The Journal of Special Education

http://sed.sagepub.com

Examining the Impacts of Early Reading Intervention on the Growth Rates in Basic Literacy Skills of At-Risk Urban Kindergarteners

Ya-yu Lo, Chuang Wang and Sherry Haskell

J Spec Educ 2009; 43; 12 originally published online May 13, 2008;
DOI: 10.1177/0022466907313450

The online version of this article can be found at:
http://sed.sagepub.com/cgi/content/abstract/43/1/12

Published by:
Hammill Institute on Disabilities

and

SAGE

http://www.sagepublications.com

The Journal of Special Education
Volume 43 Number 1
May 2009 12-28

10.1177/0022466907313450
http://journalofspecialeducation.sagepub.com
hosted at http://online.sagepub.com

Examining the Impacts of Early Reading Intervention on the Growth Rates in Basic Literacy Skills of At-Risk Urban Kindergarteners

Ya-yu Lo
Chuang Wang
University of North Carolina at Charlotte
Sherry Haskell
Carolina Developmental Therapy Services, Union County, North Carolina

This study investigated the effects of the *Scott Foresman Early Reading Intervention* (ERI) on growth rates in the early literacy skills of urban at-risk kindergarten students. Students participated in one of three groups: treatment-intensive/strategic, treatment-benchmark, and nontreatment-benchmark. Treatment group students received a 30-minute ERI program from classroom teachers 3 days a week for 5 to 14 weeks. Using multivariate analysis of variance and the hierarchical linear model, the authors compared students' benchmark and progress monitoring scores on the Phoneme Segmentation Fluency (PSF) and Nonsense Word Fluency (NWF) subtests of the *Dynamic Indicators of Basic Literacy Skills*™. Results indicated that PSF and NWF benchmark performance gaps decreased between the treatment-intensive/strategic and nontreatment-benchmark groups, indicating beneficial effects for the ERI. Additionally, the PSF and NWF progress monitoring growth rates of treatment group students during the ERI program were significantly higher than rates before treatment. Implications of early reading interventions for urban at-risk students are discussed.

Keywords: *early literacy skills; growth rates; reading interventions; at-risk students*

Urban children from economically disadvantaged families are often less prepared to begin school and demonstrate deficits in school achievement during childhood when compared to their more affluent peers (Cartledge & Lo, 2006; Hecht & Close, 2002; Lutkus, Rampey, & Donahue, 2005; Mahoney, Lord, & Carryl, 2005; Nichols, Rupley, Rickelman, & Algozzine, 2004). Research shows a strong relationship between skills that children possess upon entering school and later academic performance (Spira, Bracken, & Fischel, 2005). Of particular importance to future academic performance is the possession of emergent literacy skills such as phonological awareness, oral language skills, letter knowledge, and print concepts (Spira et al., 2005). According to Schacter and Jo (2005), low-income urban kindergarteners score one half of a standard deviation below the national average in reading achievement, and the gap increases to two standard deviations by the time they graduate from elementary school. With the recent mandates of the No Child Left Behind Act of 2001 (NCLB) to shorten the achievement gaps of disadvantaged groups and to meet the adequate yearly progress criteria for all students, the lack of reading achievement in urban children from low-income families especially presents an urgent concern for educators and parents.

To combat reading underachievement in urban children, it is important that effective reading instruction begin early and address the essentials or "Big

Authors' Note: This project is partially funded by the Bank of America Faculty Research Fund for Young Investigators in Early Childhood Development and Reading and the UNC Charlotte Urban Institute's Faculty Fellows Grants. We wish to thank anonymous reviewers for their constructive comments and Jennifer Courtney and Ellissa Brooks for proofreading our manuscript.
Address correspondence to Chuang Wang, Department of Educational Leadership, College of Education, University of North Carolina at Charlotte, 9201 University City Blvd., Charlotte, NC 28223; e-mail: cwang15@uncc.edu.

