The Impact of Intensive Reading Intervention on Level of Attention in Middle School Students

Greg Roberts, Shruti Rane, Anna-Mária Fall, Carolyn A. Denton, Jack M. Fletcher & Sharon Vaughn

Published online: 02 Jun 2014.

Submit your article to this journal

Article views: 425

View related articles

View Crossmark data

Routledge
Taylor & Francis Group

The Impact of Intensive Reading Intervention on Level of Attention in Middle School Students

Greg Roberts

The Meadows Center for Preventing Educational Risk, University of Texas at Austin

Shruti Rane

Department of Neuropsychology, Kennedy Krieger Institute, Johns Hopkins University

Anna-Mária Fall

The Meadows Center for Preventing Educational Risk, University of Texas at Austin

Carolyn A. Denton

Department of Pediatrics, University of Texas Health Science Center Houston

Jack M. Fletcher

Department of Psychology, University of Houston

Sharon Vaughn

The Meadows Center for Preventing Educational Risk, University of Texas at Austin

The purpose of the study was to estimate the impact of reading intervention on ratings of student attention over time. We used extant data from a longitudinal randomized study of a response-based reading intervention to fit a multiple-indicator, multilevel growth model. The sample at randomization was 54% male, 18% limited English proficient, 85% eligible for free or reduced-price lunch, 58% African American, and 32% Hispanic. Reading ability was measured by using the Woodcock–Johnson III Tests of Achievement. Attention was measured by using the Strengths and Weaknesses of ADHD Symptoms and Normal Behavior Scale. Findings indicate that intensive, response-based reading intervention over 3 years improved reading achievement *and* behavioral attention in middle school struggling readers, with treatment directly affecting reading, which in turn influenced attention. In the business-as-usual condition, there was no relation between improved reading and attention. The results are consistent with a correlated liabilities model of comorbidity. The results do not align with the inattention-as-cause hypothesis, which predicts that reading intervention should not affect attention. The findings do not support, but do not necessarily preclude, the phenocopy hypothesis. The results are especially pertinent for older students who may be inattentive partly because of years of struggling with reading.

Many children with reading problems are also inattentive, especially during adolescence (Franzak, 2006). Approximately 4% to 5% of the population aged 12 to 17 experience co-occurring reading disability (RD) and attention-deficit/hyperactivity disorder

Correspondence should be addressed to Greg Roberts, The Meadows Center for Preventing Educational Risk, The University of Texas at Austin, Sanchez 252 D4900, Austin, TX 78712. E-mail: gregroberts@mail.utexas.edu

(ADHD; Carroll, Maughn, Goodman, & Meltzer, 2005; Pastor & Reuben, 2008). Among those with ADHD, approximately 20% are identified with RD (Carroll et al., 2005). Of children with RD, about 9% also meet criteria for ADHD (Willcutt & Pennington, 2000). Youth with RD are about 4 times more likely to have ADHD than children without RD (Carroll et al., 2005). The inattentive symptoms associated with ADHD, rather than its hyperactivity-impulsivity dimension, largely account for the RD–ADHD link (Hinshaw, 1992; Willcutt, Betjemann, et al., 2010; Willcutt, Pennington, et al., 2010). Many children identified with RD who are below the threshold for ADHD-inattention type (i.e., not identified as ADHD) nonetheless struggle in both domains (Merrell & Tymms, 2001).

Three hypotheses have been advanced to explain the causal relationship of reading problems and inattention. The first, the *inattention-as-cause hypothesis*, assigns a causal role to inattention, with poor attention in the early stages of reading acquisition leading to increasingly pervasive reading problems (Rabiner, Coie, & The Conduct Problems Prevention Research Group, 2000; Stanovich, 1986). A second, the *phenocopy hypothesis* (Pennington, Groisser, & Welsh, 1993), assumes that reading difficulties mimic the symptoms of ADHD-related inattention (Hinshaw, 1992). Students who experience difficulties in classroom reading tasks are less engaged and appear inattentive. Attempts to replicate early support for the phenocopy hypothesis (Pennington et al., 1993) have been unsuccessful. The inattention-as-cause and phenocopy hypotheses assume that RD and ADHD are cognitively distinct disorders.

The third hypothesis, the *correlated liabilities hypothesis*, proposes that the disorders share a common subset of weaknesses or risk factors (Pennington, 2006; Willcutt, Pennington, et al., 2010). The correlated liabilities hypothesis is consistent with genetic research that indicates an overlap between RD and ADHD (Gayan et al., 2005)—neuroimaging studies suggest shared neural correlates across the two disorders (Foster, Hynd, Morgan, & Hugdahl, 2002), and neuropsychological research demonstrates common underlying deficits, such as slowed processing speed (McGrath et al., 2011). Jacobson and colleagues (Jacobson, Ryan, Denckla, Mostofsky, & Mahone, 2013; Jacobson et al., 2011) reported that children with ADHD vary in reaction time, which significantly predicts reduced oral reading fluency and may relate to the neuroanatomical underpinnings of ADHD (i.e., abnormalities in white matter pathways) that also influence language automaticity.

These three hypotheses predict different attention-related outcomes for reading intervention. The inattention-as-cause hypothesis suggests no change in teacher-reported attention due to treatment for reading difficulties. The effect on reading outcomes would be inconsistent, as well, to the extent that inattention is left untreated (e.g., medication). Hypothetically, students with inattention and RD would continue to struggle in both areas until inattention is effectively addressed. The phenocopy hypothesis predicts that appropriate reading intervention positively influences attention. Students report that instruction aligned with and responsive to struggling readers' needs is more engaging (Rabiner et al., 2000); teachers of struggling adolescents who participate in well-designed instructional programs describe their students as more attentive and more engaged (Hinshaw, 1992). Further, because inattention in the phenocopy hypothesis is represented as an affective response to inadequate instruction (Hinshaw, 1992), rather than as a distinct neuropsychological disorder, improved teacher-rated attention may occur relatively early as students are presented with meaningful opportunities to learn. A given treatment for RD may not *directly* target attention; however, to the extent that the two effects are proximal, the effect of RD intervention on attention-related outcomes may be represented best as a *direct* effect, at least in the absence of data on mediating influences. Improvements in reading and attention may covary, but the effect of RD treatment on inattention would not depend on improved reading.

Like the phenocopy hypothesis, the correlated liabilities perspective predicts effects on both reading and attention-related outcomes to the extent that treatment addresses the neuropsychological deficits shared by RD and ADHD-inattention. However, changes in attention may become evident later (compared to the onset of improved attention predicted by the phenocopy hypothesis) and only to the extent that elements of the reading intervention also target a subset of malleable competencies shared by RD and ADHD-inattention. This hypothesis is best represented as a mediation model, where RD intervention improves reading and attention to the extent that a subset of the treatment activities addresses the group of shared liabilities. Neither RD treatment nor improved reading would necessarily "cause" improved attention; instead, RD treatment would affect inattention only by improving the liabilities shared by RD and inattention.

PRIOR RESEARCH ON RD INTERVENTION AND INATTENTION

Although limited, intervention research suggests that treatment effects are more difficult to realize with children who have comorbid RD–ADHD than those with only RD or only ADHD for both reading and attention-related outcomes (e.g., Rabiner & Malone, 2004). Treatments of ADHD with stimulant medication

(Bental & Tirosh, 2008; Keulers et al., 2007), behavioral interventions (DuPaul et al., 2006; Volpe, DuPaul, Jitendra, & Tresco, 2009), and their combination (MTA Cooperative Group, 1999) are associated with improved reading, though the effects fade over time (MTA Cooperative Group, 2004). Rabiner and Malone (2004) found that participation in reading intervention does not benefit students with comorbid reading and attention problems, even though it improved the reading outcomes of children with *either* RD or inattention (i.e., those without comorbidity). Rabiner and Malone sampled first-graders.

