Article

Hammill Institute on Disabilities

Journal of Learning Disabilities
2017, Vol. 50(2) 115–127
© Hammill Institute on Disabilities 2015
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0022219415617163
journaloflearningdisabilities.sagepub.com
SAGE

Impact of Intensive Summer Reading Intervention for Children With Reading Disabilities and Difficulties in Early Elementary School

Joanna A. Christodoulou, EdD[1,2,3], Abigail Cyr, BA[1], Jack Murtagh, BA[1], Patricia Chang, MA[1], Jiayi Lin, BS[1], Anthony J. Guarino, PhD[2], Pamela Hook, PhD[2], and John D. E. Gabrieli, PhD[1,3]

Abstract

Efficacy of an intensive reading intervention implemented during the nonacademic summer was evaluated in children with reading disabilities or difficulties (RD). Students (ages 6–9) were randomly assigned to receive Lindamood-Bell's *Seeing Stars* program (n = 23) as an intervention or to a waiting-list control group (n = 24). Analysis of pre- and posttesting revealed significant interactions in favor of the intervention group for untimed word and pseudoword reading, timed pseudoword reading, oral reading fluency, and symbol imagery. The interactions mostly reflected (a) significant declines in the nonintervention group from pre- to posttesting, and (2) no decline in the intervention group. The current study offers direct evidence for widening differences in reading abilities between students with RD who do and do not receive intensive summer reading instruction. Intervention implications for RD children are discussed, especially in relation to the relevance of summer intervention to prevent further decline in struggling early readers.

Keywords

elementary, treatment, dyslexia, early identification/intervention

Specific learning disability (SLD) is the most prevalent disability category in the United States, occurring in 5% of enrolled students (U.S. Department of Education, National Center for Education Statistics, 2013). An estimated 80% of children with SLD present with difficulties in reading (Lerner, 1989; Shaywitz, Morris, & Shaywitz, 2008). Commonly, these challenges extend into adulthood (e.g., Bruck, 1992). Given the prevalence and long-term impact of reading challenges, a focus on improving outcomes for students with reading disabilities and difficulties (RD) early is a major focus in education.

Early intervention is the most effective strategy for improving short- and long-term outcomes for children with RD (Gabrieli, 2009; Gilbert et al., 2013; Schatschneider & Torgesen, 2004). However, challenges in choosing intervention programs that will be effective and delays in identifying at-risk readers (typically occurring in third grade or later in U.S. schools) are limitations to implementing early intervention (Snowling, 2013). Furthermore, insufficient research is available regarding the efficacy of available interventions for at-risk readers in early elementary school (Denton, 2012). Since the integration of response to intervention (RTI) into U.S. federal law (Individuals with Disabilities Education Act, 2004), school districts have the option of implementing this preventative framework. RTI is intended to improve the developmental trajectory of students at risk for academic difficulty via early intervention using effective research-based interventions in tiers of intensity and individualization (Denton, 2012). Schools implementing RTI carry the burden of making important program selection decisions in the context of limited time, finances, and staff resources, thus making more pressing the need for additional research on the efficacy of reading intervention programs. In the present study, we examined the efficacy of a particular reading intervention, *Seeing Stars* (Bell, 2007), administered over the summer to young children with RD.

[1]McGovern Institute for Brain Research and Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, USA
[2]MGH Institute of Health Professions, Department of Communication Sciences and Disorders, Boston, MA, USA
[3]Harvard Graduate School of Education, Cambridge, MA, USA

Corresponding Author:
Joanna A. Christodoulou, Department of Communication Sciences and Disorders, MGH Institute of Health Professions, 36 First Avenue, Boston, MA 02129, USA.
Email: jchristodoulou@mghihp.edu

Importance of Early and Intensive Reading Intervention

Early and intensive intervention yields the strongest effects for promoting reading development in RD children (Al Otaiba & Fuchs, 2002; Fletcher & Foorman, 1994; Hanselman & Borman, 2013; Kennedy, Birman, & Demaline, 1986; H. L. Swanson, 1999; Vaughn, Denton, & Fletcher, 2010; Wanzek & Vaughn, 2007). Reviews of reading intervention efficacy for children in early primary school indicate the potential for positive outcomes, particularly following small-group instruction (Wanzek & Vaughn, 2007). However, more research is needed to understand the conditions under which students with RD can benefit the most from intervention, given the persistence of difficulties for young readers (Torgesen, 2000) and legislative efforts to ensure students are competent readers by third grade (National Conference of State Legislatures, 2014).

In the RTI framework, strong evidence supports the efficacy of intensive reading instruction (Vaughn et al., 2010) delivered for 20 to 40 min for 3 to 5 days per week, with instructional time allotment increasing from kindergarten to early primary grades (Gersten et al., 2009). Without formal intervention, students who struggle to develop proficiency in reading during first grade show a .88 probability of demonstrating poor reading skills in fourth grade (Juel, 1988). When students with RD receive intervention in later years, the interventions tend to stabilize reading deficits rather than close reading achievement gaps (Torgesen, 2005). Indeed, beyond the first three years of elementary school, students with RD have significant difficulty reaching the reading fluency levels of their peers (Torgesen, Rashotte, & Alexander, 2001).

Even in studies reporting overall efficacy of interventions for students with RD, there is substantial variation in treatment response across students. Across six early intervention studies, a wide range of 8% to 44% of students remained below the 30th percentile on word reading measures despite average gains for the treatment groups at posttesting (Torgesen, 2004). These studies of effective intervention occurred with (a) early readers in kindergarten, first grade, or second grade; (b) intensive instruction ranging from 35 to 340 hr; (c) small-group or one-on-one instruction; and (d) phonologically-based programs.

Seeing Stars *Program*

Seeing Stars: Symbol Imagery for Fluency, Orthography, Sight Words, and Spelling is one of the primary reading intervention programs offered by Lindamood-Bell. The program was developed as a multisensory approach to teach phonological and orthographic awareness, sight word recognition, and comprehension in a developmental or remedial capacity. *Seeing Stars* emerged as a complement to the *Lindamood Phoneme Sequencing* (LiPS) program, which focuses on the articulatory, phonological, and auditory components of language and their connection to written language. The majority of published intervention studies have used such programs that are theoretically driven by the phonological deficit hypothesis for students with RD (A. W. Alexander & Slinger-Constant, 2004), with the goal of enhancing phonological abilities to support subsequent orthographic processing and reading. In contrast, *Seeing Stars* has a primary instructional focus on improving reading via orthographic and visual processing training, and subsequent phonological training.

The *Seeing Stars* curriculum is inspired by evidence that visual and imagery processes can support reading-relevant processes (Bell, 2007; Kosslyn, 1976; Linden & Wittrock, 1981; Long, Winograd, & Bridge, 1989). The dual coding theory underlies the *Seeing Stars* program, and postulates that cognition relies on two systems when readers are engaged with text: a system dedicated to language and another system specialized in nonverbal information processing (i.e., visual imagery; Sadoski & Paivio, 2013). The program aims to harness the potential of both language and imagery of printed text to facilitate fluent decoding and spelling, and ultimately comprehension. The cost of the program depends on several factors, including the scale of implementation, need for teacher training, and how extensively the program will be used (i.e., which level of RTI). The small-group and intensive duration parallel the instructional formats typical of effective phonologically-centered interventions for young children. For the current study, reading instruction was offered free of charge.

