in previous studies (see Figure 2A). Of more acute interest in the current experiment, however, is how group differences in activation were qualified by stimulus characteristics.

Stimulus Difficulty

The behavioral data indicate that RD readers benefited from high frequency and high imageability when reading difficult-to-decode inconsistent words. Therefore, to
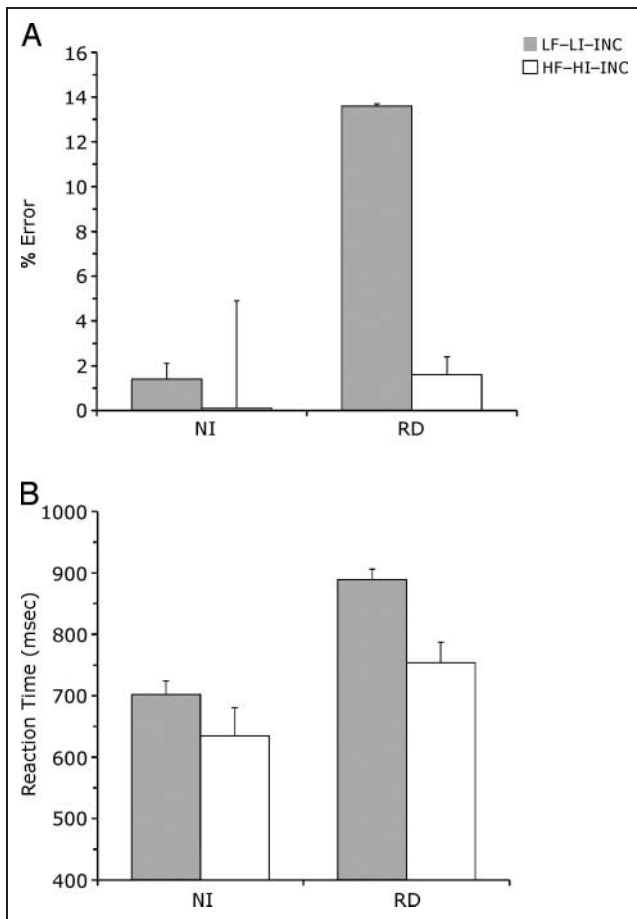
Figure 1. Mean percent error (A) and reaction time (B) for NI and RD readers on the contrast of low-frequency, low-imageable, inconsistent (LF–LI–INC) versus high-frequency, high-imageable, inconsistent (HF–HI–INC) words.

examine the maximum benefit from our stimulus manipulations, we conducted an analysis of extreme conditions (LF–LI–INC vs. HF–HI–INC words) that paralleled the analysis conducted on the behavioral data. Regions showing this Group by Stimulus difficulty interaction are shown in Figure 2B. For NI, the easier HF–HI–INC items were associated with relatively reduced activation at virtually all of these regions. For RD, by contrast, these
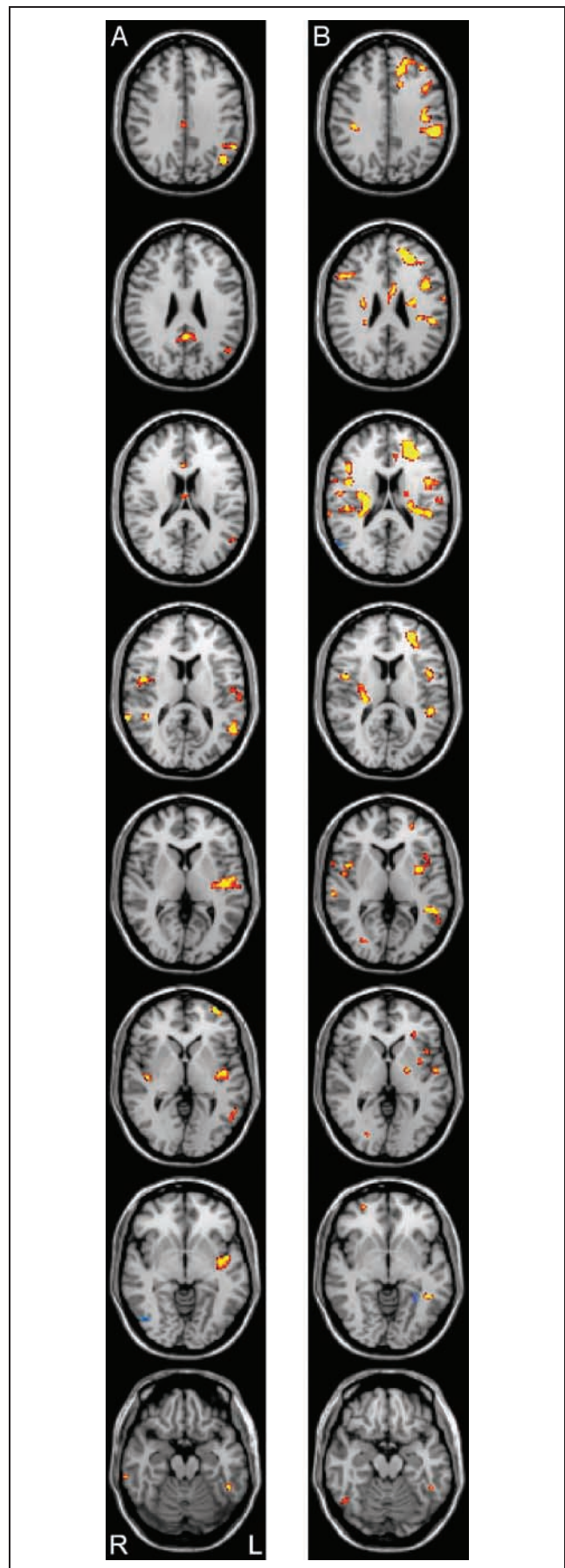
Figure 2. Omnibus group differences indicate regions where activation for NI is greater than RD (yellow/red) or where RD is greater than NI (blue/purple) (A). Group difference on the contrast of low-frequency, low-imageable, inconsistent (LF–LI–INC) versus high-frequency, high-imageable, inconsistent (HF–HI–INC) words (B). In yellow/red are those regions where NI show decreases in activation for HF–HI–INC words relative to LF–LI–INC words and RD showed increases. Images are presented at a univariate threshold of $p < .01$, corrected for mapwise false discovery rate (FDR; Genovese, Lazar, & Nichols, 2002). Images from top to bottom correspond to the following position along the z-axis in MNI space: +34, +26, +18, +12, +4, +0, −6, and −20, respectively, with the LH on the right side of the images.

easier words were associated primarily with heightened activation at key, LH reading-related regions.