Ideas" in reading (Bursuck et al., 2004; Hindson et al., 2005; National Institute of Child Health and Human Development [NICHHD], 2000; Simmons & Kame'enui, 1998). Of the five Big Ideas in beginning reading, phoneme awareness and alphabetic principle are two fundamental skills for reading acquisition (Mann & Foy, 2003; Simmons, Kame'enui, Coyne, & Chard, 2002). Phonemic awareness, focusing on processing and manipulating phonemes in spoken words, is a prerequisite for alphabetic understanding (Byrne & Fielding-Barnsley, 1989), whereas alphabetic principle (requiring one's knowledge to associate sounds with letters) is central for identifying words in print and learning to read (Vellutino, 1991). As a result, an effective reading program should strategically integrate phoneme awareness and alphabetic understanding into instruction (Good, Simmons, & Smith, 1998; NICHHD, 2000; Qi & O'Connor, 2000; Santi, Menchetti, & Edwards, 2004). In addition, authorities also contend that beginning reading instruction should be provided explicitly, systematically, frequently, and intensively in order to produce maximum improvement in reading (Cavanaugh, Kim, Wanzek, & Vaughn, 2004).

For students who are at risk for reading failure and who are nonresponsive to a universal reading program, a secondary reading intervention such as supplemental or remedial reading instruction may be most beneficial (Kamps & Greenwood, 2005; Lane & Menzies, 2003; Lane, Menzies, Munton, von Duering, & English, 2005). For example, Lane and colleagues (2002) examined the effectiveness of a supplemental early literacy program for seven first-grade students identified as at risk for antisocial behavior and who were not responsive to the schoolwide intervention. They randomly assigned students to one of two groups to participate in a two-tiered, multiple-baseline-across-groups design. Participants were assessed during baseline, intervention, postintervention, and follow-up phases. The supplemental reading program consisted of 30 lessons over a 9-week period, using Shefelbine's Phonics Chapter Book series (Shefelbine, 1998, as cited in Lane et al., 2002) and focused on phonemic awareness skills involving blending, sound-letter correspondence, high frequency words, dictation, and writing. Weekly curriculum-based measures were collected to monitor progress using the Nonsense Word Fluency (NWF) and Oral Reading Fluency (ORF) subtests of the *Dynamic Indicators of Basic Early Literacy Skills*™ (DIBELS; Good & Kaminski, 2002). The results showed that all five students who completed the intervention demonstrated increases in the NWF subtest. Four students scored moderate decreases in their ORF subtest during the intervention stage, possibly due to the application of students' newly acquired decoding skills.

Similarly, Gunn, Biglan, Smolkowski, and Ary (2000) evaluated the effects of a supplemental reading program on the phonological awareness and basic decoding skills of 256 at-risk K–3 students who were below grade level on reading or prereading skills. Students, drawn from nine schools with a sizable Hispanic population, were grouped according to grade and ethnicity (Hispanic or non-Hispanic) and randomly assigned to a treatment or control group. Reading Mastery or Corrective Reading programs were used by trained instructional assistants, who worked with the students 25 to 30 minutes daily in a pull-out setting for approximately 16 months during the 2-year study. The interventions focused on sound-letter correspondence, blending, decoding, and fluency building skills. Results were reported to determine students' progress on letter-word identification, word attack, ORF, reading vocabulary, and passage comprehension skills at three waves in time (fall and spring of Year 1 as well as spring of Year 2). Students in the intervention group outperformed the control group on the word attack skill in Year 1 and on all other measures in Year 2. In addition, the supplemental instruction showed effectiveness when ethnic background, gender, or grade levels were controlled. In their follow-up study to determine whether the effectiveness of the supplemental reading instruction was maintained after the intervention was discontinued, Gunn, Smolkowski, Biglan, and Black (2002) found that the 1-year follow-up data were consistent with the positive outcomes reported at the end of the 2-year intervention by Gunn et al. (2000).

Musti-Rao (2005) used a multiple-baseline-across-groups design to investigate the effects of the *Scott Foresman Early Reading Intervention* curriculum (ERI; Simmons & Kame'enui, 2003) as a supplemental reading program to teach phoneme awareness and alphabetic principle skills to eight K–1 at-risk urban students. The participants in each of the three small groups received three 20-minute sessions of the ERI instruction weekly (for a total of 8, 12, and 16 weeks, respectively) from the author and a teacher assistant, as coinstructors. All participants made moderate to substantial increases in the DIBELS Phoneme Segmentation Fluency (PSF) and NWF subtests after the implementation of the intervention. Additionally, four of the participants decreased their risk status by reaching the "benchmark" level on the DIBELS Spring Benchmark Assessment.