The curriculum for *Seeing Stars* calls for explicit instruction in visualizing units of increasing size: letters, syllables, words, connected text. Students are introduced to this main curricular goal at the beginning of their instruction via the analogy that just as the stars are parts of the sky, letters are parts of words. The goal of building mental visual representations of written language units follows a consistent lesson format. First, the teacher presents the task and goal of the session, then carries out the lesson, and ends with a lesson summary detailing the main elements. Students learn to visualize letters and write them in the air, both from seeing the letters in print and from hearing the sounds they make. During lessons for syllable imagery, students learn to visualize with four main activities: conjure the mental image of the target syllable; identify a specific letter among those in the syllable; report the letters in the cluster backward; and report the letters in the cluster after manipulating the letter order. The program has an explicit focus on irregular syllable patterns as well. Instruction at the whole-word level explicitly teaches both reading and spelling skills. The strategies for parsing words into more accessible units span single syllable to multisyllabic words, with the visualization focused on the discrete orthographic patterns. Students progress from visualizing orthography to learning how to visualize the semantic information they read in connected text as a strategy to monitor for meaning. Students are asked to share the mental imagery that they conjure as they read each sentence, and reread when their mental imagery does not make sense. In this way, reading comprehension is an explicit goal in the program by training students to visualize not only orthographic units, but semantic content of written text as well.

Empirical work evaluating the efficacy of *Seeing Stars* is limited yet promising. In a study investigating the impact of this reading intervention on brain structure, 11 children (ages 7–11; M = 9.1 years, SD = 1.3) with developmental dyslexia who completed an unspecified number of instructional hours over 8 weeks showed standard score gains for untimed single real word and pseudoword reading, rapid letter naming, reading comprehension, phonemic awareness, and symbol imagery (Krafnick, Flowers, Napoliello, & Eden, 2011). Although this study did not include an independent control group, these students appeared to maintain their gains 8 weeks later. As part of a study investigating the brain basis of visual processing in dyslexia, 22 students (ages 7–12; M = 9.6 years, SD = 1.4) completed 8 weeks of *Seeing Stars* and 8 weeks of either math intervention or no intervention in a randomized order, and showed statistically significant improvement for untimed single word reading, untimed pseudoword reading, and phonological awareness during the reading intervention period but not during the combined control periods (Olulade, Napoliello, & Eden, 2013). Remaining studies relied on a single case design (Rauschecker et al., 2009) or used *Seeing Stars* in combination with other reading programs (Eden et al., 2004; Sadoski & Wilson, 2006). The *Seeing Stars* program was chosen for use in this study to expand the information available on intervention efficacy and to examine whether an approach focusing primarily on orthographic training and secondarily on phonological training can be effective with students who have RD in early elementary school. Furthermore, *Seeing Stars* is recognized as a program commonly use in clinical and education settings (International Dyslexia Association, 2007).

The current randomized controlled trial investigated the impact of the *Seeing Stars* intervention for young children with RD by asking whether an intensive, small-group summer program of *Seeing Stars* intervention for children ages 6 to 9, relative to a waiting control group, has a positive effect on targeted outcomes in symbol imagery, single word and pseudoword reading, and oral reading fluency.

Method

Children were assigned randomly to an intervention group (n = 23) or a waiting control group (n = 24) that was offered equivalent intervention access after completion of the study. Participants completed assessments of reading and related skills before and after the intervention period. Intervention group participants received a reading intervention program in small groups (3–5 children per instructor) in an intensive delivery model.

Participants

Participants (6–9 years old) were recruited via community outreach efforts including contacting local area school staff, posting to parent and community online groups, and advertising on media outlets. All enrollees were native English speakers and parent questionnaire responses indicated absence of neurological or psychiatric impairments, or associated medications. Participants completed behavioral testing with trained researchers at the Massachusetts Institute of Technology (MIT). Informed consent from parents and assent from children for participation in the study, approved by the MIT Institutional Review Board, were obtained.

Participants attended three testing sessions: screening, preintervention, and postintervention. Children were initially considered for inclusion based on the presence of developmental reading difficulties and/or a current diagnosis of RD. Parent questionnaire responses indicated that 26 children held a diagnosis of a language-based learning disability (LBLD), with 13 children in the nonintervention group and 12 children in the intervention group. In addition, participants were required to score below the 25th percentile on at least two of the reading or reading subskill measures from the screening battery (see below for measures). All participants were required to demonstrate nonverbal cognitive performance (*Kaufman Brief Intelligence Test–2nd Edition*; KBIT; Kaufman & Kaufman, 2004) at or above the 16th percentile. One participant from each group, both reassigned after original group designation due to family request related to scheduling conflicts and by study staff reassignment to balance sample sizes, respectively, was omitted from the final analysis due to violation of the randomization procedure.

Following randomized assignment to the summer reading intervention group (n = 23) or the nonintervention group (n = 24), there were no significant group differences on screening measures, LBLD diagnosis, age, grade, gender, or socioeconomic status (SES; Table 1).

At the start of the study, nine children carried attention-deficit/hyperactivity disorder (ADHD) diagnoses given by external clinicians. Among this group, seven children (five on medication daily) were in the intervention group and two children (one on medication daily) were in the nonintervention group. Statistical analyses included ADHD diagnostic status as a covariate to control the potential contribution of this characteristic on group effects.

Table 1. Participant Demographic Characteristics by Group.

	Intervention group (n = 23)	Nonintervention group (n = 24)	All participants (N = 47)
Age (years)	M = 7.67 ± 0.53 $n_{6\ years}$ = 1 $n_{7\ years}$ = 15 $n_{8\ years}$ = 7 $n_{9\ years}$ = 0	M = 7.76 ± 0.69 $n_{6\ years}$ = 2 $n_{7\ years}$ = 13 $n_{8\ years}$ = 7 $n_{9\ years}$ = 2	M = 7.72 ± 0.61 White: n = 38, M_{age} = 7.77 ± 0.61 Black/African American: n = 3, M_{age} = 7.92 ± 1.01 American Indian/Alaska Native: n = 2, M_{age} = 7.17 ± 0.24 Hispanic: n = 1, age = 7.08 Multiracial: n = 1, age = 7.00 Other: n = 2, M_{age} = 7.67 ± 0.12
Grade	M = 1.43 ± 0.51	M = 1.46 ± 0.51	M = 1.45 ± 0.51
Gender	Male: n = 14 Female: n = 9	Male: n = 16 Female: n = 8	Male: n = 30 M_{age} = 7.83 ± 0.64 M_{SES} = 50.58 ± 9.97 Female, n = 17 M_{age} = 7.52 ± 0.53 M_{SES} = 51.06 ± 11.92
SES	M = 50.83 ± 10.13	M = 50.70 ± 11.28	M = 50.76 ± 10.62 White: M_{SES} = 51.47 ± 11.19 Black/African American: M_{SES} = 48.17 ± 9.70 American Indian/Alaska Native: M_{SES} = 50.75 ± 1.06 Hispanic: n = 1, SES = 35.00 Multiracial: n = 1, SES = 53.00 Other: n = 1, SES = 48.00

Note. Values are M ± SD. Socioeconomic status (SES) is based on the *Barratt Simplified Measure of Social Status*.

Measures

Participants completed the measures described below. In addition, parents completed questionnaires regarding developmental history, including diagnostic history for reading and attention disorders. Test descriptions include testing session (screening, preintervention, postintervention) and repeated administration information. Pre- and posttest scores are reported for targeted outcome measures for timed and untimed word and pseudoword reading, oral reading fluency, and symbol imagery. Remaining pretest measures are reported for sample characterization purposes. All participants completed posttesting within 3 weeks of study completion. All testing was completed before formal schooling began, except for one intervention group participant (tested after 2 days of school instruction) and four nonintervention group participants (M = 14, SD = 12.49; range = 2–28 days).