Regions of Interest

To more fully investigate stimulus-qualified reader group interactions across the set of a priori defined regions previously implicated as reading-relevant, we conducted ROI analyses. We isolated those clusters of voxels in these regions that were associated with reliable a Group × Stimulus Difficulty (LF–LI–INC vs. HF–HI–INC) interaction (see Figure 2B and Table 1). A striking pattern was observed at several regions, particularly temporo-parietal areas, including STG and SMG (see Figure 3) For NI, the easier, HF–HI–INC words were associated with reduced activation at every ROI, whereas for RD, the opposite pattern was observed (increased activation for easier tokens). Thus, although activation of key LH regions was low for inconsistent words in general for RD readers, when these difficult-to-decode words were of both high frequency and high imageability, activation levels increased modestly.

Discussion

The findings from Experiment 1 suggest latent functionality in LH regions in RD including the IFG and the STG: RD readers increased engagement of the LH reading systems for easier stimuli. However, two important points should be noted with respect to whether activation in these reading systems may be "normalized." First, despite reliable increases in major reading-related areas for HF–HI–INC stimuli relative to LF–LI–INC stimuli (see Figure 2B), activation levels were still relatively weak for RD readers compared with typical levels for NI readers. Second, the commonly seen RD compensatory response in RH posterior and bilateral IFG was still evident even on easy tokens (higher activation for RD readers), implying limits on normalization of response for this type of manipulation.

EXPERIMENT 2

In order to further examine the limits on normalization of function, we employed a second experimental manipulation—repetition. The frequency-related activation effects in Experiment 1 indirectly suggest that the number of exposures is a critical variable in increased LH responses in RD. In Experiment 2, we conduct a direct test of this notion by manipulating the number of exposures to a given token in the short term with a repetition paradigm (Katz et al., 2005; Poldrack & Gabrieli, 2001). In addition to being one of the strongest behavioral manipulations, repetition allows us to more precisely examine learning-dependent brain activation changes because we directly control short-term experience in both groups by manipulating how frequently a token is en-

countered. By employing an animacy judgment (living/nonliving) using a button press, we also measure latency and accuracy, which allows for a more precise comparison of coordinated behavioral and neurobiological changes in NI and RD groups than in Experiment 1. The goal of this repetition manipulation was to bring RD readers to a point of overlearning for repeated tokens (Adams, 1994) in order to examine whether LH systems are robustly engaged with higher learning levels.

Methods

Participants

Thirty native English speakers (17 males, 13 females) ranging from 9 years in age to 20 years (mean = 13 years) participated in exchange for payment: 16 NI and 14 RD. Of the 16 NI readers, 14 averaged standard scores >100 on the composite of the reading tests described in Experiment 1 [overall mean = 115, Woodcock Johnson basic reading (mean = 113), Woodcock Johnson spelling (mean = 119), and TOWRE total word reading (mean = 112) measures].[1] Fourteen participants met our criteria for RD with averaged standard scores <90 on a composite of these reading tests, and/or averaged standard score <100 with RD history [overall mean = 86, Woodcock Johnson basic reading (mean = 88), Woodcock Johnson spelling (mean = 88), and TOWRE total word reading (mean = 80)]. Groups did not differ on age (RD = 12.89; NI = 13.84, $t < 1.5$) or WASI performance IQ (RD = 108; NI = 111, $t < 1$). All reported normal or corrected-to-normal vision and no history of known neurological impairments. The experiment was conducted with the understanding and the written consent of each participant and all procedures were approved by the Yale University Institutional Review Board.

Stimuli and Design

Two hundred eight mid-frequency nouns were selected for the study. All words were four or five letters in length and all had regular spelling-to-sound mappings. Sixty-five percent of the items were "nonliving" and 35% were "living." Repeated and novel conditions were matched for mean frequency, length in letters, and proportion of living/nonliving trials within each imaging run. Two lists were created for counterbalancing purposes such that a subset of the novel words in the first list served as repeated words in the second list and vice versa. Participants were randomly assigned to one of the two lists.

Procedure

Eight functional imaging runs in an event-related animacy judgment (living/nonliving) paradigm employed (a) interleaved acquisition to increase the effective
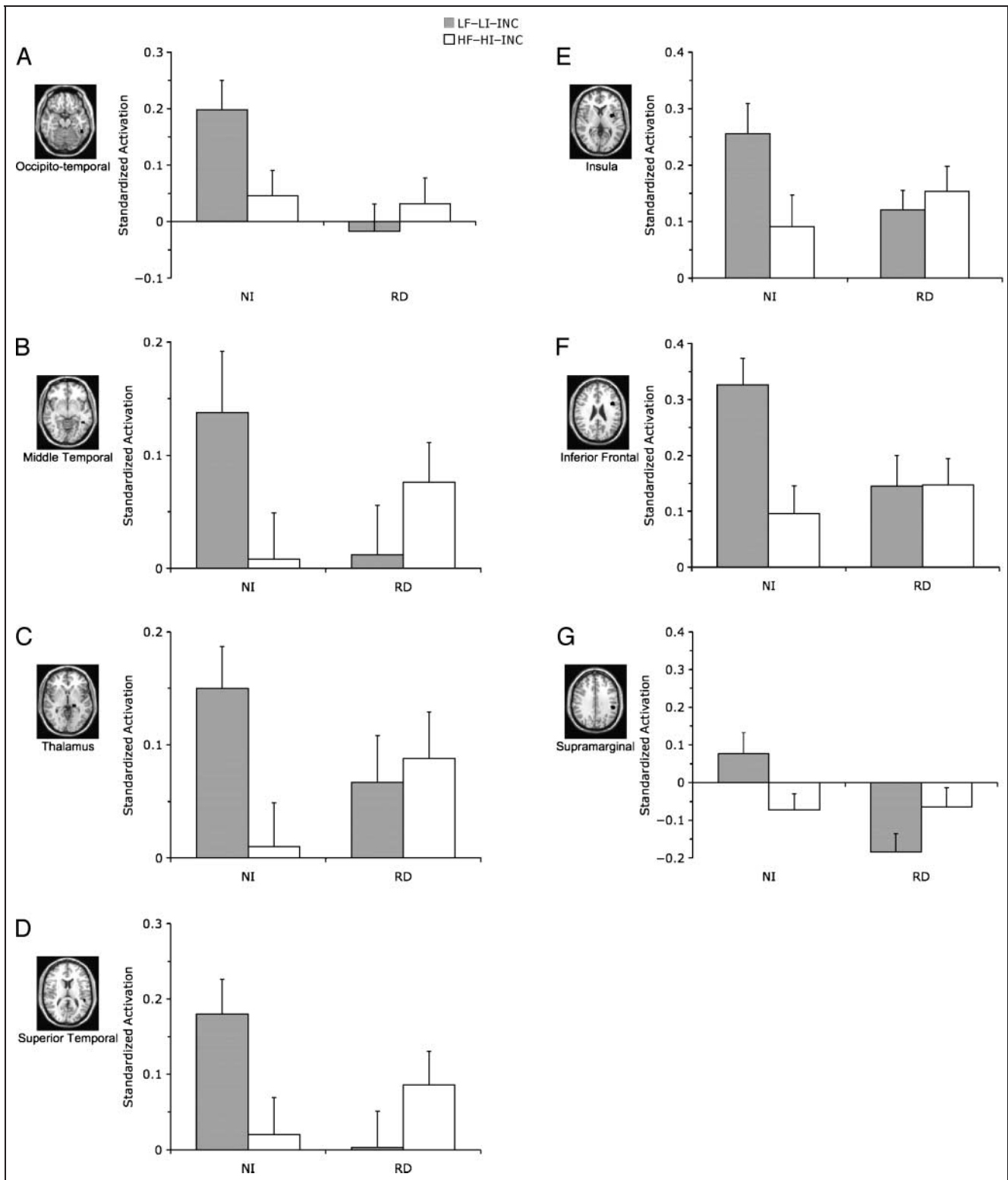
Figure 3. Standardized activation values for the reader group by stimulus difficulty contrasts of LF–LI–INC versus HF–HI–INC words in the seven LH ROIs: OT/fusiform (A), MTG (B), thalamus (C), STG (D), insula (E), IFG (F), and SMG (G).