PURPOSES OF THIS STUDY

Our focus is students in middle school. We used data from a 3-year trial of a response-based reading intervention for struggling adolescent readers (summarized in Vaughn & Fletcher, 2012) to evaluate the proposed hypotheses. We modeled student reading and student attention over time through multiple-indicator, multilevel latent variable growth models, using the slope estimates for reading and attention to test structural models associated with the three hypotheses: (a) no reading treatment effects on attention, as indicated by the inattention-as-cause hypothesis; (b) a direct reading treatment effect on attention, per the phenocopy hypothesis; and (c) a direct reading treatment effect on reading and an indirect effect on attention, as suggested by the correlated liabilities hypothesis.

Because the reading intervention was *not directly* developed for students with inattentive behavior, there were no sampling constraints on levels of attention in the original design, and we did not impose constraints in the secondary analysis reported here. Instead, we treated attention as dimensional, representing it across the range of possible levels within the group of students with RD, given the possibility that the attributes of RD and ADHD are dimensional in the population of school-aged children (Coghill & Sonuga-Barke, 2012; Lubke et al., 2007; Snowling & Hulme, 2012). Distinct groups may *not*, in fact, represent the distribution for each disorder, and individual differences may be better described in terms of degree than kind.

METHOD

School Sites

This study was conducted with Institutional Review Board approval in two large urban cities (one large district, one medium district) in the southwestern United States, with approximately half the sample from each site. Students from seven middle schools (three from the first and larger site and four from the second site) participated in the study. The three schools from the first site were classified as urban; the remaining four schools were classified as near urban, with school populations ranging in size from 633 to 1,300 students. The rate of students qualifying for free or reduced-priced lunch ranged from 56% to 86% across the schools in the larger site and from 40% to 85% in the smaller site.

Student Participants

The initial sampling frame included 2,034 fifth-grade students who had useable and eligible state test scores (the Texas Assessment of Knowledge and Skills [TAKS]) and who were expected to attend one of seven middle schools agreeing to participate in the study (Figure 1). Students were excluded if they (a) were enrolled in an alternative curriculum (i.e., life skills class), (b) performed at or below a second-grade reading level, (c) were identified with significant sensory disability (e.g., blindness, deafness), or (d) were prevented from participation due to an individualized special education plan.

We identified 768 eligible students as struggling readers, based on their spring of fifth-grade TAKS performance. These students were randomized in the early summer of 2006 within schools to the response-based, tiered condition (treatment group) or to the business-as-usual condition, using a 2:1 assignment ratio to maximize the potential benefits of treatment. Approximately 54% of the total students were male (53% in treatment and 55% in comparison), 18% were limited English proficient (18% in treatment and 17% in comparison), and 85% received free or reduced-price lunch (82% in treatment and 85% in comparison). More than half of the sample was African American (58% in treatment and 57% in comparison) and about 32% was Hispanic (32% in both conditions). Approximately 9% of the sample was White (9% in both conditions) and 1% was Asian (1% in treatment and 2% in comparison). The proportions of students in the two conditions did not differ by gender, free or reduced-price lunch status, English proficiency, or ethnicity (all $p > .05$).

The initial treatment status was maintained over the course of the study. Only students assigned initially to intervention were eligible for treatment in Year 2 and/ or Year 3 and only if they continued to struggle, based on data from the annually administered screening indicators (see Measures section for details). Students assigned to the business-as-usual condition were provided with the schools' typical programs over the 3-year period.

The catchment areas for participating middle schools in one of the districts were realigned during the summer between randomization and the onset of intervention,
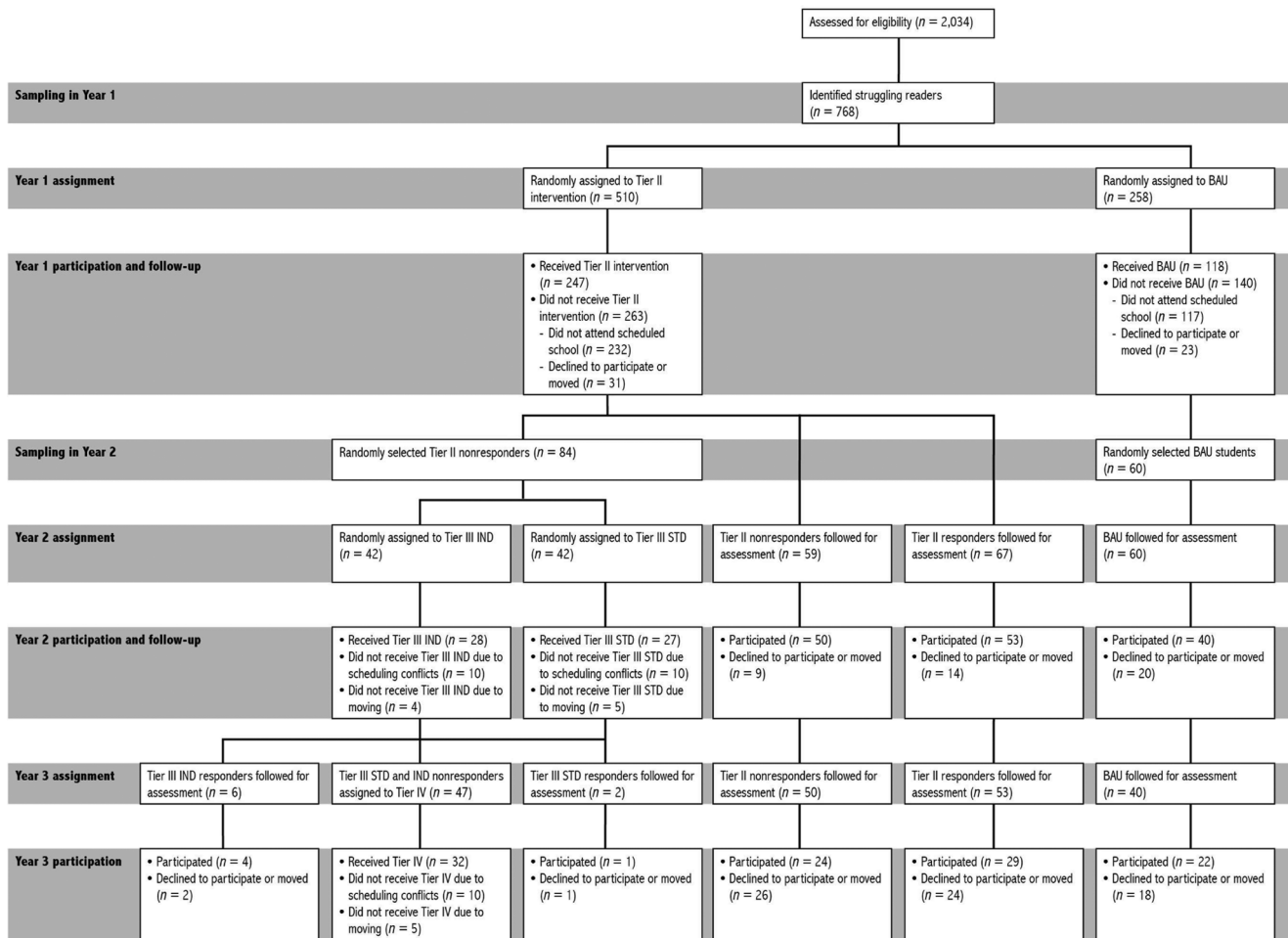
FIGURE 1 Initial assignment and ongoing participation.