Measures administered at pretest for sample characterization

***Cognitive skills, executive function, attention, behavior*.** Nonverbal cognitive ability was measured with the Matrices subtest of the KBIT (Kaufman & Kaufman, 2004). Participants identified which picture in a series fits into an increasingly complex matrix. This measure was completed during screening only, because these scores were not expected to change as a result of participation. Split-half reliability is .87 to .89 for ages 6 to 9.

Parent surveys. A parent of each participant was asked to complete questionnaires one time to index aspects of behavior (*Behavior Assessment System for Children, 2nd Edition*; BASC; Reynolds & Kamphaus, 2004), attention (*Conners' Rating Scales–Revised*; CRS; Conners, 1997), and executive functions (*Behavior Rating Inventory of Executive Function*; BRIEF; Gioia, Isquith, Guy, & Kenworthy, 2000). Internal reliability across questionnaires ranges from .72 to .98.

Oral language. The *Peabody Picture Vocabulary Test–4* (PPVT; Dunn & Dunn, 2007), a receptive vocabulary task, required participants to indicate by pointing which of four drawings corresponded to a word read aloud by the examiner. Cronbach's alpha ranges from .94 to .96. PPVT was administered during preintervention only because these scores were not expected to change as a result of study participation.

The Listening Comprehension subtest of the *Oral and Written Language Scales–II* (OWL LC; Carrow-Woolfolk, 2011) required participants to choose one of four pictures that corresponded to a sentence read aloud by the examiner. This measure was completed during preintervention. Split-half reliability ranges from .97 to .98.

Phonological awareness. Phonological awareness was measured with Elision (EL) subtest from the *Comprehensive Test* *of Phonological Processing* (CTOPP; Wagner, Torgesen, & Rashotte, 1999), which required participants to remove sounds from words read aloud by the examiner. Screening scores are reported. Reliability indices are .89 for internal consistency, .96 to .99 for interscorer, and .82 for test–retest.

Rapid automatized naming. Rapid naming skills were measured using the Letters (L) subtest of the *Rapid Automatized Naming and Rapid Alternating Stimulus Tests* (RAN/RAS; Wolf & Denckla, 2005), for which participants were asked to name letters accurately and quickly. Screening scores are reported. Reliability indices are .98 for interrater and .90 for test–retest.

Text comprehension. Reading comprehension was assessed using the Passage Comprehension (PC) subtest of the *Woodcock Reading Mastery Test–III* (WRMT; Woodcock, 2011), which is a cloze reading task for which readers provide a word missing from the sentence. Preintervention scores are reported. Split-half reliability ranges from .84 to .93.

Measures administered at pretest and posttest

Word (orthographic) recognition or decoding. For the *Symbol Imagery Test* (SIT; Bell, 2010), participants viewed cards with letters or pseudowords for 2 to 7 s and were then asked to report what they were shown. This measure was completed during preintervention and postintervention. Cronbach's alpha values range from .86 to .88. Test–retest reliability is .95.

Sight-word reading was measured with the Word Identification (WI) subtest of the WRMT, for which participants read a list of increasingly challenging words. Decoding was measured with the Word Attack (WA) subtest for which participants read increasingly challenging pseudowords. These measures were completed during preintervention and postintervention using alternate forms. Split-half reliability ranges from .92 to .98 for WI and .91 to .95 for WA. Alternate form reliability ranges from .80 to .93 for WI and from .72 to .74 for WA.

Timed reading ability was indexed by accuracy for reading real words (Sight Word Efficiency; SWE) and pseudowords (Phonemic Decoding Efficiency; PDE) within 45 s on the *Test of Word Reading Efficiency–2* (TOWRE; Torgesen, Wagner, & Rashotte, 2012). These measures were completed during preintervention and postintervention using alternate forms. Reliability estimates across subtests are .91 to .92 for alternate forms, .90 to .91 for test–retest, and .99 for interscorer.

Connected text fluency. For the Oral Reading Fluency (ORF) subtest of the *Dynamic Indicators of Basic Early Literacy Skills–6* (DIBELS; Good & Kaminski, 2002), participants read three connected-text passages for 1 min each. Scores were calculated based on number of words read accurately. This measure was completed during screening and postintervention using Benchmark 3 forms for both time points. Test–retest reliability indices range from .92 to .97 (Tindal, Marston, & Deno, 1983).

Intervention

The intervention group received *Seeing Stars* in an intensive delivery model during 6 weeks of the summer. This group did not receive other Lindamood-Bell program components; *Seeing Stars* was provided in isolation. Participants attended the program 5 days per week for 4 hr each day, totaling 100 to 120 hr of instruction. The sessions consisted of curricular activities with a short break (5–10 min per hour). Instruction occurred in groups of 3 to 5 children assigned by reading level and was delivered by teachers trained by and working for Lindamood-Bell who alternated classrooms each hour daily, which resulted in comparable exposure to each instructor and minimized or eliminated potential teacher effects.

Efforts to maintain treatment fidelity included teacher training and teacher monitoring. First, teachers in the study were staff members of and trained by the Lindamood-Bell program. The company offers, on average, 80 hr of formal instruction followed by about 80 hr of clinical observation in the teacher training process. Second, during the study, there were three levels of monitoring in class implementation efficacy offered by a project consultant manager, a regional director, and a corporate director of instruction. Observations informed recommendations from these supervisors to modify curriculum implementation. The attrition rate was zero.

Elective Activities

Parents of participants were asked to report additional reading remediation their child received during the study period. Of the 23 intervention participants, 20 (via parents) responded to the survey; only 1 reported additional remediation of 2 hr per week for 6 weeks at a local learning center. Responses from 23 out of the 24 nonintervention participants indicated that 7 children had private tutors averaging 1.6 hr per week for 6.2 weeks; 5 attended a commercial reading program or summer school averaging 10 hr per week for 6.6 weeks; and 11 received no reading instruction. In contrast to the intensive 100 to 120 hr of instruction for the intervention group, the nonintervention group reading-related instruction ranged from 0 to 80 hr of total instruction (M = 14.2, SD = 23.2). As reported below, correlation analyses between hours dedicated to summer reading instruction and outcomes on reading outcome measures were conducted to evaluate the potential impact within groups and among all students.

Table 2. Pretest Behavioral Scores for Characterization by Group.

Measure	Intervention group (n = 23; $M \pm SD$)	Nonintervention group (n = 24; $M \pm SD$)
KBIT Matrices	107.17 ± 11.89	105.75 ± 14.98
PPVT	112.17 ± 10.85	109.08 ± 10.75
OWL Listening Comprehension	104.83 ± 9.33	104.75 ± 9.10
CTOPP Elision	8.05 ± 1.53	8.43 ± 2.29
RAN Letters	90.32 ± 9.13	93.13 ± 11.10
WRMT Passage Comprehension	90.96 ± 11.49	90.79 ± 7.47
BASC Externalizing Problems	47.36 ± 7.72	45.55 ± 5.45
BASC Internalizing Problems	48.14 ± 8.04	47.95 ± 8.95
BASC Behavioral Symptoms	50.21 ± 6.02	48.50 ± 6.96
BASC Adaptive Skills	50.23 ± 7.63	50.90 ± 7.66
CRS Global Index	52.29 ± 8.37	53.32 ± 9.36
CRS *DSM-IV* Total	54.62 ± 9.12	54.18 ± 8.68
BRIEF Behavioral Regulation	50.65 ± 9.76	51.29 ± 7.90
BRIEF Metacognition	58.43 ± 11.66	54.54 ± 8.46
BRIEF Global Executive Composite	55.96 ± 10.98	53.58 ± 8.22

Note. BASC = *Behavior Assessment System for Children, 2nd Edition*; BRIEF = *Behavior Rating Inventory of Executive Function*; CRS = *Conners' Rating Scales–Revised*; CTOPP = *Comprehensive Test of Phonological Processing*; KBIT = *Kaufman Brief Intelligence Test–2nd Edition*; OWL = *Oral and Written Language Scales–II*; PPVT = *Peabody Picture Vocabulary Test–4*; RAN = *Rapid Automatized Naming*; WRMT = *Woodcock Reading Mastery Test–III*. Scores are age-based standard scores for KBIT, PPVT, RAN, WRMT (M = 100, SD = 15); CTOPP (M = 10, SD = 3); BASC, CRS, and BRIEF (M = 50, SD = 10).