Yurick (2006) further extended Musti-Rao's (2005) study to examine the effects of the ERI on the phoneme awareness and alphabetic principle skills of 61 experimental kindergarten students (38 non-English language learners and 23 English language learners [ELLs]) across three high-poverty urban schools. Six instructional assistants and two graduate students delivered the ERI instruction three to five times per week for 25 to 30 minutes per instructional session. Using a pretest-posttest nonequivalent control group design, the author reported that the treatment groups (including treatment-ELL and treatment-non-ELL groups) made substantial gains on the Letter-Word Identification and Word Attack subtests of the *Woodcock Johnson Tests of Achievement III* when compared to the control group. Furthermore, the proportion of treatment group students reaching the benchmark level as indicated by the DIBELS recommendations substantially improved at the end of the school year, whereas approximately 22% of the control group students regressed to either strategic or intensive levels.

Collectively, these studies indicated that a supplemental reading program using a direct and explicit instruction approach could be a viable practice in preventing reading failure among at-risk students. Despite support from the existing research indicating the effectiveness of supplemental early reading programs, current research has two major limitations. First, the examination of performance level in students has been traditionally the predominant approach in screening and evaluating children's reading outcomes, with very limited attention to growth or slope (Speece, 2005). In responding to a recent call for the response-to-intervention effort, researchers have suggested a dual focus of level and rate to more adequately identify intervention nonresponders and/or make adequate instructional decisions to enhance students' learning (Burns & Senesac, 2005; Fuchs, 2003; McMaster, Fuchs, Fuchs, & Compton, 2005). Some researchers have started to evaluate the efficacy of secondary/supplemental reading interventions by comparing growth in basic literacy skills of students receiving interventions and those without interventions (e.g., Bursuck et al., 2004; Kamps & Greenwood, 2005). Continuous research is needed.

Second, an important value of single-subject research designs in evaluating the effectiveness of a practice has relied on comparing participants' performance patterns of repeated measures before and during the intervention phase to document experimental control (Horner et al., 2005). While the use of visual analysis allows one to interpret results in a single-subject research study based on the mean performance, slope or trend of behavior changes, and the fluctuation or variability of behavioral performance, these measures are descriptive in nature and can be limited in the generalization of the results. The application of the hierarchical linear model (HLM) can provide necessary support to data analyses using inferential statistics in single-subject research (Horner, 2006).

The current study was a part of a larger study, extending Musti-Rao's (2005) and Yurick's (2006) studies, which investigated the effects of a supplemental early reading intervention program (i.e., ERI) and intensive tutoring on the basic literacy skills and social behavior of 58 urban kindergarten students in four classes through single-subject research methodology. The focus of the current study was to evaluate the extent to which the ERI program improved treatment students' DIBELS benchmark scores and progress monitoring growth rates with regard to phoneme awareness and letter-sound correspondence skills when compared to those of low-risk students receiving no interventions. Specifically, this study sought first to determine the differences between the treatment and control groups in their PSF and NWF subtest scores of the DIBELS Winter and Spring Benchmark Assessments. Second, to explore the differential growth rates in the participants and to address the research limitations described previously, we conducted HLM procedures (Raudenbush & Bryk, 2002) in data analyses to provide statistical evidence on the participants' progress of change in basic reading skills and to identify reasons for the progress.

Method

Sampling Selection Procedure

Participants were selected from a pool of 58 kindergarten students across four classes in an urban K–5 Title I school, located in a metropolitan southeastern city in the United States. At the time of the study, the school served approximately 600 students, of which 61% were African American, 37% were Latino, 1% were multiracial, and 1% were Caucaisan or Asian American. Students living with one parent comprised 73% of the population, and 94% of the students received free or reduced lunch. All 58 students received the DIBELS Fall, Winter, and Spring Benchmark Assessments as a part of the schoolwide literacy assessment