sampling rate of the hemodynamic response (Josephs, Turner, & Friston, 1997), (b) multiple randomized or "jittered" trial durations (4–7 sec) (Miezin, Maccotta, Ollinger, Petersen, & Buckner, 2000), and (c) and six "null" trials

(Friston et al., 1995) to improve our estimate of baseline activation. Each 6:18 minute run consisted of 56 trials in which six words were presented six times in a pseudorandom fashion with 20 intermixed tokens serving as

unrepeated control words. All participants completed at least six runs, for a minimum of 216 repetition trials and 120 novel trials across runs. On each trial, a word appeared in the center of screen for 2500 msec and participants were instructed to indicate, as quickly as possible, whether the word came from the category of living or nonliving objects via a right-hand button press. Participants pressed buttons on a response pad with the middle finger for "living" responses and the index finger for "nonliving" responses. In-scanner behavioral measures (i.e., reaction time and accuracy) were collected for all subjects using PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993). Response timing started at the onset of the stimulus presentation and continued until the end of the trial. Participants received 16 practice trials before functional scanning began in order to familiarize themselves with the task and setup.

fMRI Image Acquisition and Analysis

Image acquisition and preprocessing were conducted as described in Experiment 1 except that high-resolution anatomical images were obtained for 3-D reconstruction (sagittal MPRAGE acquisition, FA = $45°$; TE = 4.66 msec; TR = 2000 msec; FOV = 25.6×25.6 cm; 1 mm slice thickness, no gap; $256 \times 256 \times 1$ NEX; 28 slices total). For each subject, a nonlinear transformation was then obtained using BioImageSuite (Papademetris et al., 2003; www.bioimagesuite.org), mapping between the subject-space high-resolution anatomic and the standard brain space defined by the MNI "Colin" brain. Prior to across-subjects analysis, this transformation was applied to the single-subject activation maps (described below), with trilinear interpolation, into 2 mm isotropic MNI space.

For single-subject event-related analysis, a regression-based method was utilized, allowing for direct estimation of the hemodynamic response for each trial type, at each voxel separately, without prior specification of a reference function (Miezin et al., 2000). Parameters from this regression model were then used to uniquely estimate the mean response for each condition from -3 to $+15$ sec relative to stimulus onset. Subject activation maps were then created for each condition using the regression estimates to calculate the mean difference in activity for an activation period (3–8 sec post trial onset) relative to a baseline period (0–3 sec prior to trial onset). Linear contrasts for effects of interest, including the evoked response of each trial type, simple subtractions among trial types, main effects, and interactions, were applied to these regression estimates to obtain contrast images for each subject. Across subjects, each voxel in these contrast images was tested versus zero with an F test, implementing a mixed-model or repeated measures ANOVA (Holmes & Friston, 1998; Woods, 1996; Kirk, 1982).

ROI coordinates were defined by peak activation sites within a Group by Linear trend analysis performed to isolate voxels where there was a linear trend across the six presentations that differed by reader group. From this analysis, we identified sites in the seven LH regions that have been previously implicated in reading (cf., Price, 2000; Pugh, Mencl, Jenner, et al., 2000; Pugh, Mencl, Shaywitz, et al., 2000; Posner et al., 1999) and were examined in Experiment 1. For the thalamus, STG, and SMG, including voxels that passed the threshold of $p < .01$, FDR corrected, yielded a volume of less than 100 mm^3; thus we adjusted the threshold to include voxels that passed $p < .05$, FDR, to obtain a more stable descriptor of activation of these regions. MNI coordinates and significance levels for peak activation, as well as the volume for the seven regions, are presented in Table 2. RH regions were examined based on previous studies implicating RD differences and the presence of a Group by Linear trend interaction (Pugh, Mencl, Jenner, et al., 2000; Pugh, Mencl, Shaywitz, et al., 2000).

Results

Behavioral Analysis

Separate ANOVAs were performed on both accuracy and latencies to correct responses. Reading group was the sole between-subjects factor and repetition was the within-subject factor. Due to the relatively small number of stimuli in each exposure condition, we collapsed the six exposures into three periods: early (first and second exposures), middle (third and fourth exposures), and late (fifth and sixth exposures). This three-level coding is employed for both behavioral and fMRI ROI analyses. Accuracy analyses (Figure 4A) revealed a main effect of reader group [$F(1, 28) = 4.61, p < .05$], a marginal effect of repetition ($p < .10$), and no interaction. Latency analyses shown in Figure 4B revealed main effects of reading group [$F(1, 28) = 7.01, p < .05$] and repetition [$F(2, 56) = 36.6, p < .001$]. The Reading group by Repetition interaction was not significant ($F < 1$). These data indicate that effects of repetition were facilitative as expected and were of similar magnitude for both NI and

Table 2. Regions of Interest (ROIs) Showing Reader Group × Linear Trend Interaction