Assessed for eligibility (*n* = 2,034)

Sampling in Year 1
Identified struggling readers (*n* = 768)

Year 1 assignment
Randomly assigned to Tier II intervention (*n* = 510)
Randomly assigned to BAU (*n* = 258)

Year 1 participation and follow-up
- Received Tier II intervention (*n* = 247)
- Did not receive Tier II intervention (*n* = 263)
  - Did not attend scheduled school (*n* = 232)
  - Declined to participate or moved (*n* = 31)

- Received BAU (*n* = 118)
- Did not receive BAU (*n* = 140)
  - Did not attend scheduled school (*n* = 117)
  - Declined to participate or moved (*n* = 23)

Sampling in Year 2
Randomly selected Tier II nonresponders (*n* = 84)
Randomly selected BAU students (*n* = 60)

Year 2 assignment
Randomly assigned to Tier III IND (*n* = 42)
Randomly assigned to Tier III STD (*n* = 42)
Tier II nonresponders followed for assessment (*n* = 59)
Tier II responders followed for assessment (*n* = 67)
BAU followed for assessment (*n* = 60)

Year 2 participation and follow-up
- Received Tier III IND (*n* = 28)
- Did not receive Tier III IND due to scheduling conflicts (*n* = 10)
- Did not receive Tier III IND due to moving (*n* = 4)

- Received Tier III STD (*n* = 27)
- Did not receive Tier III STD due to scheduling conflicts (*n* = 10)
- Did not receive Tier III STD due to moving (*n* = 5)

- Participated (*n* = 50)
- Declined to participate or moved (*n* = 9)

- Participated (*n* = 53)
- Declined to participate or moved (*n* = 14)

- Participated (*n* = 40)
- Declined to participate or moved (*n* = 20)

Year 3 assignment
Tier III IND responders followed for assessment (*n* = 6)
Tier III STD and IND nonresponders assigned to Tier IV (*n* = 47)
Tier III STD responders followed for assessment (*n* = 2)
Tier II nonresponders followed for assessment (*n* = 50)
Tier II responders followed for assessment (*n* = 53)
BAU followed for assessment (*n* = 40)

Year 3 participation
- Participated (*n* = 4)
- Declined to participate or moved (*n* = 2)

- Received Tier IV (*n* = 32)
- Did not receive Tier IV due to scheduling conflicts (*n* = 10)
- Did not receive Tier IV due to moving (*n* = 5)

- Participated (*n* = 1)
- Declined to participate or moved (*n* = 1)

- Participated (*n* = 24)
- Declined to participate or moved (*n* = 26)

- Participated (*n* = 29)
- Declined to participate or moved (*n* = 24)

- Participated (*n* = 22)
- Declined to participate or moved (*n* = 18)

Note. Sections labeled *sampling* represent cases available for randomization or for tracking at the outset of each school year. *Participation and follow-up* represent students who participated as assigned and completed the end-of-year test battery. *Sampling* is not indicated for the final year because all available cases from Year 2 were continued in Year 3. Initial randomization (*Year 1 assignment*) represents the independent variable; we made intent-to-treat assumptions and used a full information likelihood estimator, so analysis included all available data for students randomized at that point. BAU = business-as-usual group; Tier III IND = individualized Tier III intervention; Tier III STD = standardized Tier III intervention.

diverting a subsample of the students to nonparticipating sites. This change resulted in 419 of the 768 originally randomized students enrolling in one of the seven participating middle schools—278 students who were originally randomized to the reading treatment group and 141 students randomized to business as usual, representing an approximate 40% loss in both conditions. For the larger sample (*n* = 768), there were no differences on average TAKS scores at pretest (*p* > .05). The participating (*n* = 419) and reassigned (*n* = 349) students did not differ on key demographics, including race, gender, primary language status, and free or reduced-price lunch status, and there were no interactions with group. Also, within the final treatment and business-as-usual groups (*n* = 419), there were no differences in demographics or TAKS scores across conditions at pretest (*p* > .05). The proportion of students

not available in the fall of sixth grade did not differ between the two experimental conditions (*p* > .05). Of the 419 returning students, 365 completed Year 1 of the study, 247 from the treatment group and 118 from business as usual.

Teacher Participants

Researchers hired experienced teachers to provide interventions for each of the 3 years. Teachers varied each year (see descriptions of teachers for Year 1 in Vaughn et al., 2010; for Year 2 in Vaughn et al., 2011; and for Year 3 in Vaughn et al., 2012), but all had high levels of knowledge and considerable experience teaching students with RD. The research team provided intervention teachers with approximately 60 hr of

professional development each year prior to implementation. Teachers also participated in biweekly staff development meetings and received regular (once every 1–2 weeks) on-site feedback and coaching. The research staff hired and managed the intervention teachers, and students participated during their elective period, thus minimizing threats due to treatment contamination.

Intervention

The response-based, tiered instructional model was the independent variable for all years of the study. All students (both conditions) received Tier I, or core, reading instruction from their regular teachers; however, treatment students also received supplemental Tier II reading intervention during Year 1 of the study. Students who inadequately responded to this Tier II intervention moved to more intensive Tier III intervention in study Year 2. Likewise, students who responded inadequately to Tier III received Tier IV intervention during study Year 3. Thus, in the treatment condition, the intervention intensity varied based on student response, but students stayed in their original group, treatment or business as usual, for the entire study.

This design is a departure from those typical in intervention research. In this case, the experimental manipulation was *access to more intensive instruction* and included 1 to 3 years of intensive reading support, depending on student response. Inadequate responders in the treatment condition received more intensive instruction. Business-as-usual students received the intervention their school offered. The experimental reading program, designed for students reading more than 2 years below grade level, supplied additional instructional intensity; however, it was not the independent variable because its application was conditional on student response. We describe business as usual, Tier II, Tier III, and Tier IV in the following sections.

Business as usual. Business as usual represents at-risk readers' typical school experience. Each year of the study, researchers provided content area (e.g., social studies, science) teachers with 6-hr professional development sessions on embedding evidence-based vocabulary and reading comprehension instruction across the school day, benefiting all students. In addition, struggling readers in the larger school district attended a 45-min reading class and an 85-min English language arts class daily. Treatment students in this district received both the 45-min intervention provided by the district and the experimental intervention. Treatment and business-as-usual students in the smaller school district attended a 50-min English language arts class daily. The smaller district did not offer additional reading classes for struggling readers, meaning that

treatment students in the larger district received two daily reading classes and treatment students in the smaller district received one daily reading class. Both districts provided rigorous test preparation to all students each year before the statewide high-stakes reading assessment.

We also collected data on additional reading instruction beyond school-provided opportunities (e.g., after-school tutoring). For both the initial business-as-usual sample and the treatment sample, about 25% of the students received this type of additional reading instruction, averaging about 30 hr total ($SD = 15$ hr) across the 3 years of the study.

Tier II. In Project Year 1, students in the treatment condition participated in secondary intervention classes of 10 to 12 students during their elective period (business-as-usual students attended their regular school-provided electives in all years of the study). Intervention was provided in daily 50-min sessions for about 160 sessions per school year.