Table 1
Groupings of Participants by Classes and Ethnicity

	Control Group	Treatment Groups	
Grouping Method	Nontreatment-Benchmark (n = 25)	Treatment-Benchmark (n = 5)	Treatment-Intensive/Strategic (n = 17)
By classes			
Class 1 (n = 14)	7	0	5
Class 2 (n = 15)	3	0	5
Class 3 (n = 14)	7	0	5
Class 4 (n = 15)	8	5	2
By ethnicity			
African American male	7	2	7
African American female	11	2	2
Latino American male	3	0	5
Latino American female	2	1	2
Asian American male	1	0	1
Multiracial American male	1	0	0

Note: The numbers refer to the total number of students enrolled in each class or group.

program. However, only the Winter and Spring Benchmark results were used in the current study as pretest and posttest, respectively. Although data on the Fall Benchmark were available, the Winter Benchmark Assessment results were the basis for participant selection to serve as a more precise pretest indicator than the Fall Benchmark scores. After the completion of the Winter Benchmark, students were placed into three groups according to the instructional recommendations suggested and generated by the DIBELS assessment reporting system (Good, Simmons, Kame'enui, Kaminski, & Wallin, 2002). The "intensive" group (needing substantial intervention) comprised 4 students, the "strategic" group (needing additional intervention) included 14 students, and the "benchmark" group (at grade level) consisted of 40 students.

All students in the intensive and strategic groups were purposefully selected to receive intervention because of their risk status in basic literacy skills based upon their performance on the DIBELS Winter Benchmark. One student in the intensive group was excluded from the data analysis because he received not only the supplemental reading instruction but also additional peer tutoring support on phonological awareness skills, making it difficult to separate the effects of treatments in group data analyses. As a result, 17 students served as the treatment-intensive/strategic group in the current study.

Among the 40 students in the benchmark group, 5 did not reach the benchmark level in either the PSF or NWF subtest on the Winter Benchmark Assessment and were included to receive intervention as the treatment-benchmark group. Finally, a stratified random sampling method was used to select 25 nontreatment-benchmark students from the remaining 35 students in the benchmark group for comparison purposes.

Participants

The sampling selection described above identified 47 participants placed in three groups: treatment-intensive/strategic (n = 17), treatment-benchmark (n = 5), and nontreatment-benchmark (n = 25). Table 1 shows a breakdown of the participants in their groupings by classes and ethnicity. Nontreatment-benchmark students were selected for comparison purposes to control for teacher variables and to see if the treatment students progressed differently (i.e., growth rates) from the nontreatment-benchmark students. If they were different, we wanted to know how they were different and whether the differences were related to the interventions. It is important to repeat that the current study was part of a larger literacy and tutoring support study involving four kindergarten classes that was evaluated primarily through a single-subject, multiple-baseline research methodology. All students in the large study who fell below the Winter Benchmark level received the supplemental reading program in a staggered manner to improve their basic literacy skills. As a result, a comparable control group of students at the strategic or intensive level receiving no supplemental reading instruction was not available for comparion. Furthermore, the small number of treatment-benchmark students (n = 5) was also limited due to the nature of the student characteristics in the selected participants.

Lesson Instructors

Four classroom teachers and one graduate student (all were African American women) served as the reading program instructors. Teacher 1 (30 years old), holding a master's degree in elementary education, was certified to teach K–5 students and had 9 years of teaching experience in kindergarten. She also had experiences with Direct Instruction, the *Open Court Reading Program*, and other phonics-based literacy programs. Teacher 2 (49 years old) was certified to teach elementary education and had 18 years of teaching experience at the elementary level. She had experiences with various reading programs including the *Open Court Reading Program*, *Reading Mastery*, *Break Through to Literacy, Literature Experience*, and *Corrective Reading*. Teacher 3 (25 years old) was certified to teach elementary education and was in her fourth year of teaching. She had experience with the *Open Court Reading Program, Reading Mastery*, and *Hooked on Phonics* during her teaching career. Teacher 4 (42 years old), holding a provisional K–5 license, had 2 years of experience as a lead teacher and 23 years as a teacher assistant, assisting with and instructing in various phonics-based reading programs for young children. The graduate student was seeking special education licensure at the time of the study to teach students with mild to moderate disabilities. As a part of her graduate coursework, she received training on Direct Instruction.