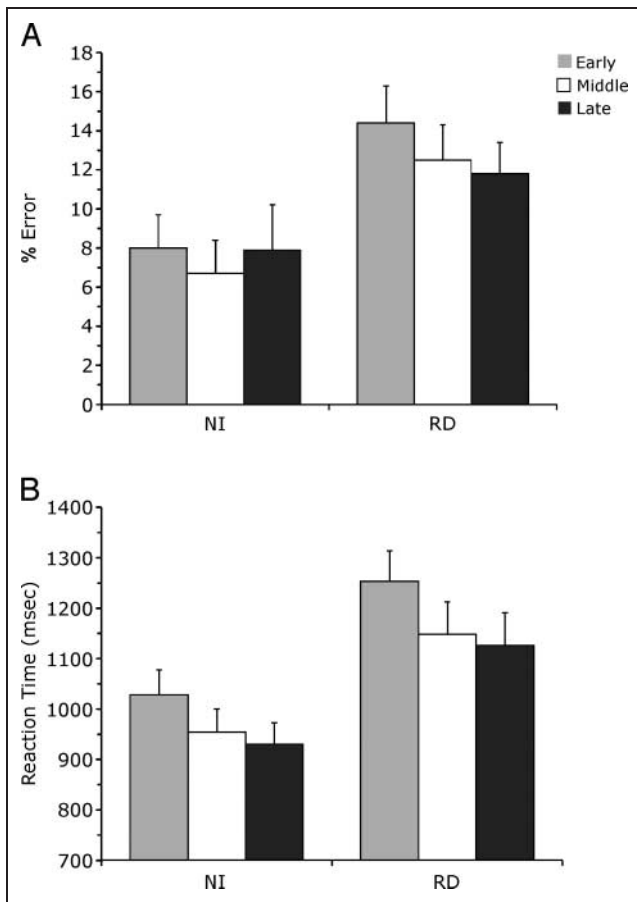| Region	x	y	z	p	Volume (mm^3)
L. Occipito-temporal/fusiform	−46	−44	−20	.0002	112
L. Middle temporal	−58	−34	−6	.0002	880
Thalamus	−26	−32	8	.0024	216
L. Superior temporal	−66	−14	10	.0002	432
L. Insula	−44	2	−2	.0011	248
L. Inferior frontal	−56	18	24	.0002	216
L. Supramarginal	−34	−64	56	.0042	560

Figure 4. Mean percent error (A) and reaction time (B) for NI and RD readers for early, middle, and late exposures to words.

RD groups. A general advantage for NI was observed but performance in RD was, nonetheless, quite good on this task.

fMRI Analysis

The main effect analysis for group differences revealed the commonly seen underactivation of wide numbers of regions in RD relative to NI readers (see Figure 5A). The Reading group by Linear repetition comparison revealed reading group differences in the direction of the repetition effect in a number of regions, including the LH OT, MTG, thalamus, STG, insula, IFG, and SMG (see Figure 5B).

Figure 5. The group difference on unrepeated words (A) shows regions where activation for NI is greater than RD (yellow/red) or where RD is greater than NI (blue/purple) ($p < .001$, FDR corrected). The Group by Linear repetition interaction shows regions where NI decreased activation across the six exposures to a word (yellow/red) and RD increased activation across exposures ($p < .01$, FDR corrected). Images from top to bottom correspond to the following position along the z-axis in MNI space: $+46$, $+36$, $+26$, $+14$, $+8$, -4, -12, -12, and -20, respectively, with the LH on the right side of the images.
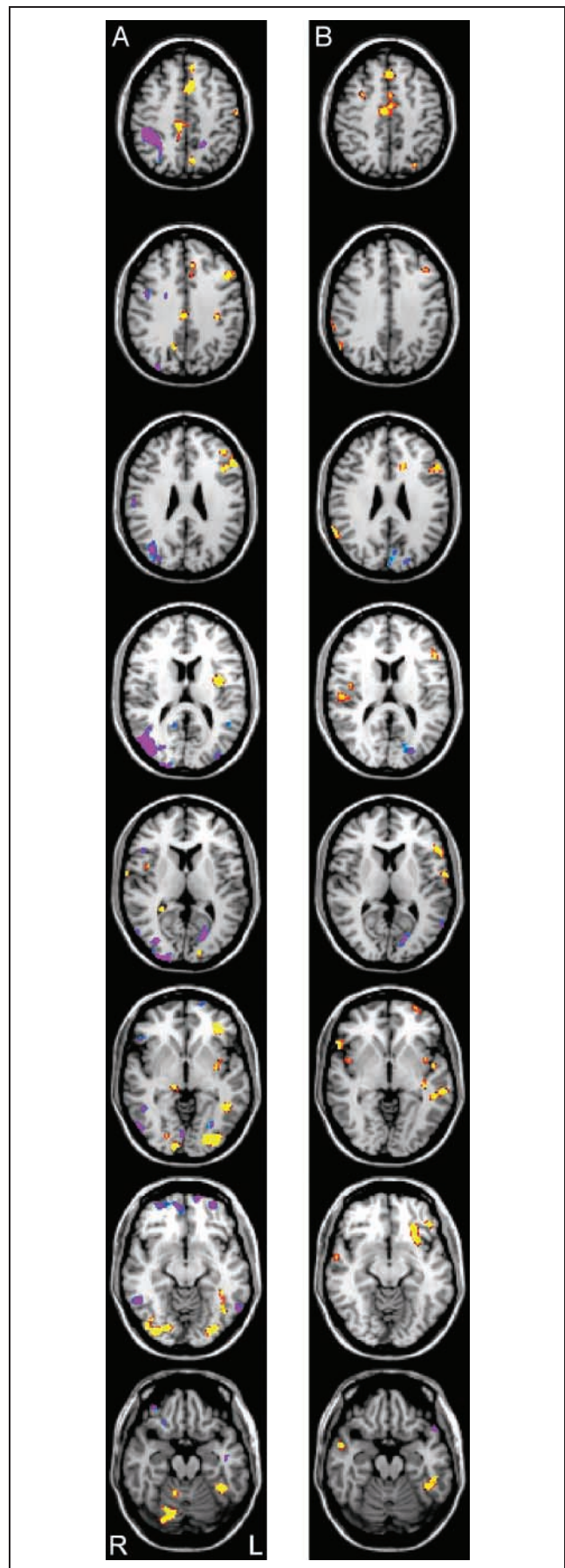
These interactions, shown in more detail in the following ROI Analyses section, indicate differential effects of repetition on activation for the two groups in these regions. On early trials, increased activation for NI relative to RD is seen at multiple regions, including the LH OT, STG, insula, IFG, MFG, cerebellum, and thalamus, along with RH sites including the OT, IFG, and MTG. On late trials, the differences are more circumscribed and limited to the OT, extrastriate, and LH insula. Heightened RH activation for RD is apparent at OT and MTG sites, and LH prefrontal increases in RD are seen. Critically, activation differences in several reading-related sites such as the STG and the IFG are no longer apparent on late trials.

Regions of Interest

Figure 6 presents the activation levels for each of the seven ROIs described earlier at each exposure period for the two reader groups. The pattern in most of these regions is such that, for NI readers, reduction in activation from early to middle to late was seen, whereas for RD the opposite pattern was observed. Indeed, for the RD readers, each region exhibits significant activation on late trials (fifth and sixth exposures combined). Note that in the SMG, STG, and insula, a shift from deactivation to activation in RD is seen by the middle trials, whereas in the OT the shift occurs by the late trials. Moreover, in most regions, we observed activation decreases for NI but with some sustained activity even on late trials, whereas the IFG shows deactivation on late trials. This is consistent with our previous study (Katz et al., 2005) suggesting that premotor activation is eliminated in silent reading tasks with multiple repetitions (and increased efficiency) for skilled readers. Of note is the activation pattern in the MTG that is uniformly low for NI but increases dramatically across repetitions in RD.