In the 1st year, instructors sequentially implemented a three-phase standardized treatment protocol; each phase lasted several months. *Phase I* initially focused on word study and fluency, increasingly moving to vocabulary, sentence and paragraph meaning, and overall comprehension. Repeated reading practice of 10- to 15-min daily supported fluency. Students tracked their own progress through oral reading fluency assessments. Advanced strategies for decoding multisyllabic words (word study) were taught through the REWARDS program (Archer, Gleason, & Vachon, 2005). Progress in the program depended on individual students mastering sounds and word reading. Students received instruction in and practiced daily with letter sounds, letter combinations, and affixes. To decode and spell multisyllabic words, students were taught to use a segmentation strategy. Daily vocabulary instruction drew from the words in the texts used for instruction and practice. Instructors provided student-friendly definitions of the words along with examples and nonexamples of appropriate use of the new words. Reading, writing, and verbal use reinforced word knowledge. To develop text comprehension, students answered questions of varying levels of difficulty (literal and inferential) during and after reading a passage. These questions helped to check for understanding and model active thinking during reading. Students learned how to use text to answer questions and justify their answers.

Phase II focused on vocabulary and comprehension, revisiting the Phase I word study and fluency skills and strategies when needed. In addition to the previously described Phase I vocabulary work, students learned about word relatives and parts of speech (e.g., *install, installation, installable*) and how to apply word

study principles to spelling (encoding vs. decoding). Again, the vocabulary words were from the texts students read in class, including both narrative (e.g., novels) and expository (e.g., informational text) sources. Teachers previewed the vocabulary words and their spelling and then previewed the passage with students. Then, teachers guided students in an initial reading, asking questions to model active thinking and to check for understanding. To process and summarize information, students completed a graphic organizer while reading. Students also completed writing activities to summarize the new content.

In *Phase III*, students learned how to apply word-level and comprehension practices to the expository texts they encountered in social studies and science. This phase placed particular emphasis on comprehension and critical thinking at the sentence, paragraph, and multi-paragraph levels.

After 9 months, participants in the Tier II intervention made statistically significant gains on measures of spelling, word attack, phonemic decoding efficiency, and passage comprehension (TAKS) compared to students in business as usual (Vaughn et al., 2010). Decoding, fluency, and reading comprehension performance, as measured by the Woodcock–Johnson III Tests of Achievement (WJ-III; Woodcock, McGrew, & Mather, 2001), favored the treatment group, but the gains did not differ statistically from those of business-as-usual students. The median effect size across all reading outcome measures was .16 (Cohen's *d*).

Tier III and Tier IV. In Project Year 2 (Tier III), treatment students who did not pass the TAKS were randomly assigned to standardized treatment or individualized treatment (see Figure 1). To provide a more intensive intervention, instruction was delivered in groups of five students per interventionist during students' elective period. Teachers in the individualized treatment varied the instructional focus, the time allocated daily to each instructional task, and instructional pacing over time for each student, based on the results of weekly curriculum-based measures developed for the intervention. In addition, teachers implemented a motivational plan based on students' interests and a goal-setting plan.

In Project Year 3 (Tier IV), class size for low responders was reduced further to two to four students per group, again during students' elective period. Individualized instructional programs were developed for each student, based on the previously described protocol. Business as usual in Year 2 and Year 3 was similar to that in Year 1.

In a 9-month efficacy trial of the Tier III interventions, students in both the standardized and individualized treatment groups outperformed students in the business-as-usual group ($n = 59$ and $n = 22$ in fall of 2007) on assessments of decoding, fluency, and reading comprehension. These differences were statistically significant. Intervention type did not moderate the pattern of effects, although students in the standardized treatment slightly outperformed individualized students on word attack. Vaughn et al. (2011) provide a complete description of the Tier III intervention.

Tier IV included a 50-min intensive reading intervention daily for eighth graders with intractable reading problems. This intervention was effective, as measured by the WJ-III Word Identification ($ES = 0.49$) and Reading Comprehension ($ES = 1.20$) subtests. However, despite 3 years of intensive intervention, treatment students continued to lack grade-level reading proficiency.

Allocation across the 3-year study. Figure 1 summarizes allocation and follow-up for groups of students after initial assignment. In addition to the expected attrition, we reduced the business-as-usual sample during the summer between Year 1 and Year 2 to accommodate the project budget (see Figure 1). We randomly selected approximately 50% ($n = 60$) of the business-as-usual group for ongoing follow-up. Our purpose was to select a sufficiently large sample to reliably estimate the performance of the 118 business-as-usual students who participated in Year 1 and were subject to follow-up in subsequent years.

We also randomly selected approximately 80% of the Tier II inadequate responders (143 of 180 cases; we no longer followed the remaining 37) for assignment or follow-up in Year 2. From this group, we randomly assigned 84 students to one of the two Tier III instructional conditions—individualized and standardized—in Year 2. We carried out this additional randomization within the group of students initially randomized to treatment, preserving the internal validity of initial assignment. Our purpose at the time was to test differences in the teacher-directed versus more standardized approaches to reading intervention. We found no comprehension-related differences in the effects of the different Tier III interventions (Vaughn & Fletcher, 2012), although both groups outperformed the group of selected, but not assigned, Tier II inadequate responders ($n = 59$).

In Year 2, we followed the full sample of Tier II responders ($n = 67$), whom we provided no further intervention. These Tier II responders maintained the initial reading gains in the remaining years, despite receiving no additional intervention. This group accounted for about a third of the sample at the end of Year 1 that was assigned to treatment (67 of 210 students), and we reasoned that these students would continue to represent the upper third of performers in the treatment group. This group of Tier II responders continued to

score better than about 67% of students in the sample, on average.

In Year 3, sample sizes reduced due to the usual attrition, schedule conflicts, and students declining to participate. The sample size in Year 3 ($n = 112$) represented 28% of the original group of 419 students. About 207 of the 307 cases that stopped participating in the project did so because the family moved; because the child declined to participate; or for treatment students, because of scheduling conflicts. The remaining 100 students were randomly removed from the sample, as just described.

To evaluate differential attrition over the 3 years, we regressed treatment condition, completer status (i.e., did or did not attrite), and the interaction of condition and completer status on the WJ-III pretest measures described next and key demographic variables. We found no significant condition by completer status effects for gender, English proficiency, ethnicity, free or reduced-price lunch status, special education, or Time 1 scores on the WJ-III subtests, suggesting no significant differential attrition. Further, we refit the reading model (described next), using listwise deletion (i.e., included only cases with data at all time points) rather than full information maximum likelihood. Parameters in the revised model did not differ statistically from the values based on full information maximum likelihood, supporting the model's validity and the randomness of missing data.

Measures

W–J III tests of achievement. Reading ability was measured by using the Letter-Word Identification, Word Attack, and Passage Comprehension subtests of the WJ-III (Woodcock et al., 2001). A latent measure of *reading* was created based on *W* scores. Coefficient alphas in the Grade 6 sample were .97 at pretest and .92 at posttest for Letter-Word Identification, .93 at pretest and .99 at posttest for Word Attack, and .94 at pretest and .85 at posttest for Passage Comprehension.