Measurements

The DIBELS (Good & Kaminski, 2002) benchmark and progress monitoring measures were used to evaluate the participants' early literacy skills. In general, the DIBELS assessments are widely used and have excellent technical adequacy (Fuchs, Fuchs, & Compton, 2004; Good & Kaminski, 2002; Good, Kaminski, Simmons, & Kame'enui, 2001; Kaminski & Good, 1996; Speece, Mills, Ritchey, & Hillman, 2003).

DIBELS consists of various standardized, individually administered measures designed to assess students' phonological awareness, alphabetic principle, and fluency with connected text. For kindergarten students, three DIBELS measures were available throughout the winter and spring and could serve as dependent variables: LNF (Letter Naming Fluency), PSF, and NWF. For the purpose of directly evaluating phoneme awareness and letter-sound correspondence skills, only PSF and NWF subtests were used in this study. This focus is supported by research findings that young children's fluency skills in letter-sound relationships (e.g., phonemic segmentation fluency and letter sound fluency) may account for a small, yet unique and significant, amount of variance in word reading and/or spelling (Kame'enui, Simmons, Good, & Harn, 2001; Ritchey & Speece, 2006). Additionally, NWF has been suggested to be a valid measure of early reading and risk status not only for first-grade children but also for children beginning kindergarten in the spring, based on its moderate to strong concurrent and predictive validity coefficients, moderate to significant unique variance contribution to other reading measures, and sensitivity to poor reader status (Speece et al., 2003).

The PSF requires a student to verbally produce individual phonemes of three-, four-, and five-phoneme words fluently in 1 minute. For example, when orally presented with the word "meet," the student will need to say "/m/ /ea/ /t/" to receive three possible points for the word. This measure yields a score for the number of correct phonemes a student orally produces within a 1-minute period. Alternate-form reliability for PSF was .88 for kindergarten children (Kaminski & Good, 1996). Concurrent validity of PSF with the readiness cluster score of the *Woodcock-Johnson Psycho-Educational Battery* was .54 in the spring of kindergarten (Good et al., 2001). Concurrent validity estimates ranged from .43 to .65 on other measures of cognitive ability and school readiness (Kaminski & Good, 1996). Predictive validity of spring kindergarten PSF with spring first-grade *Woodcock-Johnson Psycho-Educational Battery* was .68, and that with curriculum-based measurement ORF was .62 (Good et al., 2001).

The NWF assesses a student's fluency to verbally produce individual sounds of or the whole word for various unfamiliar, nonsense words within a 1-minute period. In this measure, the examiner presents 50 randomly ordered VC or CVC nonsense words on an 8.5-inch × 11-inch sheet of paper to the student. The examiner instructs the student to either verbally produce the individual sound of each letter or say the whole word. For example, when presented with the word "zad," the student could say "/z/ /a/ /d/" or the whole word, "/zad/," to receive a total of three letter-sounds correct. Good et al. (2001) reported concurrent validity with the Woodcock-Johnson readiness cluster score (i.e., visual auditory learning and letter identification) ranging between .35 in May and .59 in February (median coefficient = .52) with samples of 70 to 242 children. The predictive validity coefficients from October of first grade to May of first grade were .71 with respect to passage reading fluency and .52 with respect to the Woodcock-Johnson reading cluster score.

Procedures

General procedures. Treatment students received the supplemental reading instruction in a staggered manner based on the application of a single-subject, multiple-baseline-across-participants design to investigate the functional relationships between the treatment and the dependent variables (Tawney & Gast, 1984). Prior to the baseline data collection, all participants were administered the DIBELS Winter Benchmark Assessment. The assessment results were used to select participants and their groupings (i.e., nontreatment-benchmark, treatment-benchmark, or treatment-intensive/strategic). The DIBELS PSF and NWF progress monitoring assessments were then administered weekly with the treatment group students and biweekly with the control group students to evaluate their progress. The supplemental reading intervention program was provided to three treatment-intensive/strategic students who were considered to be most at risk (first tier) once stability in their baseline data was observed, with the remaining 19 treatment group students continuing with the baseline condition. Once improvements on the progress monitoring performance of the first three treatment-intensive/strategic students were observed, three additional treatment-intensive/strategic students (second tier) with stable baseline data started to receive the intervention. Similarly, once performance improvements were noted for the second-tier students, the intervention was implemented with the remaining 16 students. Finally, the DIBELS Spring Benchmark Assessment was administered to all participants as a postintervention measure.