Discussion

The findings from Experiment 2 demonstrate that repetition has a similar, facilitative effect on reaction time and accuracy for both NI and RD readers, but has an opposite effect on BOLD activation. In crucial LH regions, RD readers show the often-reported deactivation early but show reliable increases between three and six exposures, whereas NI readers demonstrated continued reduction in activation with increased exposure. The current findings reinforce our conclusion from Experiment 1 that the LH systems in RD are poorly tuned but can respond when processing words is made easier. Of note, although there are still some residual differences in LH regions such as the insula, and slightly elevated RH response in posterior ventral areas for RD even with multiple exposures, the overall activation in several LH regions appears to be robust and not reliably different as performance improves.

The similar pattern of findings in the two experiments suggests that these learning-related increases in RD generalize to both overt naming tasks (Experiment 1) and silent reading tasks (Experiment 2). Note that the regions of maximum activation tend to differ somewhat within the broadly defined ROIs but this is not surprising given the differing demands of naming and silent lexical access. In order to identify those voxels which showed reliable group interactions with the stimulus manipulations in both experiments, presumably task invariant reading sites, we conducted an intersect analysis (Hadjikhani & Roland, 1998), which identified those voxels for NI that reduced activation across stimulus difficulty (Experiment 1), repetition (Experiment 2), and increased activation across stimulus difficulty and repetition for RD. As seen in Figure 7, overlapping sites are found at the IFG, insula, STG, and MTG that show this pattern in each experiment. Thus, despite rather varied response demands, a core set of areas, mainly in LH cortex, showed opposite activation changes in NI and RD as a function of stimulus difficulty.

GENERAL DISCUSSION

The current findings suggest a degree of latent functionality in LH neurocircuitry for reading in RD readers. The behavioral results from Experiment 1 (on both naming latencies and accuracy) demonstrate that factors such as frequency and imageability enhance performance for RD and NI readers on difficult-to-decode (inconsistent) words. Similarly, repetition of tokens in Experiment 2 facilitated processing on repeated tokens for both groups of readers. This work extends recent studies examining stimulus difficulty and effort effects on activation differences (Hoeft et al., 2007; Cao et al., 2006), by demonstrating significant increases in neural activity in LH systems for adolescent RD as a consequence of learning and experience.

More striking was the differential pattern of brain activation across both experiments. In Experiment 1, the often-reported group differences in activation (lower BOLD signal across the LH reading-related circuitry in RD relative to NI) were qualified by stimulus difficulty. Specifically, on the most difficult words (LF–LI–INC), RD readers demonstrated reduced activation relative to NI across almost all reading-relevant zones; indeed, for these items, no reliable activation was seen in key temporo-parietal sites including the STG and the SMG. For easier stimulus types (e.g., HF–HI–INC words), RD readers demonstrated *increased* activation relative to the hardest stimuli at these same LH regions, suggesting that these cortical networks are poorly trained in RD but not wholly unavailable during reading performance (Pugh, Mencl, Shaywitz, et al., 2000). Although NI readers showed stable activation of these regions for all stimulus types, when compared to RD readers, the NI group showed the opposite pattern of modulation for
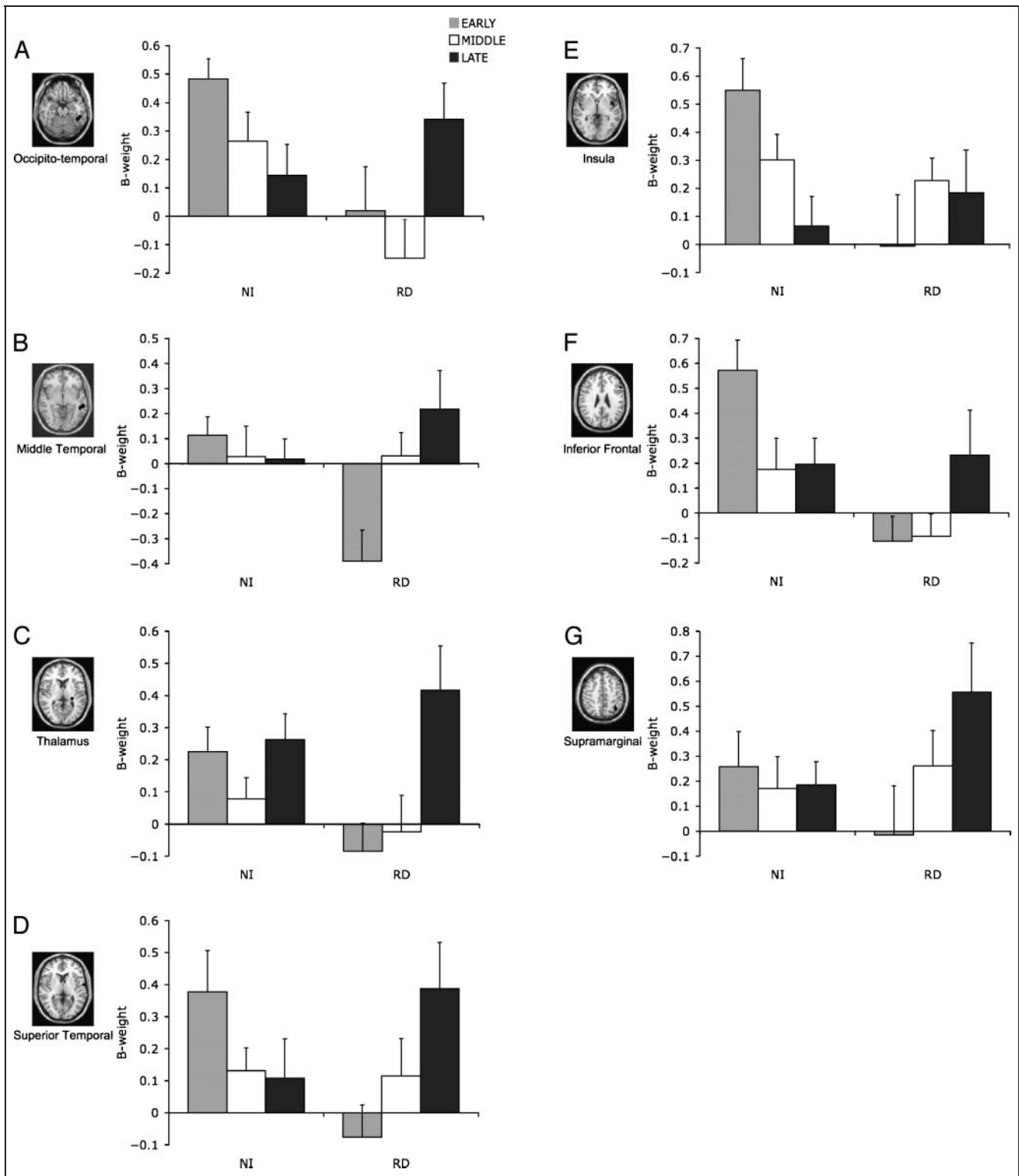
Figure 6. Activation values for the Reader group by Linear repetition interaction in the seven LH ROIs: OT/fusiform (A), MTG (B), thalamus (C), STG (D), insula (E), IFG (F), and SMG (G).

easier-to-process words with relatively *reduced* activation apparent across the LH reading circuit—presumably reflecting increased processing efficiency (Katz et al., 2005; Sandak, Mencl, Frost, & Pugh, 2004; Sandak, Mencl, Frost, Rueckl, et al., 2004; Poldrack & Gabrieli, 2001).