Strengths and weaknesses of ADHD symptoms and normal behavior scale. Students' *attention* was measured by using the Attention Deficit subtest of the Strengths and Weaknesses of ADHD Symptoms and Normal Behavior Scale (SWAN; Swanson et al., 2006). The Attention Deficit subtest measures attention, using a 7-point scale ranging from far below average to far above average. We used Items 1 to 9 to estimate latent values for inattention (Ullebo, Breivik, Gillberg, Lundervold, & Posserud, 2012). Increased inattentive behavior leads to lower scores on the SWAN rating scale. The SWAN conceptualizes the *Diagnostic and Statistical Manual of Mental Disorders* symptom domains of ADHD as a continuum that covers the full range of relevant aptitudes, characterizing this component of the ADHD construct as dimensional (Coghill & Sonuga-Barke, 2012). Arnett et al. (2013) found that the SWAN is comparable to the Disruptive Behavior Rating Scale in validity, reliability, and heritability and the SWAN is comparably heteroscedastic, making it a preferred measure of positive attention and impulse regulation behaviors.

Procedures

The WJ-III subtests were administered before and after each of the 3 years of intervention. We used WJ-III data from the fall and spring of Year 1 and from the spring of Year 2 and Year 3. Research staff members trained and monitored the test administrators, who were blind to treatment condition, and all tests were administered in quiet settings in the schools as part of a larger test battery. The child's language arts teachers (different in each grade) completed the SWAN attention rating scale each spring, producing three data points. These raters were blind to treatment group assignment.

Analysis

We estimated trends in reading and attention by using a multiple-indicator, multilevel (three-level) latent variable growth model, which maximizes the advantages of structural equation modeling in a longitudinal context by explicitly modeling measurement error in observed variables and by constraining measurement variance across time points. This model yields more precise trend estimates (Wu, Li, & Zumbo, 2007), provides a more rigorous basis for validity claims about the structure of a given data set (Valentine & McHugh, 2007), and offers a framework for evaluating measurement characteristics across time.

Level 1 was conceptualized as the measurement model, with reading and attention estimated as latent continuous factors at points in time. For reading, we used *W* scores from the three WJ-III subtests to estimate status on each of four occasions. We estimated attention on each of three occasions, using items from the SWAN. Measurement error in structural models is specifically estimated, so that for any given occasion (assuming a well-fit model), predicted values are more reliable estimates than the observed scores on which they are based. At Level 2, trends in factor scores (estimated in Level 1) were fit for reading and attention over the 3-year period. We modeled reading as curvilinear, with intercept, slope, and a quadratic term that represented the deceleration in slope over time. Attention was modeled as linear. We modeled the variation in growth-related parameters

at Level 3, including treatment effects due to the reading intervention and interrelationships among the growth parameters for reading and attention (i.e., the models suggested by the three hypotheses described in this article's introduction).

The intercept at Time 1 in business as usual was fixed at 0 in the reading and attention growth models, providing a baseline for contrasts across time and between groups. Temporal invariance, a requirement in multiple-indicator, multilevel models, was evaluated and constraints were imposed as necessary (Meredith, 1993). The assumption of equal errors over time is generally untenable (Wu et al., 2007), so we allowed residual variances to estimate freely. Cross-group measurement invariance was assumed, given the randomized nature of the design and the imposition of temporal constraints in the measurement model. We conducted the analyses by using Mplus 7.0 (Muthén & Muthén, 2012).

RESULTS

Sample values at each time point for reading and for attention are in Table 1. Students were about 1 standard deviation below the normative mean on the Passage Comprehension subtest of the WJ-III prior to intervention. Treatment students made gains in standard scores relative to the normative group over the 3 years. Business-as-usual students lost ground compared to their age group, as indicated by the decreasing average standard score over time. Raw scores on the SWAN items suggest increased levels of reported attention in the treatment group. There were no differences over time in the business-as-usual group.

The baseline multiple-indicator, multilevel model with reading and attention trends estimated simultaneously fit the data well, $\chi^2(721) = 1225.27$, comparative fit index = .96, Tucker-Lewis index = .96, root mean square error of approximation (RMSEA) = .039, RMSEA confidence interval [.035, .042]. For the reading model, longitudinal invariance for factor loadings was supported ($\Delta_{\chi2}/\Delta df = 5.289/6$, $p > .01$); invariance in intercepts was not, and the model was refit with the intercept for Word Attack freely estimated across the four measurement occasions. Full measurement invariance was supported for the attention model ($\Delta_{\chi2}/\Delta_{df} = 25.853/16$, $p > .01$, for factor loadings and $\Delta_{\chi2}/\Delta_{df} = 27.97/18$, $p > .01$, for intercepts).

Estimated factor scores for reading at Time 1 (i.e., intercept means) did not differ across the treatment and business-as-usual groups ($\Delta_{\chi2} = 1.00$, $\Delta_{df} = 1$, $p = .317$). For the entire sample (i.e., independent of

TABLE 1
Scores for WJ-III Reading and SWAN

	Fall, Sixth Grade			Spring, Sixth Grade			Spring, Seventh Grade			Spring, Eighth Grade		
	M	SD	n	M	SD	n	M	SD	N	M	SD	n
Achieved W Score, WJ-III												
Letter-Word Identification												
Business as Usual	501.60	21.86	141	508.60	22.77	118	519.7	22.35	40	518.72	22.93	22
Treatment	500.91	21.24	278	509.60	23.75	247	521.9	21.52	158	526.11	17.54	90
Passage Comprehension												
Business as Usual	492.59	14.41	141	495.55	13.62	118	499.48	14.94	40	499.11	15.27	22
Treatment	492.29	13.27	278	496.35	13.48	247	503.59	12.90	158	506.58	10.44	90
Word Attack												
Business as Usual	499.11	18.18	141	501.38	15.75	118	505.46	14.23	40	503.11	16.76	22
Treatment	498.04	17.60	278	502.30	18.68	247	508.90	16.81	158	509.17	16.04	90
Standard Score, WJ-III												
Letter-Word Identification												
Business as Usual	91.59	12.42	141	93.07	13.44	118	95.63	14.21	40	90.36	13.75	22
Treatment	91.74	12.59	278	94.13	14.25	247	97.09	13.20	158	95.44	11.22	90
Passage Comprehension												
Business as Usual	86.45	12.49	141	86.88	12.04	118	86.95	14.37	40	83.05	15.64	22
Treatment	86.57	11.08	278	87.87	11.54	247	90.85	11.55	158	90.54	10.63	90
Word Attack												
Business as Usual	95.53	11.33	141	95.55	10.04	118	95.60	9.61	40	90.95	11.82	22
Treatment	95.18	11.03	278	96.53	12.07	247	98.59	11.94	158	95.98	11.98	90
Raw Score, SWAN												
Business as Usual				30.99	12.28	138	31.84	10.46	25	31.64	12.06	25
Treatment				32.16	10.56	287	31.54	12.2	83	36.28	10.87	94