Baseline. During baseline, all participants received district-prescribed reading instruction (i.e., SRA/McGraw-Hill *Open Court Reading*) for 90 minutes in their respective classrooms. Open Court uses explicit and systematic skill instruction to teach kindergarteners foundational skills, including phonemic awareness, sound-letter correspondence, phonics, fluency, word knowledge, and use of morphemes to enhance decoding skills and learn patterns in the language (SRA/McGraw-Hill, 2005). No supplemental reading intervention was provided for individual students who were performing below benchmark level.

Teacher training. Prior to the commencement of the intervention, each of the lesson instructors received 40 minutes of training from the first author for 2 days on the supplemental reading curriculum (i.e., ERI). On Day 1, the first author explained the essential features of the ERI program and its lessons. The teachers then viewed a videocassette containing a video demonstration of an ERI lesson. The videocassette, a part of the ERI kit, illustrated a teacher delivering an ERI lesson with a group of three students and showed what a 30-minute ERI lesson looked like. On Day 2, the first author demonstrated how to deliver each activity of an ERI lesson, followed by instructors being paired up to practice delivering a lesson. Feedback was provided to instructors throughout the practice. The first author and a trained graduate assistant were present initially during each ERI session for at least 1 week to provide support to each classroom teacher. Periodic classroom visits to the ERI sessions were made by the first author or the graduate assistant thereafter to provide instructors with continuous consultation and support.

Early reading intervention. Treatment group students received 30 minutes of supplemental reading instruction using the ERI program 3 days a week from their classroom teachers and a graduate student in a staggered format, as described previously. Each lesson instructor delivered the ERI lessons to a group of three to six students. The ERI program is an evidence-based early reading program with validated instructional designs (Simmons & Kame'enui, 2003). It was developed to improve reading achievement of at-risk K–1 students. The ERI curriculum contains 126 lessons, within which students receive instruction on phonological awareness, alphabet understanding, word reading, and writing development for each lesson. Prior to the beginning of the ERI program, all treatment students were individually administered a curriculum placement test to determine their entry point in the program. The placement test consisted of six subtests: (a) Letter Names and Sounds, (b) First Sounds in Words, (c) Whole Word Segmentation, (d) Sound-Letter, (e) Whole Word Sound-Letter, and (f) Beginning Word-Reading tests. Students started with the lessons that matched their existing skills (i.e., Lessons 1, 43, 61, or 73). Six treatment-intensive/strategic students started with Lesson 1 (Lesson Groups 1 and 2), five treatment-intensive/strategic students began with Lesson 43 (Group 3), five treatment-intensive/strategic students started with Lesson 61 (Group 4), and the remaining six students began with Lesson 73 (including five treatment-benchmark students and one treatment-intensive/strategic student). All of the 30-minute ERI sessions, set aside from the regular Open Court Reading Program, took place in the teacher's workroom or in the corner of the classroom. During the majority of the intervention sessions, two teacher assistants were available on a rotating basis to provide instruction to the remaining students while the classroom teachers delivered the supplemental reading lessons to the treatment group students.

Benchmark assessments. The DIBELS Winter and Spring Benchmark Assessments were individually administered to all participants by their classroom teachers in December of 2005 and May of 2006, respectively. All classroom teachers administered the assessments by following the DIBELS administration and scoring guide (Good & Kaminski, 2002). For the purpose of the study, the PSF and NWF subtest scores on both benchmark assessments were compared for data analyses among treatment-intensive/strategic students, treatment-benchmark students, and nontreatment-benchmark students.

Progress monitoring assessments. The DIBELS PSF and NWF progress monitoring assessments were administered weekly with the treatment group students by the classroom teachers and biweekly with the nontreatment group students by two graduate assistants. Each of the PSF and NWF subtests contains 20 alternate forms for monitoring progress. Each participant was administered the progress monitoring measures, beginning with the first alternate form and continuing in its sequence regardless of students' absences. The administration of a PSF and an NWF progress monitoring assessment altogether was approximately 4 minutes for each student.