Analysis

To assess if the groups (intervention or nonintervention) were comparable at baseline, a Hotelling's T^2 or two-group between-subjects multivariate analysis of variance was conducted on all pretest measures. Age-based standardized scores were used for all analyses, except for DIBELS for which raw scores were used.

To examine intervention effects across outcome measures (SIT, WRMT WI, WRMT WA, TOWRE SWE, TOWRE PDE, DIBELS ORF), a MANCOVA was conducted on the six posttest scores with the pretest scores as the covariates. The MANCOVA was selected because the dependent variables were statistically and conceptually related, and it is more statistically powerful than analysis of gain scores if the groups did not differ significantly at pretest (Weinfurt, 2000). The multiple dependent variables reflect more accurately the reality of the phenomenon under study. In addition, multivariate analysis limits the inflation of Type I error rates (Meyers, Gamst, & Guarino, 2012). The evaluations of assumptions were all met.

Separate MANCOVAs were conducted with the ADHD diagnostic status as the covariate to identify whether initial trends were maintained. Given the importance of acknowledging the role of ADHD status among students with RD, this additional analysis offers a deeper evaluation of the relation between variables; however, outcomes regardless of ADHD status are reported as the main findings to ensure comparability to the majority of published research on reading efficacy.

To examine individual variability, change scores were calculated (posttest minus pretest) for standardized word reading outcomes as a composite (WRMT WI, WRMT WA, TOWRE SWE, TOWRE PDE) to determine how many children in each group showed no change or positive change of scores, as compared to decline indicated by negative change. In addition, correlation analysis was used to identify the relation between these change scores and summer reading instructional hours. A final analysis evaluated the proportion of children who met stringent research-based criteria for RD at the end of the study.

Results

Pretest: Screening and Preintervention Comparison of Groups

Intervention and nonintervention groups did not differ on demographic characteristics, including age, $t(45) = 0.50$, $p = .62$, gender, $\chi^2(1, N = 47) = 0.17$, $p = .68$, or SES, as indicated by matched scores on the *Barratt Simplified Measure of Social Status* (Barratt, 2006), $t(40) = 0.04$, $p = .97$ (Table 1).

Participants from the intervention and nonintervention groups exhibited no significant differences across reading and related measures before intervention began ($\alpha = .001$). On average, the groups performed within normal limits for their age and did not show significant group differences on nonverbal cognitive abilities (KBIT); expressive language (PPVT); oral language comprehension (OWL LC); phonological awareness (CTOPP EL); rapid letter naming (RAN L); passage comprehension (WRMT PC); BASC indices of externalizing problems, internalizing problems, behavioral symptoms, or adaptive skills; CRS indices of attention; or BRIEF indices of executive function (Table 2).

Table 3. Pretest and Posttest Behavioral Scores for Intervention Measures by Group.

	Intervention group (n = 23)		Nonintervention group (n = 24)				
	Time 1	Time 2	Time 1	Time 2			
Measure	$M \pm SD$	$M \pm SD$	$M \pm SD$	$M \pm SD$	F	p	η_p^2
SIT	89.57 ± 8.88	99.32 ± 9.16	87.57 ± 8.71	84.21 ± 9.53	20.52	<.001	.406
WRMT WI	85.04 ± 11.59	85.65 ± 9.90	85.29 ± 8.50	78.58 ± 9.10	17.47	<.001	.368
WRMT WA	87.43 ± 10.22	88.09 ± 9.06	87.04 ± 11.96	79.04 ± 9.39	9.96	.004	.249
TOWRE SWE	84.22 ± 11.41	80.04 ± 12.50	84.33 ± 10.53	78.83 ± 11.15	0.88	.355	.029
TOWRE PDE	79.32 ± 9.32	80.48 ± 9.44	80.26 ± 10.43	72.86 ± 9.60	7.92	.009	.209
DIBELS ORF	30.31 ± 25.62	40.48 ± 29.11	35.61 ± 25.77	35.88 ± 22.36	4.50	.042	.130

Note. DIBELS = *Dynamic Indicators of Basic Early Literacy Skills–6*; ORF = Oral Reading Fluency; PDE = Phonemic Decoding Efficiency; SIT = *Symbol Imagery Test*; SWE = Sight Word Efficiency; TOWRE = *Test of Word Reading Efficiency–2*; WA = Word Attack; WI = Word Identification; WRMT = *Woodcock Reading Mastery Test–III*. Age-based standard scores (M = 100, SD = 15) are reported for assessments except DIBELS ORF (raw scores).

In addition, reading and reading-related skills did not differ between groups at pretest (Table 3), based on MANOVA results indicating that there were no univariate or multivariate outliers at $\alpha = .001$. Assumptions of normality, homogeneity of variance-covariance matrices, linearity, and multicollinearity were all met. Using the Wilks's lambda criterion, the composite dependent variate was not significantly affected by group membership, $F(6, 34) = 0.43$, $p = .85$.

Group Comparison: Intervention Effects

Descriptive statistics and MANCOVA results for each posttest adjusting for pretest scores are reported in Table 3. There was a significant difference between groups as indicated by the Wilks's lambda criterion, $F(6, 25) = 6.17$, $p < .001$, $\eta^2 = .597$ (a very large effect size). Univariate analyses were conducted on each dependent measure separately to determine the locus of the statistically significant multivariate group effect. To control for Type I errors, Bonferroni adjustment was employed. Posttest scores were significantly higher for the intervention group than the nonintervention group on SIT, WRMT WI, WRMT WA, TOWRE PDE, and DIBELS ORF (Figure 1). There was no significant group difference on posttest TOWRE SWE scores.

Three trends appeared to be underlying significant group differences. For untimed real-word reading (WRMT WI), untimed pseudoword reading (WRMT WA), and timed pseudoword reading (TOWRE PDE), the intervention group showed no statistically significant changes in scores while the nonintervention group declined significantly. On oral reading fluency (DIBELS ORF), the intervention group improved while the nonintervention group showed no statistically significant change. On symbol imagery (SIT), the intervention group improved while the nonintervention group declined.

MANCOVAs were recalculated to include ADHD diagnostic status as a covariate to identify whether initial trends were maintained. Results were consistent for all measures regarding statistical significance.