Note. Raw scores represent the mean sum total of teacher responses on Items 1 to 9 on the SWAN. Latent means are used in the structural modeling. WJ-III = Woodcock-Johnson III Tests of Achievement; SWAN = Strengths and Weaknesses of ADHD Symptoms and Normal Behavior Scale.
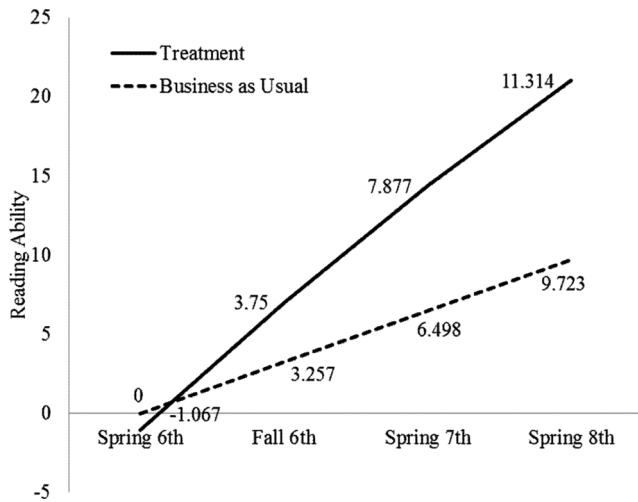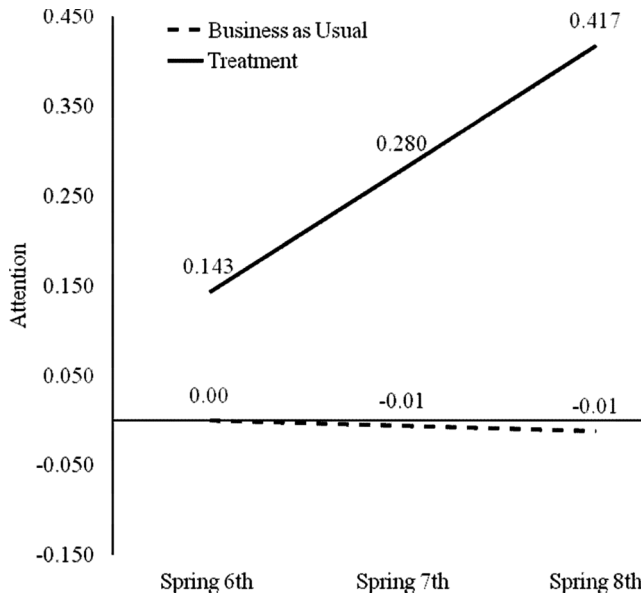
FIGURE 2   Growth trajectory of reading for the treatment and business-as-usual groups. *Note.* Values represent latent estimates of average reading. Time 1 in the business-as-usual group was fixed at 0. All other estimates (across time within the business-as-usual group and for all time points in the treatment) are relative to that baseline.

treatment status), students' reading ability improved over time ($b = 4.43$, $SE = .366$; $p < .001$), although the treatment group made greater gains than business as usual, as suggested in Figure 2. For attention, the intercept in the treatment group was .143 ($SE = .131$, $p > .05$) and the slope estimate was .14 ($SE = .06$, $p < .05$). The slope for business as usual was virtually 0 ($\beta = -0.006$, $SE = .117$, $p > .05$). Groups did not differ on intercept

FIGURE 3   Growth trajectory of attention for the treatment and business-as-usual groups. *Note.* Values represent latent estimates of average reading. Time 1 in the business-as-usual group was fixed at 0. All other estimates (across time within the business-as-usual group and for all time points in the treatment) are relative to that baseline.

($p = .27$), but they did differ on slope ($p = .02$). Figure 3 summarizes trends in the attention data for the treatment and business-as-usual groups across three time points.

To evaluate the statistical significance of treatment group differences in reading and attention and to fully specify the relationships among slope in attention, slope in reading and reading treatment as they relate to the three hypotheses, we fit the structural models in Figure 4 in Level 3 of the multiple-indicator, multilevel models. The first model (top of Figure 4) specified a direct effect of reading treatment on trends in attention. The regression coefficient was .175 ($p = .143$).
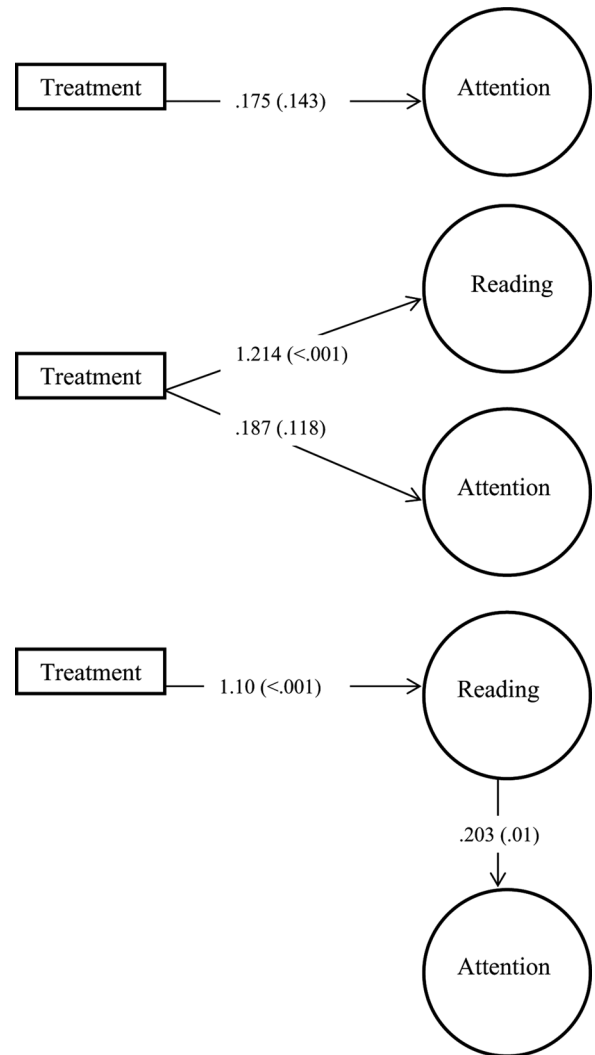

FIGURE 4   Structural relationships among reading slope, attention slope, and reading treatment. *Note.* Values represent unstandardized coefficients. *Treatment* was dummy-coded (0 = business as usual; 1 = treatment). *Reading* and *Attention* are slope estimates from multiple-indicator, multilevel latent variable growth models (see Roberts, Vaughn, Fletcher, Stuebing, & Barth, 2013). Regression coefficients are unstandardized; *p* values are in parentheses.

The middle model in Figure 4, with direct effects for reading treatment on reading slope and attention slope, indicated a significant direct effect for reading treatment on reading slope ($b = 1.214$, $p < .001$) and no direct effect on attention ($b = .187$, $p = .118$). The phenocopy hypothesis would predict a significant direct effect on attention, as discussed in the introduction. The bottom model in Figure 4 represented the correlated liabilities hypothesis, with a direct effect for reading treatment on reading performance and a relationship between improved reading and teacher-reported behavioral attention over time. The model indicated that improved reading ($b = 1.10$, $p < .001$) is associated with improved behavioral attention ($b = .203$, $p = .01$), as rated by students' English language arts teachers. There was no direct effect for treatment on attention. The three multiple-indicator, multilevel models with the Level 3 structural specifications (i.e., direct effect on attention, direct effects on reading and attention, direct effect on reading and an effect of reading on attention) all fit the data well, with RMSEA upper bounds for the 90% confidence intervals no greater than .05.

DISCUSSION

Many children with reading problems are also inattentive, regardless of whether they meet criteria for ADHD. Findings from this study indicate that intensive, response-based reading intervention over 3 years improved reading achievement *and* behavioral attention in middle school struggling readers—treatment directly affected reading, which in turn influenced attention. The results are consistent with a correlated liabilities model of comorbidity. The results do not align with the inattention-as-cause hypothesis, which predicts that reading intervention should not affect attention. The findings do not support, but do not necessarily preclude, the phenocopy hypothesis, which suggests a direct effect of reading treatment on attention.