Interobserver Agreement Data

Two graduate students were trained to collect the interobserver agreement (IOA) data on the DIBELS PSF and NWF progress monitoring assessments. The IOA data were collected for 100% of sessions for the 25 nontreatment-benchmark students. Due to the high frequency of the data collection for the 22 treatment group students and the availability of the graduate students, the IOA data for the treatment group students were collected for only 15% of the sessions. The IOA data for the nontreatment-benchmark students were conducted by having both graduate students sit approximately one foot apart from each other and collect the data simultaneously on site. Because the classroom teachers administered the assessments for the treatment group students, each classroom teacher was instructed to audiotape the assessments through a digital audio recorder. The graduate students then listened to the tapes and independently scored the assessments for IOA checks. The percentage of the IOA between the two observers on each assessment was calculated by dividing the smaller score by the larger score and multiplying by 100.

On the PSF assessment, the mean percentage of the IOA data was 99% (range 88%–100%) for the nontreatment-benchmark students and 95.8% (range 47%–100%) for the treatment group students. On the NWF assessment, the mean percentage of the IOA data was 98.1% (range 88%–100%) for the nontreatment-benchmark students and 93.8% (range 43%–100%) for the treatment group students. It is important to note that all individual IOA data were higher than 80% agreement, except during five PSF and six NWF IOA checks because of the poor recording quality and background noise in one particular session.

Treatment Fidelity

Treatment fidelity data were collected for 15% of the intervention sessions across all lesson groups. Each lesson instructor was observed during lesson implementation to evaluate the extent to which the lesson was delivered as planned. Using an eight-step treatment fidelity checklist developed by the first author (i.e., following script, modeling, using manipulatives, feedback, providing individual practice, error correction, reinforcement, lesson timing), a graduate student circled "Yes" if the instructor completed a step/component accurately and marked "No" if the instructor failed to complete the step/component. Treatment fidelity was calculated by dividing the number of "Yes" responses by the total number of steps (i.e., 8) and multiplying by 100. A percentage of 100 treatment fidelity on the lesson implementation was achieved for all instructors.

Experimental Design

A quasi-experimental design was conducted. The treatment group students were selected to receive the supplemental reading program in a staggered, delayed format at three varying points in time. As a result, 3 treatment-intensive/strategic students received the intervention for 14 weeks, another 3 for 10 weeks, and the remaining 11 treatment-intensive/strategic students and the 5 treatment-benchmark students received the intervention for 5 weeks. The staggered, delayed intervention was designed according to (a) the existing literacy skills of the participants that allowed the experimenters to determine if changes in the dependent variables were due to the treatment and (b) the purposes of a larger study from which the current study derived. The nontreatment-benchmark students from the same classes where the treatment group students were instructed served as the control group.

Data Analyses

The multivariate analysis of variance (MANOVA) was conducted to compare the benchmark scores among the three participant groups. The overall alpha level was set at .05.

The increase rates for the PSF and NWF progress monitoring assessments were examined using individual growth curve analysis with HLM. Due to its flexibility to allow different measurement points across individuals in a certain period of time and its robustness against Type I errors across a range of sample sizes and types of serial dependence such as autocorrelation and moving average (Jenson, Clark, Kircher, & Kristjansson, 2007; Raudenbush & Bryk, 2002), HLM has been widely used as a method to examine the growth (rate of change) of students' reading and writing improvement in general (Speece, Ritchey, Cooper, Roth, & Schatschneider, 2004; Taylor, Pearson, & Peterson, 2005) and of exceptional children's growth of literacy in particular (Van den Noorgate & Onghena, 2003a). In addition, HLM has been proposed by a group of scholars as a meta-analytic strategy in analyzing treatment effects of single-subject design research (Shadish & Rindskopf, 2007; Van den Noorgate & Onghena, 2003b). One of the challenges in meta-analyses is to control the treatment conditions and outcome measures across various single-subject designs, especially when these designs are conducted by various researchers. In our study, each individual student received treatment within a single-subject design research, but we controlled both the treatment conditions and outcome measures through a group design.