Although the data indicate that RD readers increased engagement of the LH reading systems for easier stimuli, Experiment 1 was not definitive regarding whether activation in these reading systems may be "normalized." Despite modest increases in major reading-related areas
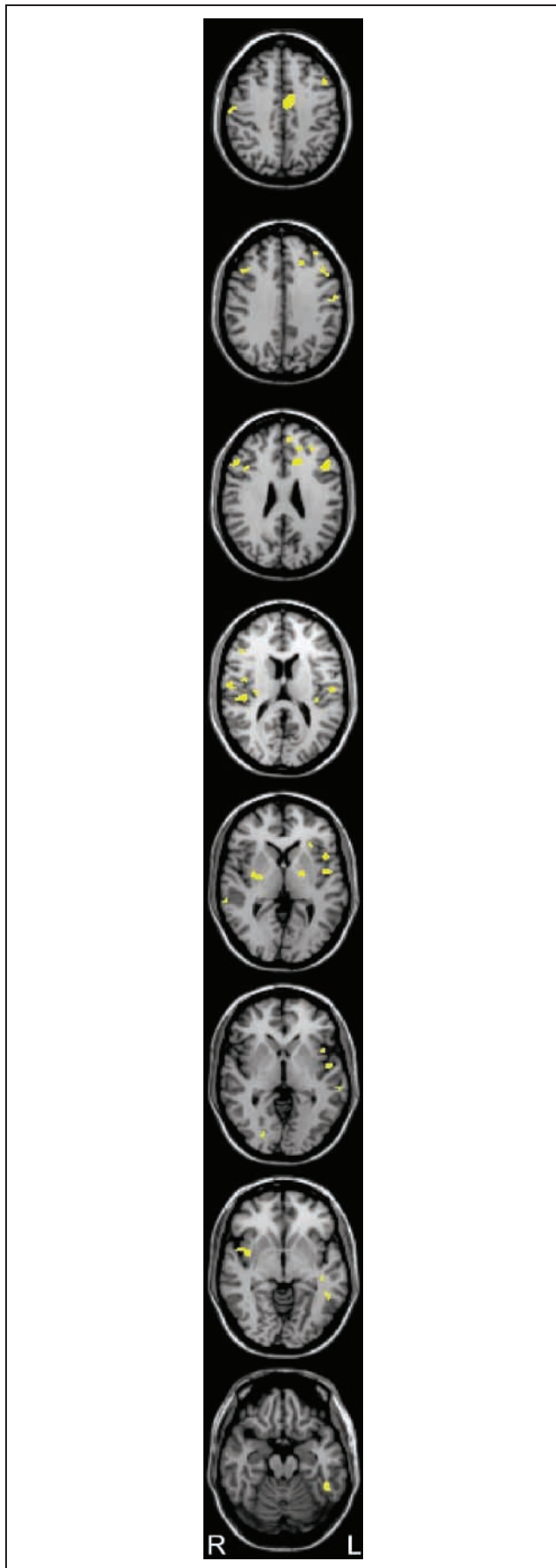
for HF–HI–INC stimuli (see Figure 2B), activation levels were still relatively weak for RD readers compared with typical levels for NI readers. Moreover, the commonly seen RD compensatory response in the RH posterior and bilateral IFG was evident even on easy tokens, implying limits on normalization of response.

Experiment 2 provided clear evidence for a shift from deactivation to robust activation of most key LH regions with repetition in RD readers. We can conclude from these experiments that many important reading-related regions are capable of engaging in print processing for stimuli that are made easier to process either through repetition-induced learning or increased top–down support from imageability or frequency manipulations.

Given that phonologically sensitive subregions of the LH (i.e., IFG, SMG, STG) are generally found to be underengaged in RD (including in the current study for more difficult stimuli), the increased responsiveness in RD for high-frequency/high-imageable inconsistent tokens relative to low-frequency/low-imageable inconsistent tokens in naming at each of the key phonologically tuned areas (Pugh et al., 1996) may reflect increased communication or resonance with semantically tuned networks; that is, stabilization of a poorly tuned phonological coding system via support from nonphonological systems. Experiment 2 is empirically straightforward with regard to increased activation in phonologically tuned regions, but whether this reflects semantic reinforcement (given that the task involves animacy judgment), orthographic/phonological reinforcement, or both of these things is not fully answerable. One possible answer is suggested by the pattern of activation in the MTG (see Figure 6), a region strongly associated with lexical–semantic processing (Frost et al., 2005; Sandak, Mencl, Frost, Rueckl, et al., 2004; Price, 2000; Pugh, Mencl, Jenner, et al., 2000). In the MTG, NI readers showed minimal and unchanging activation whereas RD demonstrated initial deactivation and then robust activation by middle to late trials. This might reflect differential sensitivity to semantic support in RD readers. In general though, we see that relatively poorly tuned phonologically sensitive regions are positively affected by ameliorative factors for RD in both Experiments 1 and 2.

The current results suggest a neurobiological learning curve wherein NI and RD readers start at very different points on an inverted U-shaped relationship between learning and neural activation. Neuroimaging studies of perceptual and motor skill learning in nonimpaired populations have demonstrated that initial skill acquisition (unskilled performance) is associated with increased

Figure 7. Intersect analysis showing voxels that showed a Reader group by Stimulus difficulty effect and a Reader group by Repetition effect ($p < .05$, FDR corrected in each; conjoint threshold of $p < .0025$). Images from top to bottom correspond to the following position along the z-axis in MNI space: +42, +34, +26, +14, +2, −2, −6, and −20, respectively, with the LH on the right side of the images.

activation in task-specific cortical areas, whereas continued practice of an acquired skill tends to be associated with task-specific decreases in activation in the same cortical regions. (e.g., Katz et al., 2005; Poldrack & Gabrieli, 2001). Other studies of skilled reading have also shown different patterns of learning-related changes in brain activation as a function of item familiarity. For example, Henson, Price, Rugg, Turner, and Friston (2002) found that repetition of familiar real words was associated with decreases in cortical activation, whereas repetition of (initially unfamiliar) pseudowords was associated with increased activation in the same regions. Thus, this *learning-curve* hypothesis suggests that whether learning is associated with increased or decreased activation depends upon the degree to which processing is overlearned and automatic (how far along the learning curve it is). With regard to developmental trajectories, our cross-sectional studies have suggested that beginning readers show low activation of LH posterior regions, particularly the occipito-temporal region, but increase with experience (Shaywitz et al., 2002). However, once these systems are in place, increased routinization for stimuli (such as the repetition effect in the current Experiment 2) results in drops in activation (thus, a nonmonotonic relation between activation and experience in nonimpaired readers).