Significance

The findings are noteworthy in several respects. The randomized design allowed for causal inferences about the effect of treatment. The 3-year time frame offered a longer-than-usual treatment interval, an acknowledged need in reading intervention research (Suggate, 2012), and the longitudinal design provided analytic flexibility for evaluating change over time and for examining interrelations among change parameters. The use of latent variable statistical models provided more precise estimates of targeted constructs than models that use manifest variables.

The age of sample participants is also unique. RD in adolescents is an increasingly prevalent topic of research (Vaughn & Fletcher, 2012). However, the role of attention has received far less scrutiny from researchers working with older struggling readers, and its causal relationship with reading-related processes in adolescent struggling readers has been, to our knowledge, unexamined. Across the age spectrum, the majority of studies that include reading and other academic outcomes do so within the context of ADHD treatment and in samples of students with a primary diagnosis of ADHD (Barbaresi, Katusic, Colligan, Weaver, & Jacobsen, 2007; Jensen et al., 2007; McGee & Share, 1988). Our study treated attention dimensionally as a means of explicating the respective roles of RD and inattention in the outcomes of older students and the potential efficacy of reading intervention in both domains. This framework is especially pertinent for older students, who may be inattentive partly because of years of struggling with reading.

Future Research and Clinical Implications

Our findings suggest that treatment for RD may influence inattention by addressing liabilities that underlie both disorders. Executive functions are a prime candidate in this respect given their importance to both reading performance *and* attention (Sesma, Mahone, Levine, Eason, & Cutting, 2009). Teaching students to plan, to monitor comprehension, and to use linguistic context appears to improve reading as well as attention, and future studies should measure these, and related executive skills, as potential mediators of reading treatment's effect on inattention. Future work might also consider innovative research designs as a means of distinguishing the phenocopy and correlated liabilities hypotheses. For example, randomizing features of a reading intervention that are thought to support executive functions would allow for internally valid comparisons of RD treatment effects. Presumably, RD treatments designed without explicit (or implicit) instruction on or support of executive functions would produce a different pattern of attention-related outcomes than RD interventions that include such components. If two treatments were comparable in all other respects, improved attention in the executive function group would represent compelling evidence in favor of the correlated liabilities view. Clinically, the findings suggest that knowledge about a student's reading ability and about his or her exposure and response to evidence-based reading intervention can inform diagnosis and treatment for inattention. Persistent inattention in adolescents with RD who have participated in and responded to evidence-based, intensive reading intervention may represent a different set of clinical options than cases where reading intervention

has been unavailable or untried or cases where reading intervention has been ineffective in addressing a student's reading disability.

Limitations

We analyzed extant data from a longitudinal study of reading intervention with struggling older readers; the design and measures were developed accordingly, creating several impositions for the present study. We were unable to specify mechanisms that may explain the changes in attention due to reading intervention, primarily because we did not measure likely variables.

We also used a sample of students who did not have clinically diagnosed comorbid ADHD and RD. We argued for a dimensional conceptualization of attention but recognize that our sample, even given a continuous characterization of ADHD, was heterogeneous. This heterogeneity was desirable because it allowed us to capture variation in attention that may be especially important for adolescents who fall below common thresholds for ADHD. RD can also be conceptualized dimensionally (Snowling & Hulme, 2012), but our primary focus in the original study was adolescent readers below a certain threshold (and a broader study population—students reading at levels above the threshold, e.g., would have greatly increased the cost of the intervention and the study). We do not know how the findings might differ in a clinical sample diagnosed with ADHD. Attrition was considerable, as discussed earlier. Although the evidence suggests that data were missing at random, overall findings of the study should be considered with this caveat in mind.

Finally, although measures like the SWAN are consistent with diagnostic criteria for ADHD, teacher ratings may reflect more than simple "inattention," including behaviors aligned with more general components of self-regulation and other executive functions. Research that uses data from a variety of measures and other cognitive assessments may help to clarify this issue.

FUNDING

REFERENCES

Archer, A. L., Gleason, M. M., & Vachon, V. (2005). *REWARDS intermediate: Multisyllabic reading strategies*. Longmont, CO: Sopris West.

Arnett, A. B., Pennington, B. F., Friend, A., Willcutt, E. G., Byrne, B., Samuelsson, S., & Olson, R. K. (2013). The SWAN captures variance at the negative and positive ends of the ADHD symptom dimension. *Journal of Attention Disorders*, *17*, 152–162. doi:10.1177/1087054711427399

Barbaresi, W. J., Katusic, S. K., Colligan, R. C., Weaver, A. L., & Jacobsen, S. J. (2007). Long-term school outcomes for children with attention-deficit/hyperactivity disorder: A population-based perspective. *Journal of Developmental & Behavioral Pediatrics*, *28*, 265–273.

Bental, B., & Tirosh, E. (2008). The effects of methylphenidate on word decoding accuracy in boys with attention-deficit/hyperactivity disorder. *Journal of Clinical Psychopharmacology*, *28*, 89–92. doi:10.1097/jcp.0b013e3181603f0e

Carroll, J. M., Maughn, B., Goodman, R., & Meltzer, H. (2005). Literacy difficulties and psychiatric disorders: Evidence for comorbidity. *Journal of Child Psychology and Psychiatry*, *46*, 524–532.

Coghill, D., & Sonuga-Barke, E. J. S. (2012). Annual research review: Categories versus dimensions in the classification and conceptualisation of child and adolescent mental disorders—Implications of recent empirical study. *Journal of Child Psychology and Psychiatry*, *53*, 469–489.

DuPaul, G. J., Jitendra, A. K., Volpe, R. J., Tresco, K. E., Lutz, J. G., Vile Junod, R. E., . . . Mannella, M. C. (2006). Consultation-based academic interventions for children with ADHD: Effects on reading and mathematics achievement. *Journal of Abnormal Child Psychology*, *34*, 635–648.

Foster, L. M., Hynd, G. W., Morgan, A. E., & Hugdahl, K. (2002). Planum temporale asymmetry and ear advantage in dichotic listening in developmental dyslexia and attention-deficit/hyperactivity disorder (ADHD). *Journal of the International Neuropsychological Society*, *8*, 22–36.

Franzak, J. K. (2006). Zoom: A review of the literature on marginalized adolescent readers, literacy theory, and policy implications. *Review of Educational Research*, *76*, 209–248.

Gayan, J., Willcutt, E. G., Fisher, S. E., Francks, C., Cardon, L. R., Olson, R. K., . . . DeFries, J. C. (2005). Bivariate linkage scan for reading disability and attention-deficit/hyperactivity disorder localizes pleiotropic loci. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *46*, 1045–1056. doi:10.1111/j.1469-7610.2005.01447.x

Hinshaw, S. (1992). Externalizing behavior problems and academic underachievement in childhood and adolescence: Causal relationships and underlying mechanisms. *Psychological Bulletin*, *111*, 127–155.

Jacobson, L. A., Ryan, M., Denckla, M. B., Mostofsky, S. H., & Mahone, E. M. (2013). Performance lapses in children with attention-deficit/hyperactivity disorder contribute to poor reading fluency. *Archives of Clinical Neuropsychology*, *28*, 672–683. doi:10.1093/arclin/act048

Jacobson, L. A., Ryan, M., Martin, R. B., Ewen, J., Mostofsky, S. H., Denckla, M. B., & Mahone, E. M. (2011). Working memory influences processing speed and reading fluency in ADHD. *Child Neuropsychology*, *17*, 209–224.