Individual Differences in Intervention Response

Reading improvement relative to pretest scores was achieved by 50% of the participants in the intervention group ($M = -0.81$, $SD = 6.39$) as indicated by composite change scores (posttest minus pretest) greater than zero based on standardized reading outcome measures (WRMT WI, WRMT WA, TOWRE SWE, TOWRE PDE). This test excluded DIBELS, which lacks standardized scores, and SIT, which is closely related to the intervention program and could inflate this index. Among students in the nonintervention group ($M = -7.18$, $SD = 4.48$), only one participant showed a mean gain on these reading outcome measures based on the composite change score. Score declines were noted in 50% of the intervention group and in 95% of the nonintervention group. Using the benchmark of positive change scores, analysis indicated that the proportion of students in each group that met this benchmark differed significantly, $\chi^2(1, N = 43) = 10.93$, $p = .001$.

We calculated effect sizes (Cohen's d) for the outcome measures to examine the magnitude of the differences between the intervention and nonintervention groups. Calculations were based on post- minus pretest standardized scores (i.e., change scores). Untimed word reading (WRMT WI) yielded an effect size of 0.96 (intervention: $M = 0.61$, $SD = 7.90$; nonintervention: $M = -6.70$; $SD = 7.30$). Untimed pseudoword reading (WRMT WA) yielded an effect size of 0.87 (intervention: $M = 0.65$, $SD = 11.61$; nonintervention: $M = -8.00$; $SD = 7.95$). Timed word reading (TOWRE SWE) yielded an effect size of 0.19 (intervention: $M = -4.17$, $SD = 7.14$; nonintervention: $M = -5.50$; $SD = 7.03$). Timed pseudoword reading (TOWRE

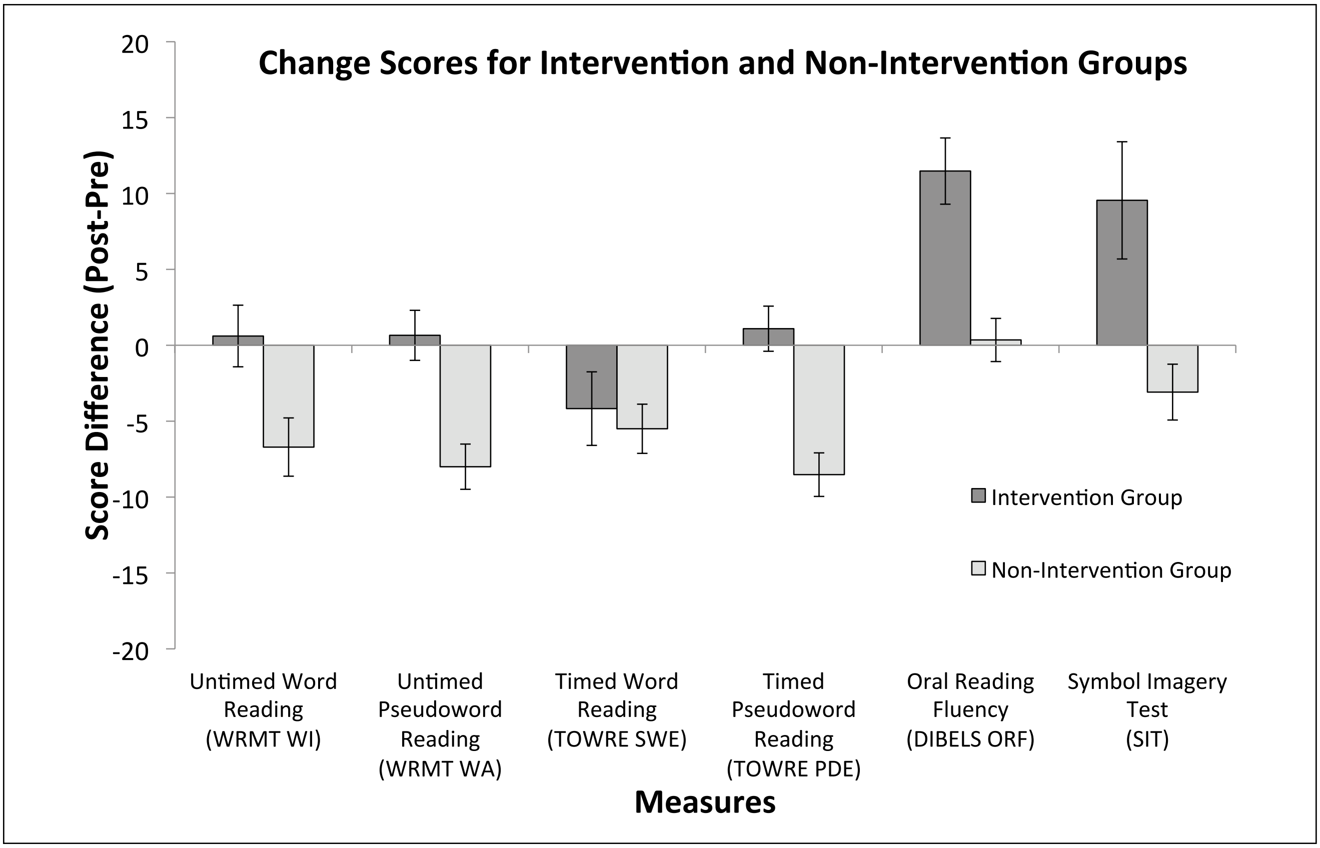

Figure 1. Change scores by group for intervention measures.
Note. DIBELS = *Dynamic Indicators of Basic Early Literacy Skills–6*; ORF = Oral Reading Fluency. Standard error bars included. Standard score differences are reported for all standardized measures; raw score differences are reported for DIBELS ORF.

PDE) yielded an effect size of 1.08 (intervention: M = 1.09, SD = 10.47; nonintervention: M = −8.52; SD = 6.96). Oral reading fluency (DIBELS ORF) yielded an effect size of 0.76 (intervention: M = 11.48, SD = 18.52; nonintervention: M = 0.35; SD = 9.02). The SIT yielded an effect size of 1.32 (intervention: M = 9.55, SD = 9.73; nonintervention: M = −3.09; SD = 9.40).

Although the nonintervention group did not receive instruction delivered through the study, 58% of children in this group received formal summer reading instruction according to parent reports. To determine the potential influence of this instruction on reading outcomes, a correlation between number of hours participating in summer reading instruction and a composite score of the four standardized word reading outcome measure change scores (posttest minus pretest) was used, which indicated a nonsignificant relationship (r = .22, p = .37).

Furthermore, we evaluated the proportion of children in each group who met stringent research-based criteria for RD, as indicated by a standard score below the 25th percentile on at least two of four single word reading measures (WRMT WI, WA; TOWRE SWE, PDE). Although the majority of children retained their RD status from pre to post in both groups, one child (from the intervention group) no longer met criteria, while several newly met RD criteria (2 from intervention group; 5 from nonintervention group) at the end of the study.

Discussion

This study investigated the impact of an intensive reading intervention program, *Seeing Stars*, implemented in the nonacademic summer for early readers with RD. Although intervention and control groups began with similar pretest scores, significant benefits of intervention were observed on the majority of word reading, pseudoword reading, and text reading outcome measures, with children who received the intervention showing relative gains (through maintenance or improvement) across the summer and children in the control group exhibiting relative stagnation or score declines. Both groups declined significantly on speeded real word reading. For the symbol imagery measure, which was most aligned with the content of the intervention, students receiving the intervention improved significantly, whereas students in the control group declined significantly. Group differences in reading outcomes were not attributable to SES or baseline differences in language ability, phonological awareness, rapid naming, executive function, behavior, or attention because there were no differences between groups on any of these measures. Overall, there were statistically significant benefits for students receiving the *Seeing Stars* intervention relative to students in the nonintervention control group.