In the current study, RD readers appear to start at a very low point on this curve but repetition and learning result in increases. This does not necessarily imply a simple developmental delay account of specific RD because these regions are clearly less than ideally organized even after many years of reading experience in our adolescent RD readers, and this suggests some degree of compromise in the neural circuitry. But the robust activation seen after multiple exposures in Experiment 2 does lend itself to speculation that these systems are trainable (indeed, it will be critical in future studies to push these systems further to test limits on learning effects).

One question that the data from Experiment 2 raise is why the initial (first exposure) activation response in LH is so low given that RD readers have certainly seen the words used in this experiment thousands of times. The most straightforward hypothesis, and one that would point to a very specific learning problem in RD, is that these readers fail to consolidate the learning experience into longer-term neural changes in processing and organization. Thus, the system might be available for processing but might fail to demonstrate savings with longer-term modulation of connections. If this turns out to be the case (for all or some subtypes of RD), then this would shift focus away from exclusive focus on simple mapping deficits toward more systematic investigation of the mechanisms of explicit or implicit learning. There is, nonetheless, some indication of neural consolidation in Experiment 1, where both frequency and imageability (a semantic variable) were associated with a heightened

LH response even without local repetition. However, the increases for HF–HI–INC tokens in that experiment were modest. In any event, the mechanisms underlying an apparent consolidation deficit in RD will require us to explore a new line of dynamic learning paradigms to measure long-term learning under varied training conditions (e.g., Sandak, Mencl, Frost, & Pugh, 2004; Sandak, Mencl, Frost, Rueckl, et al., 2004).

The current results, along with our previous work that has shown nonuniformity in hemodynamic effects of learning as a function of stimulus type (Sandak, Mencl, Frost, Rueckl, et al., 2004), have some important methodological and design implications as well. These findings reveal the importance of controlling stimulus factors in order to derive a more precise understanding of brain–behavior relations in RD. Although most published reports show lower LH posterior activation in RD, the results of these experiments indicate that the extent of this difference is dependent on stimulus difficulty. From a design consideration, we argue that dynamic designs, which parametrically examine the ways in which learning modulates relative activation across distributed systems, will allow for the development of a more detailed theory of the neurobiological mechanisms of reading and provide a framework for examining systems-level differences in RD. Indeed, it seems plausible that in searching for biomarkers that are diagnostic in this condition, response patterning to tasks (measuring dynamic changes with learning) will prove more substantive than static group differences.

The current results provide some constraint on the sort of hypotheses we entertain regarding neurobiological mechanisms in RD. Whatever the biologic mechanism (or mechanisms) that engender risk for RD, this mechanism must be of the sort that results in a neurocircuitry that is relatively disrupted in general, but is, nonetheless, not so fundamentally compromised that a more typical reading response cannot be induced. Indeed, recent intervention studies with at-risk or RD children indicate increased engagement of all these LH areas following intensive remediation (Shaywitz et al., 2004; Temple et al., 2003; Simos et al., 2002).

Various biological accounts have been proposed to explain this LH dysfunction, including a suggestion of a higher numbers of cortical dysplasias or ectopias (Galaburda, 1992), reduced myelination in white matter tracts (Klingberg et al., 2000) connecting anterior and posterior language zones, or abnormalities in gray matter development (Miller, Sanchez, & Hynd, 2003). Obviously, the current findings do not directly assess any of these speculations but they do suggest a clear biologic constraint: The systems are weakened but not wholly dysfunctional in even severe older RD readers. The general notion of a ''developmental lesion'' at critical LH systems in RD (cf., Eden & Zeffiro, 1998) would appear to be inconsistent with data suggesting functional activation of these systems under certain conditions (see

Pugh, Mencl, Shaywitz, et al., 2000 for similar conclusions with functional connectivity analyses). Speculatively, these findings seem most consistent with accounts that posit ''noisy'' or unstable neural systems (Sperling, Lu, Manis, & Seidenberg, 2005).

Summary

In Experiment 1, effects of imageability and frequency on behavioral performance in NI and RD participants were similar. Both groups show facilitation but, as predicted, the benefit was larger for RD. Effects of imageability and frequency on brain activation in phonologically tuned subsystems in NI and RD were wholly dissimilar. For NI readers, easier words were associated with relatively reduced activation. For RD readers, easier words resulted in increased activation. Experiment 2 replicated this pattern with a simple and direct manipulation of on-line learning through stimulus repetition. Thus, the phonologically tuned subsystems in adolescent RD readers appear to be poorly trained but not wholly disrupted. Further studies will be required to test the limits on learning in these LH systems in RD.

Acknowledgments

Note

1. Two of the 16 readers were unavailable for testing, but their overall performance on the in-scanner task was 97% and 98% correct and they had no history of reading difficulties.

REFERENCES

Adams, M. J. (1994) Beginning to read: Thinking and learning about print. Cambridge, MA: MIT Press.

Bruck, M. (1992). Persistence of dyslexics' phonological deficits. Developmental Psychology, 28, 874–886.

Cao, F., Bitan, T., Chou, T. L., Burman, D. D., & Booth, J. R. (2006). Deficient orthographic and phonological representations in developmental dyslexics revealed by brain activation patterns. Journal of Child Psychology and Psychiatry, 40, 1041–1050.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. Behavioral Research Methods, Instruments, and Computers, 25, 257–271.

Eden, G. F., & Zeffiro, T. A. (1998). Neural systems affected in developmental dyslexia revealed by functional neuroimaging. Neuron, 21, 279–282.

Friston, K. J., Ashburner, J., Frith, C. D., Poline, J.-B., Heather, J. D., & Frackowiak, R. S. J. (1995). Spatial registration and normalization of images. Human Brain Mapping, 2, 165–189.

Frost, S. J., Mencl, W. E., Sandak, R., Moore, D. L., Rueckl, J., Katz, L., et al. (2005). An fMRI study of the trade-off between semantics and phonology in reading aloud. NeuroReport, 16, 621–624.