Jensen, P. S., Arnold, L. E., Swanson, J., Vitiello, B., Abikoff, H. B., Greenhill, L. L., . . . Hur, K. (2007). Follow-up of the NIMH MTA study at 36 months after randomization. *Journal of American Academy of Child & Adolescent Psychiatry*, *46*, 989–1002.

Keulers, E. H., Hendriksen, J. G., Feron, F. J., Wassenberg, R., Wuisman-Frerker, M. G., Jolles, J., & Vles, J. S. (2007).

Methylphenidate improves reading performance in children with attention deficit hyperactivity disorder and comorbid dyslexia: An unblinded clinical trial. *European Journal of Paediatric Neurology*, *11*, 21–28.

Lubke, G. H., Muthén, B. O., Moilanen, I. K., McGough, J. J., Loo, S. K., Swanson, J. M., . . . Smalley, S. L. (2007). Subtypes vs. severity differences in attention deficit hyperactivity disorder in the Northern Finnish Birth Cohort (NFBC). *Journal of the American Academy of Child & Adolescent Psychiatry*, *46*, 1584–1593.

McGee, R., & Share, D. (1988). Attention deficit disorder-hyperactivity and academic failure: Which comes first and what should be treated? *Journal of the American Academy of Child and Adolescent Psychiatry*, *27*, 318–325.

McGrath, L. M., Pennington, B. F., Shanahan, M. A., Santerre-Lemmon, L. E., Barnard, H. D., Willcutt, E. G., . . . Olson, R. K. (2011). A multiple deficit model of reading disability and attention-deficit/hyperactivity disorder: Searching for shared cognitive deficits. *Journal of Child Psychology and Psychiatry*, *52*, 547–557. doi:10.1111/j.1469-7610.2010.02346.x

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*, 525–543.

Merrell, C., & Tymms, P. B. (2001). Inattention, hyperactivity, and impulsiveness: Their impact on academic achievement and progress. *British Journal of Educational Psychology*, *71*, 43–56.

MTA Cooperative Group. (1999). A 14-month randomized clinical trial of treatment strategies for attention-deficit/hyperactivity disorder. *Archives of General Psychiatry*, *56*, 1073–86.

MTA Cooperative Group. (2004). National Institute of Mental Health Multimodal Treatment Study of ADHD follow-up: 24-month outcomes of treatment strategies for attention-deficit/hyperactivity disorder. *Pediatrics*, *113*, 754–761.

Muthén, B., & Muthén, L. (2012). *Mplus user's guide*. Los Angeles, CA: Muthén and Muthén.

Pastor, P. N., & Reuben, C. A. (2008). Diagnosed attention deficit hyperactivity disorder and learning disability: United States, 2004–2006. *Vital and Health Statistics*, *10*(237), 1–14.

Pennington, B. F. (2006). From single to multiple deficit models of developmental disorders. *Cognition*, *101*, 385–413. doi:10.1016/j.cognition.2006.04.008

Pennington, B. F., Groisser, D., & Welsh, M. C. (1993). Contrasting cognitive deficits in attention deficit hyperactivity disorder versus reading disability. *Developmental Psychology*, *29*, 511–523. doi:10.1037/0012-1649.29.3.511

Rabiner, D. L., Coie, J. D., & The Conduct Problems Prevention Research Group. (2000). Early attention problems and children's reading achievement: A longitudinal investigation. *Journal of the American Academy of Child and Adolescent Psychiatry*, *39*, 859–867.

Rabiner, D. L., & Malone, P. S. (2004). The impact of tutoring on early reading achievement for children with and without attention problems. *Journal of Abnormal Child Psychology*, *32*, 273–284. doi:10.1023/B:JACP.0000026141.20174.17

Roberts, G., Vaughn, S., Fletcher, J. M., Stuebing, K., & Barth, A. (2013). Effects of a response-based, tiered framework for intervening with struggling readers in middle school. *Reading Research Quarterly*, *48*(3), 1–18.

Sesma, H. W., Mahone, E. M., Levine, T., Eason, S. H., & Cutting, L. E. (2009). The contribution of executive skills to reading comprehension. *Child Neuropsychology*, *15*, 232–246.

Snowling, M. J., & Hulme, C. (2012). Annual research review: The nature and classification of reading disorders: A commentary on proposals for *DSM-5*. *Journal of Child Psychiatry and Psychology*, *53*, 593–607.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*, 360–406.

Suggate, S. B. (2012). Why what we teach depends on when: Grade and reading intervention modality moderate effect size. *Developmental Psychology*, *46*, 1556–1579.

Swanson, J. M., Schuck, S., Mann, M., Carlson, C., Hartman, K., Sergeant, J. A., . . . McCleary, R. (2006). *Categorical and dimensional definitions and evaluations of symptoms of ADHD: The SNAP and SWAN rating scales*. Irvine: University of California, Irvine.

Ullebo, A. K., Breivik, K., Gillberg, C., Lundervold, A. J., & Posserud, M. (2012). The factor structure of ADHD in a general population of primary school children. *Journal of Child Psychology and Psychiatry*, *53*, 927–936.

Valentine, J. C., & McHugh, C. M. (2007). The effects of attrition on baseline comparability in randomized experiments in education: A meta-analysis. *Psychological Methods*, *12*, 268–282. doi:10.1037/1082-989X.12.3.268

Vaughn, S., Cirino, P. T., Wanzek, J., Wexler, J., Fletcher, J. M., Denton, C. A., & Francis, D. J. (2010). Response to intervention for middle school students with reading difficulties: Effects of a primary and secondary intervention. *School Psychology Review*, *39*, 3–21.

Vaughn, S., & Fletcher, J. M. (2012). Response to intervention with secondary school students. *Journal of Learning Disabilities*, *45*, 244–256.

Vaughn, S., Wexler, J., Leroux, A., Roberts, G., Denton, C., Barth, A., & Fletcher, J. (2012). Effects of intensive reading intervention for eighth-grade students with persistently inadequate response to intervention. *Journal of Learning Disabilities*, *45*, 515–525.

Vaughn, S., Wexler, J., Roberts, G., Barth, A. E., Cirino, P. T., Romain, M., & Denton, C. A. (2011). The effects of individualized and standardized interventions on middle school students with reading disabilities. *Exceptional Children*, *77*, 391–407.

Volpe, R. J., DuPaul, G. J., Jitendra, A. K., & Tresco, K. E. (2009). Consultation-based academic interventions for children with attention deficit hyperactivity disorder: Effects on reading and mathematics outcomes at 1-year follow-up. *School Psychology Review*, *38*, 5–13.

Willcutt, E. G., Betjemann, R. S., McGrath, L. M., Chhabildas, N. A., Olson, R. K., DeFries, J. C., & Pennington, B. F. (2010). Etiology and neuropsychology of comorbidity between RD and ADHD: The case for multiple-deficit models. *Cortex*, *46*, 1345–1361.

Willcutt, E. G., & Pennington, B. F. (2000). Comorbidity of reading disability and attention-deficit/hyperactivity disorder: Differences by gender and subtype. *Journal of Learning Disabilities*, *33*, 179–191.

Willcutt, E. G., Pennington, B. F., Duncan, L., Smith, S. D., Keenan, J. M., Wadsworth, S., . . . Olson, R. K. (2010). Understanding the complex etiologies of developmental disorders: Behavioral and molecular genetic approaches. *Journal of Developmental & Behavioral Pediatrics*, *31*, 533–544.

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

Wu, A. D., Li, Z., & Zumbo, B. D. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. *Practical Assessment, Research and Evaluation*, *12*(3), 1–26.