Three aspects of the significant benefits are noteworthy. First, although the intervention group improved significantly on two measures, three of the group differences resulted from no pre to post changes in the intervention group versus significant declines in the nonintervention group. Second, within the intervention group there was diversity of benefits, with half of the students showing some gain and half showing no gain. Future investigations with larger samples can examine factors that contribute to efficacy at the individual level. This diversity was in striking contrast to the nonintervention group in which 95% of students failed to exhibit a pre to post composite gain.

Seeing Stars *Intervention Efficacy*

Although the current study extends our knowledge regarding the efficacy of the *Seeing Stars* curriculum for students in early elementary school with RD, two aspects of the findings were unexpected. First, it was surprising that the intervention had no influence on timed single real word reading given significant influences on reading of words and pseudowords in timed and untimed contexts, as well as oral reading fluency. However, a possible explanation is thatwthe group receiving intervention was learning strategies that directed attention to letter-sound correspondence patterns. Thus, as students implemented these pattern recognition strategies in words that they had previously been reading by sight, reading speed may have slowed compared to their initial rates. The relatively short duration of the intervention (6 weeks) may have been a factor as well given that over a longer period of time students can become more adept at applying their knowledge of letter patterns with automaticity. Second, more broadly, it was not anticipated that for several measures the benefit of the intervention would be expressed as the absence of a decline that occurred in the group assigned to intervention. However, the impact of summer months on readers with difficulties or disabilities can be detrimental for reading outcomes. Notably, students with RD do not typically make a year of progress in an academic year, so maintaining standard scores during the summer months (when they may be most vulnerable to reading skill regression) should be considered an index of progress even though students with RD did not, on average, close the gap separating them from their typically reading peers.

Despite these unexpected aspects of the results, the intervention effects were substantial. The current study showed the magnitude of the differences in the means favored the intervention over the nonintervention group ($d = 0.19–1.32$) across measures. Focusing on the standardized word and pseudoword measures (WRMT, TOWRE), performance differences translated to a learning advantage of two months for the intervention group over the nonintervention group. Typically developing readers show average annual reading gains expressed with an effect size of 0.97 from first to second grade and 0.60 from second to third (Hill, Bloom, Black, & Lipsey, 2007).

Although the current study's implementation during the summer differs from the typical academic-year timeframe used in other investigations of *Seeing Stars* (Krafnick et al., 2011; Olulade et al., 2013), the results converge to indicate promising outcomes on average. The age range in this study was somewhat narrower and younger (6–9 years) in comparison to the prior studies (7–11 years, Krafnick et al., 2011; 7–12 years, Olulade et al., 2013) and included a larger sample size (n = 23) of students receiving intervention (n = 11, Krafnick et al., 2011; n = 22 in Olulade et al., 2013), in addition to a separate control group. The current study used a more condensed treatment duration of 4 hr per weekday for 6 weeks, while comparable studies implemented treatment over 8 weeks (Krafnick et al., 2011), with a single study specifying the duration of 3 hr per weekday for a total of 120 hr of instruction given (Olulade et al., 2013) and neither study reporting size of instructional groups. Across the current and previous studies, the groups receiving intervention showed relative benefits on measures of untimed single real and pseudoword reading (Krafnick et al., 2011; Olulade et al., 2013), with the distinction that only in the current study was the benefit of the treatment driven by the absence of decline on word-level measures in the intervention group. Notably, Olulade et al. (2013) report significant gains in word and pseudoword reading following intervention compared to no significant differences during a control period that followed. Compared to gains in symbol imagery measures reported in Krafnick et al. (2011), the current study suggests similar gains in symbol imagery for the intervention group. Taken together, evidence suggests that *Seeing Stars* can improve word-reading trajectories for early elementary school children in an intensive and condensed time frame and in small groups.

The current study shares characteristics with other programs evaluated after at least 100 hr of instruction for children with RD in kindergarten through third grade (Wanzek & Vaughn, 2007). The current study implemented intervention intensively over 6 weeks; in comparison, other interventions studied lasted 5 months to longer than 1 year without showing a robust relative difference across programs on student outcomes (Wanzek & Vaughn, 2007). The current study findings converge with research indicating that optimized instruction favors smaller groups, offered in first grade, combining phonics and text instruction (Wanzek & Vaughn, 2007). Notably, reading intervention studies tend to investigate the progress of students during the academic year, when reading curricula are replaced or enhanced with research-based programs.

Despite condensed, intensive, small-group instruction, not all students were able to demonstrate immediate reading gains. It is possible that for some children, the lack of positive change in reading outcomes reflected efforts that countered a tendency to show skill regression during the summer months. Alternatively, these outcomes may have indicated an attempt to apply strategies that were not yet mastered, or a need for alternative curricular content, or longer intervention duration. This possibility lends itself to research that investigates how to optimize intervention strategies that are individualized. Future studies can investigate the long-term effects of the intervention program, in comparison to other curricular approaches and across ability levels. Longitudinal data are required, particularly in light of evidence indicating that without sustained instructional support across early elementary school grades, children do not maintain reading gains they acquired after a year of supportive instruction in first grade (Connor et al., 2013).

Limitations

A study limitation is the absence of two additional control groups that would have provided further insights into the basis of the intervention findings. First, there was no active control group receiving an intervention with a curriculum unlikely to help with reading (e.g., on math or arts), but similar to the *Seeing Stars* program in regards to other influences of small-group instruction for a similar duration and intensity. A challenge for this research is that parents of young children with RD are disinclined to make the necessary efforts to have their children attend a summer program designed to not help with reading. An interesting approach to this dilemma is the within-subject cross-over design in which children receive both intervention and active-control conditions in a counterbalanced order (e.g., Olulade et al., 2013). Second, a typically reading control group could also have revealed whether the decline in the nonintervention group was greater than that which occurs in typical students across the summers after first or second grades.

Another study limitation is that findings cannot necessarily be generalized to other instructional settings (e.g., in school during the academic calendar), or intensity (i.e., variations on 4 hr per day). However, a meta-analysis of interventions for students with RD indicated optimized outcomes with intensive instruction delivered in early elementary school in small groups (Vaughn & Wanzek, 2014). Another potential limitation is that treatment fidelity monitored by Lindamood-Bell staff was subjective, and did not include quantitative metrics. However, the teachers who implemented the *Seeing Stars* program were trained and experienced instructors. The lack of consistent standards for collecting and reporting fidelity information is a major limitation for education intervention research more broadly, and a critical requisite for future work that considers measures including checklists, observations, recordings, and details on intervention dosage (E. Swanson, Wanzek, Haring, Ciullo, & McCulley, 2013). Finally, the difference between groups in the number of instructional days of schooling before posttesting was completed is a limitation, although it was nonintervention participants who received more schooling and this ought to have only attenuated intervention efficacy.

Implications for Research and Policy

In addition to measuring the efficacy of the *Seeing Stars* curriculum, this study offers the unique contribution of evaluating intervention efficacy for young students with RD during the summer months, with the pre to post design also affording insights into developmental trajectories. The absence of academic instruction in the summer often leads to a partial loss of what had been learned, such as reading skills, during the school year, a loss termed the "summer slump." Such a slump may be particularly pernicious for young students with RD who are already behind their peers in reading ability at the end of first or second grade and who experience a widening reading achievement gap during the summer that is evident on their return for the next school year. Indeed, the nonintervention group exhibited this widening gap by demonstrating declining reading scores across the summer equivalent to one and a half months. The summer declines in RD students are in sharp contrast to summer gains reported for typical readers in the same early elementary school grades (although beyond third grade summer declines are observed; Cooper, Nye, Charlton, Lindsay, & Greathouse, 1996). This widening reading gap during the summer months has been noted in studies of lower SES students (K. L. Alexander, Entwisle, & Olson, 2007a, 2007b; Burkam, Ready, Lee, & LoGerfo, 2004; Cooper et al., 1996; Heyns, 1987; McCoach, O'Connell, Reis, & Levitt, 2006; Mraz & Rasinski, 2007), and students eligible for special education (Shaw, 1982) or diagnosed with language impairments (Morgan, Farkas, & Wu, 2011).