Galaburda, A. M. (1992). Neurology of developmental dyslexia. Current Opinion in Neurology and Neurosurgery, 5, 71–76.

Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. Neuroimage, 15, 870–878.

Hadjikhani, N., & Roland, P. E. (1998). Cross-modal transfer of information between the tactile and the visual representations in the human brain: A positron emission tomographic study. Journal of Neuroscience, 18, 1072–1084.

Henson, R. N., Price, C. J., Rugg, M. D., Turner, R., & Friston, K. J. (2002). Detecting latency differences in event-related BOLD responses: Application to words versus pseudowords and initial versus repeated face presentations. Neuroimage, 15, 83–97.

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

Holmes, A. P., & Friston, K. J. (1998). Generalizability, random effects, and population inference. Neuroimage, 7, S34.

Jared, D., McRae, K., & Seidenberg, M. S. (1990). The basis of consistency effects in word naming. Journal of Memory and Language, 29, 687–715.

Josephs, O., Turner, R., & Friston, K. (1997). Event-related fMRI. Human Brain Mapping, 5, 243–248.

Katz, L., Lee, C. H., Tabor, W., Frost, S. J., Mencl, W. E., Sandak, R., et al. (2005). Effects of printed word repetition in lexical decision and naming on behavior and brain activation. Neuropsychologia, 43, 2068–2083.

Kirk, R. E. (1982). Experimental design: Procedures for the social sciences. Belmont, CA: Wadsworth.

Klingberg, T., Hedehus, M., Temple, E., Salz, T., Gabrieli, J. D., Moseley, M. E., et al. (2000). Microstructure of temporo-parietal white matter as a basis for reading ability: Evidence from diffusion tensor magnetic resonance imaging. Neuron, 5, 493–500.

Miezin, F. M., Maccotta, L., Ollinger, J. M., Petersen, S. E., & Buckner, R. L. (2000). Characterizing the hemodynamic response: Effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. Neuroimage, 11, 735–759.

Miller, C. J., Sanchez, J., & Hynd, G. W. (2003). Neurological correlates of reading disabilities. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), Handbook of learning disabilities. New York: Guilford Press.

Papademetris, X., Jackowski, A. P., Schultz, R. T., Staib, L. H., & Duncan, J. S. (2003). Computing 3D non-rigid brain registrations using extended robust point matching for composite multisubject fMRI analysis. In R. E. Ellis & T. M. Peters (Eds.), Medical image computing and computer assisted intervention (pp. 788–795). Berlin: Springer-Verlag.

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

Poldrack, R. A., & Gabrieli, J. D. E. (2001). Characterizing the neural mechanisms of skill learning and repetition priming: Evidence from mirror reading. *Brain, 124,* 67–82.

Posner, M. I., Abdullaev, Y. G., McCandliss, B. D., & Sereno, S. C. (1999). Anatomy, circuitry and plasticity of word reading. In J. Everatt (Ed.), *Reading and dyslexia: Visual and attentional processes* (pp. 137–162). London: Routledge.

Price, C. J. (2000). The anatomy of language: Contributions from functional neuroimaging. *Journal of Anatomy, 197,* 335–359.

Pugh, K., Mencl, W. E., Shaywitz, B. A., Shaywitz, S. E., Fulbright, R. K., Skudlarski, P., et al. (2000). The angular gyrus in developmental dyslexia: Task-specific differences in functional connectivity in posterior cortex. *Psychological Science, 11,* 51–56.

Pugh, K. R., Mencl, W. E., Jenner, A. R., Katz, L., Frost, S. J., Lee, J. R., et al. (2000). Functional neuroimaging studies of reading and reading disability (developmental dyslexia). *Mental Retardation and Developmental Disabilities Research Reviews, 6,* 207–213.

Pugh, K. R., Shaywitz, B. A., Shaywitz, S. A., Constable, R. T., Skudlarski, P., Fulbright, R. K., et al. (1996). Cerebral organization of component processes in reading. *Brain, 119,* 1221–1238.

Sandak, R., Mencl, W. E., Frost, S., & Pugh, K. R. (2004). The neurobiological basis of skilled and impaired reading: Recent findings and new directions. *Scientific Studies of Reading, 8,* 273–292.

Sandak, R., Mencl, W. E., Frost, S. J., Rueckl, J. G., Katz, L., Moore, D., et al. (2004). The neurobiology of adaptive learning in reading: A contrast of different training conditions. *Cognitive, Affective, & Behavioral Neuroscience, 4,* 67–88.

Sarkari, S., Simos, P. G., Fletcher, J. M., Castillo, E. M., Breier, J. I., & Papanicolaou, A. C. (2002). The emergence and treatment of developmental reading disability: Contributions of functional brain imaging. *Seminars in Pediatric Neurology, 9,* 227–236.

Seidenberg, M. S., Waters, G. S., Barnes, M., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior, 23,* 383–404.

Shaywitz, B., Shaywitz, S., Blachman, B., Pugh, K. R., Fulbright, R., Skudlarski, P., et al. (2004). Development of left occipito-temporal systems for skilled reading following a phonologically based intervention in children. *Biological Psychiatry, 55,* 926–933.

Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Mencl, W. E., Fulbright, R. K., Constable, R. T., et al. (2002). Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biological Psychology, 52,* 101–110.

Shaywitz, S. E., Shaywitz, B. A., Fulbright, R. K., Skudlarski, P., Mencl, W. E., Constable, R. T., et al. (2003). Neural systems for compensation and persistence: Young adult outcome of childhood reading disability. *Biological Psychiatry, 54,* 25–33.

Simos, P. G., Fletcher, J. M., Bergman, E., Breier, J. I., Foorman, B. R., Castillo, E. M., et al. (2002). Dyslexia-specific brain activation profile becomes normal following successful remedial training. *Neurology, 58,* 1203–1213.

Sperling, A. J., Lu, Z., Manis, F. R., & Seidenberg, M. S. (2005). Deficits in perceptual noise exclusion in developmental dyslexia. *Nature Neuroscience, 8,* 862–863.

Strain, E., & Herdman, C. M. (1999). Imageability effects in word naming: An individual differences analysis. *Canadian Journal of Experimental Psychology, 53,* 347–359.

Strain, E., Patterson, K., & Seidenberg, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1140–1154.

Strain, E., Patterson, K., & Seidenberg, M. S. (2002). Theories of word naming interact with spelling–sound consistency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 207–214.

Temple, E., Deutsch, G. K., Poldrack, R. A., Miller, S. L., Tallal, P., Merzenich, M. M., et al. (2003). Neural deficits in children with dyslexia ameliorated by behavioral remediation: Evidence from functional MRI. *Proceedings of the National Academy of Sciences, U.S.A., 100,* 2860–2865.

Woods, R. P. (1996). Modeling for intergroup comparisons of imaging data. *Neuroimage, 4,* S84–S94.