It is important that the present study shows that remediation delivered in the summer can prevent, and by some measures narrow, the reading gap that is otherwise exacerbated in the summer for students with RD. The intervention group maintained their reading ability across the summer on nearly all measures, and significantly improved their scores on a few measures. The present findings align with prior studies reporting that summer programs can reduce or prevent summer slump in students characterized as having learning disabilities (Cornelius & Semmel, 1982), falling short of literacy benchmarks (Zvoch & Stevens, 2011, 2013), coming from low SES homes (Johnston, Riley, Ryan, & Kelly-Vance, 2015; Kim & Quinn, 2013), or enrolled in schools with programs to foster summer reading in minority children (Kim & White, 2008). Thus, the summer may offer a significant opportunity to support students with reading difficulty in maintaining or improving their reading skills rather than falling even farther behind, especially students from low SES environments (Allington & McGill-Franzen, 2013), and to place those students on a better trajectory for learning to read and for all the other educational benefits and opportunities associated with effective reading.

Acknowledgments

We thank participants, their families, and school communities. We thank our colleagues for their support: Kelly Halverson, Joseph McIntyre, Tyler Perrachione, Keri-Lee Garel. We thank our assessment team: Dorice Moise, Lisa Bloom, Lianne Grasso, Kristen Vecchio, Vickie Chan, Sarah Mackenzie, Colleen Witt, Melissa Boulay, Trish Collins. We thank Lindamood-Bell affiliates: Paul Worthington, Healy Sammis, Rebekah Rogers, Danielle O'Connor, Amy Abbondanzio, Kileigh Donoghue, Rachel Burke, Matt Sorrento, Ben Campbell, Lynn Flowers, Dave Hungerford, Liz Szporn, Jody Gilles, Christina Seremetis, Julie Bogdanski. We thank our affiliates: Patricia Saxler, Mahsa Ershadi, April Maddy, Jahan Naghshineh, Jonathan Walters; Lilla Gabrieli, Louisa Kania, Jessie Hild, Gina Distefano, David Driscoll; Jenny Lu, Katherine Dere; Florence Lai, Ronna Fried.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research (design, recruitment, pre and post behavioral testing, data analysis) was supported by the Ellison Medical Foundation and the Halis Family Foundation; the intervention was supported by Lindamood-Bell Learning Processes.

References

Alexander, A. W., & Slinger-Constant, A. M. (2004). Current status of treatments for dyslexia: Critical review. *Journal of Child Neurology*, *19*, 744–758.

Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2007a). Lasting consequences of the summer learning gap. *American Sociological Review*, *72*, 167–180.

Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2007b). Summer learning and its implications: Insights from the Beginning School Study. *New Directions for Youth Development*, *114*, 11–32.

Allington, R., & McGill-Franzen, A. (2013). Eliminating summer reading setback: How we can close the rich/poor reading achievement gap. *Reading Today*, *30*(5), 10–11.

Al Otaiba, S., & Fuchs, D. (2002). Characteristics of children who are unresponsive to early literacy intervention. *Remedial and Special Education*, *23*, 300–316.

Barratt, W. (2006). *Barratt simplified measure of social status*. Terre Haute: Indiana State University.

Bell, N. (2007). *Seeing stars*. San Luis Obispo, CA: Gander.

Bell, N. (2010). *Symbol imagery test*. Avila Beach, CA: Gander.

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

Burkam, D. T., Ready, D. D., Lee, V. E., & LoGerfo, L. F. (2004). Social-class differences in summer learning between kindergarten and first grade: Model specification and estimation. *Sociology of Education*, *77*, 1–31.

Carrow-Woolfolk, E. (2011). *Oral and Written Language Scales* (2nd ed.). Los Angeles, CA: Western Psychological Services.

Conners, C. K. (1997). *Conners' Rating Scales–Revised*. North Tonawanda, NY: Multi-Health Systems.

Connor, C. M., Morrison, F. J., Fishman, B., Crowe, E. C., Al Otaiba, S., & Schatschneider, C. (2013). A longitudinal cluster-randomized controlled study on the accumulating effects of individualized literacy instruction on students' reading from first through third grade. *Psychological Science*, *24*, 1408–1419.

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research*, *66*, 227–268.

Cornelius, P. L., & Semmel, M. I. (1982). Effects of summer instruction on reading achievement regression of learning. *Journal of Learning Disabilities*, *15*, 409–413.

Denton, C. A. (2012). Response to intervention for reading difficulties in the primary grades: Some answers and lingering questions. *Journal of Learning Disabilities*, *45*, 232–243.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test* (4th ed.). San Antonio, TX: PsychCorp.

Eden, G. F., Jones, K. M., Cappell, K., Gareau, L., Wood, F. B., Zeffiro, T. A., . . .Flowers, D. L. (2004). Neural changes following clinical study remediation in adult developmental dyslexia. *Neuron*, *44*, 411–422.

Fletcher, J. M., & Foorman, B. R. (1994). Issues in definition and measurement of learning disabilities: The need for early intervention. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 185–200). Baltimore, MD: Brookes.

Gabrieli, J. D. E. (2009). Dyslexia: A new synergy between education and cognitive neuroscience. *Science*, *325*, 280–283.

Gersten, R., Compton, D., Connor, C. M., Dimino, J., Santoro, L., Linan-Thompson, S., & Tilly, W. D. (2009). *Assisting students struggling with reading: Response to intervention and multi-tier intervention in the primary grades* (NCEE 2009-4045). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. Retrieved from http://ies.ed.gov/ncee/wwc/publications/practiceguides/

Gilbert, J. K., Compton, D. L., Fuchs, D., Fuchs, L. S., Bouton, B., Barquero, L. A., & Cho, E. (2013). Efficacy of a first-grade responsiveness-to-intervention prevention model for struggling readers. *Reading Research Quarterly*, *48*, 135–154.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). *Behavior Rating Inventory of Executive Function*. Lutz, FL: Psychological Assessment Resources.

Good, R. H., & Kaminski, R. A. (Eds.). (2002). *Dynamic Indicators of Basic Early Literacy Skills* (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement.

Hanselman, P., & Borman, G. (2013). The impacts of Success for All on reading achievement in grades 3–5: Does intervening during the later elementary grades produce the same benefits as intervening early? *Educational Evaluation and Policy Analysis*, *35*, 237–251.

Heyns, B. (1987). Schooling and cognitive development: Is there a season for learning? *Child Development*, *55*, 6–10.

Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2007). *Empirical benchmarks for interpreting effect sizes in research*. New York, NY: MDRC. Retrieved from http://www.mdrc.org/publication/empirical-benchmarks-interpreting-effect-sizes-research

Individuals with Disabilities Education Act of 2004, Pub. L. No. 108-446, §612, 2686 Stat. 118 (2004).

International Dyslexia Association. (2007). *Framework for informed reading and language instruction: Matrix of multisensory structured language programs*. Retrieved from http://www.interdys.org/EWEBEDITPRO5/UPLOAD/MSL2007FINALR1.PDF

Johnston, J., Riley, J., Ryan, C., & Kelly-Vance, L. (2015). Evaluation of a summer reading program to reduce summer setback. *Reading and Writing Quarterly*, *31*, 334–350.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, *80*, 437–447.

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Brief Intelligence Test* (2nd ed.). Minneapolis, MN: Pearson.

Kennedy, M. M., Birman, M., & Demaline, B. (1986). *The effectiveness of Chapter 1 services. Second interim report from the National Assessment of Chapter 1*. Washington, DC: Office of Educational Research and Improvement.

Kim, J. S., & Quinn, D. M. (2013). The effects of summer reading on low-income children's literacy achievement from kindergarten to grade 8: A meta-analysis of classroom and home interventions. *Review of Educational Research*, *83*, 386–431.

Kim, J. S., & White, T. G. (2008). Scaffolding voluntary summer reading for children in grades 3 to 5: An experimental study. *Scientific Studies of Reading*, *12*, 1–23.

Kosslyn, S. M. (1976). Using imagery to retrieve semantic information: A developmental study. *Child Development*, *47*, 434–444.

Krafnick, A. J., Flowers, D. L., Napoliello, E. M., & Eden, G. F. (2011). Gray matter volume changes following reading intervention in dyslexic children. *NeuroImage*, 57, 733–741.

Lerner, J. (1989). Educational interventions in learning disabilities. *Journal of the American Academy of Child and Adolescent Psychiatry*, *28*, 326–331.

Linden, M. A., & Wittrock, M. C. (1981). The teaching of reading comprehension according to the model of generative learning. *Reading Research Quarterly*, *17*, 44–57.

Long, S. A., Winograd, P. N., & Bridge, C. A. (1989). The effects of reader and text characteristics on reports of imagery during and after reading. *Reading Research Quarterly*, *19*, 353–372.

McCoach, B. D., O'Connell, A. A., Reis, S. M., & Levitt, H. A. (2006). Growing readers: A hierarchical linear model of children's reading growth during the first 2 years of school. *Journal of Educational Psychology*, *98*, 14–28.

Meyers, L. S., Gamst, G., & Guarino, A. J. (2012). *Applied multivariate research: Design and Implementation* (2nd ed.). Thousand Oaks, CA: Sage.

Morgan, P. L., Farkas, G., & Wu, Q. (2011). Kindergarten children's growth trajectories in reading and mathematics: Who falls increasingly behind? *Journal of Learning Disabilities*, *44*, 472–488.

Mraz, M., & Rasinski, T. V. (2007). Summer reading loss. *Reading Teacher*, *60*, 784–789.

National Conference of State Legislatures. (2014). *K-3 reading proficiency and early literacy*. Retrieved from http://www.ncsl.org/research/education/third-grade-reading.aspx

Olulade, O. A., Napoliello, E. M., & Eden, G. F. (2013). Abnormal visual motion processing is not a cause of dyslexia. *Neuron*, *79*, 1–11.

Rauschecker, A. M., Deutsch, G. K., Ben-Shacharb, M., Schwartzman, A., Perry, L. M., & Dougherty, R. F. (2009). Reading impairment in a patient with missing arcuate fasciculus. *Neuropsychologia*, *47*, 180–194.

Reynolds, C. R., & Kamphaus, R. W. (2004). *Behavior Assessment System for Children* (2nd ed.). Bloomington, MN: Pearson.

Sadoski, M., & Paivio, A. (2013). *Imagery and text: A dual coding theory of reading and writing* (2nd ed.). New York, NY: Routledge.

Sadoski, M., & Wilson, V. L. (2006). Effects of a theoretically based large-scale reading intervention in a multicultural urban school district. *American Educational Research Journal*, *43*, 137–154.

Schatschneider, C., & Torgesen, J. K. (2004). Using our current understanding of dyslexia to support early identification and intervention. *Journal of Child Neurology*, *19*, 759–765.

Shaw, V. T. (1982). *Retention of selected reading and arithmetic skills by learning disabled pupils and non-disabled pupils over summer vacation*. Unpublished master's thesis, California State College, Stanislaus.

Shaywitz, S. E., Morris, R., & Shaywitz, B. A. (2008). The education of dyslexic children from childhood to young adulthood. *Annual Review of Psychology*, *59*, 451–475.

Snowling, M. J. (2013). Early identification and interventions for dyslexia: A contemporary view. *Journal of Research in Special Educational Needs*, *13*, 7–14.

Swanson, E., Wanzek, J., Haring, C., Ciullo, S., & McCulley, L. (2013). Intervention fidelity in special education and general education research journals. *Journal of Special Education*, *47*, 3–13.

Swanson, H. L. (1999). Reading research for students with LD: A meta-analysis of intervention outcomes. *Journal of Learning Disabilities*, *32*, 504–532.

Tindal, G., Marston, D., & Deno, S. L. (1983). *The reliability of direct and repeated measurement* (Res. Rep. 109). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities.

Torgesen, J. K. (2000). Individual differences in response to early interventions in reading: The lingering problem of treatment resisters. *Learning Disabilities Research & Practice*, *15*, 55–64.

Torgesen, J. K. (2004). Lessons learned from research on interventions for students who have difficulty learning to read. In P. McCardle & V. Chhabra (Eds.), *The voice of evidence in reading research* (pp. 355–382). Baltimore, MD: Brookes.

Torgesen, J. K. (2005). Recent discoveries on remedial interventions for children with dyslexia. In M. J. Snowling & C. Hulme (Eds.), *The science of reading* (pp. 521–537). Malden, MA: Blackwell.

Torgesen, J. K., Rashotte, C. A., & Alexander, A. (2001). Principles of fluency instruction in reading: Relationships with established empirical outcomes. In M. Wolf (Ed.), *Dyslexia, fluency, and the brain* (pp. 333–355). Parkton, MD: York Press.

Torgesen, J. K., Wagner, R., & Rashotte, C. (2012). *Test of Word Reading Efficiency* (2nd ed.). Austin, TX: PRO-ED.

U.S. Department of Education, National Center for Education Statistics. (2013). *Digest of education statistics, 2012* (NCES 2014-015). Washington, DC: Author.

Vaughn, S., Denton, C. A., & Fletcher, J. M. (2010). Why intensive interventions are necessary for students with severe reading disabilities. *Psychology in the Schools*, *47*, 432–444.

Vaughn, S., & Wanzek, J. (2014). Intensive interventions in reading for students with reading disabilities: Meaningful impacts. *Learning Disabilities Research & Practice*, *29*, 46–53.

Wagner, R., Torgesen, J. K., & Rashotte, C. (1999). *Comprehensive Test of Phonological Processing*. Austin, TX: PRO-ED.

Wanzek, J., & Vaughn, S. (2007). Research-based implications from extensive early reading interventions. *School Psychology Review*, *36*, 541–561.

Weinfurt, K. P. (2000). Repeated measures analysis: ANOVA, MANOVA and HLM. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding more multivariate statistics* (pp. 317–361). Washington, DC: American Psychological Association.

Wolf, M., & Denckla, M. B. (2005). *Rapid automatized naming and rapid alternating stimulus tests*. Austin, TX: PRO-ED.

Woodcock, R. W. (2011). *Woodcock Reading Mastery Tests* (3rd ed.). San Antonio, TX: Pearson.

Zvoch, K., & Stevens, J. J. (2011). Summer school and summer learning: An examination of the short- and longer term changes in student literacy. *Early Education and Development*, *22*, 649–675.

Zvoch, K., & Stevens, J. J. (2013). Summer school effects in a randomized field trial. *Early Childhood Research Quarterly*, *28*, 24–32.