In our analyses, growth curve trajectories were used for individual students' performances across time, measured weekly or biweekly, and then related to their initial literacy status (intensive, strategic, or benchmark). Two-level HLM was used. At Level 1, each individual student's PSF or NWF scores were predicted by the student's baseline PSF or NWF scores and the number of weeks the student received the treatment. At Level 2, the baseline PSF or NWF scores and the growth rates were predicted by the groups into which the students were classified. The students were grouped in two ways: (a) by treatment condition (intensive/strategic vs. benchmark) and (b) by the length of treatment (14 weeks, 10 weeks, 5 weeks, and no intervention). These variables to represent the method of grouping were centered around the group mean with the intercept representing the average outcome for each group. To estimate the treatment group students' growth rate before and during the treatment, only those students from the two treatment groups (treatment-intensive/strategic and treatment-benchmark) were included in the first round of analyses. In the second round of analyses, all participants were included in order to compare the growth rates between treatment group students during the intervention and nontreatment group students.

Two sets of models were fit for statistical analyses: (a) unconditional models (without predictors) were used to examine the mean and variance of the within-subject parameters, and (b) conditional models (with predictors) were used to estimate the cause of the variance of the within-subject parameters. The unconditional models for PSF and NWF were random intercepts with random slopes, suggesting that both the overall mean scores and growth rates varied across students. Unconditional models were fit first as an exploratory analysis to provide useful empirical evidence to determine which variable to include in the individual growth equation for evaluating subsequent conditional models (Raudenbush & Bryk, 2002). A two-step strategy for the conditional model was used to avoid redundant groupings (Compton, 2000). First, simple conditional models were run to examine each variable individually. Second, the variables significant at the first step were examined simultaneously (complete conditional model) at a significance level of $p < .05$.

Results

PSF and NWF Benchmark Scores

Table 2 shows the descriptive statistics of the PSF and NWF benchmark scores for the treatment-intensive/strategic students, treatment-benchmark students, and the nontreatment-benchmark students. No statistically significant interaction effects were identified between PSF, NWF, and groups, $F(2, 43) = 0.67$, $p > .05$; between PSF and NWF, $F(1, 43) = 0.76$, $p > .05$; or between PSF and groups, $F(2, 43) = 0.39$, $p > .05$. However, the interaction effect between NWF and groups was statistically significant, $F(2, 43) = 7.25$,

Table 2
Means and Standard Deviations of PSF and NWF for the Participants

	Control Group	Treatment Groups	
Dependent Variables	Nontreatment-Benchmark (n = 25)	Treatment-Benchmark (n = 5)	Treatment-Intensive/Strategic (n = 17)
Winter PSF	40.56 (13.04)	45.75 (8.96)	10.06 (7.07)
Spring PSF	59.42 (7.99)	55.40 (3.78)	43.06 (19.84)
Winter NWF	41.92 (14.69)	42.00 (4.69)	13.53 (12.64)
Spring NWF	65.83 (20.26)	54.60 (9.34)	45.94 (13.53)

Note: The numbers in parentheses are standard deviations. PSF = Phoneme Segmentation Fluency subtest; NWF = Nonsense Word Fluency subtest.

$p < .05$. Although this interaction was statistically significant, the interaction was found to be orthogonal (the means for all three groups were increasing). Therefore, we proceeded to analyze the main effects.

Our data showed that the supplemental ERI program had positive impacts on the treatment students' PSF and NWF benchmark scores. Specifically, in the winter the treatment-intensive/strategic students had significantly lower PSF scores, $F(2, 43) = 46.54$, $p < .05$, partial $\eta^2 = .67$, and NWF scores, $F(2, 43) =$ 25.21, $p < .05$, partial $\eta^2 = .51$, than the treatment-benchmark and nontreatment-benchmark students. No statistically significant differences were noted between the treatment-benchmark and nontreatment-benchmark students (ps > .05). In the spring, the treatment-intensive/strategic students continued to show significantly lower PSF scores, $F(2, 43) = 8.05$, $p < .05$, partial $\eta^2 = .26$, and NWF scores, $F(2, 43) = 7.54$, $p < .05$, partial $\eta^2 = .24$, than the nontreatment-benchmark students. No statistically significant differences, however, were found between either the treatment-intensive/strategic and treatment-benchmark students or the treatment-benchmark and nontreatment-benchmark students (ps > .05). This was partly due to the small sample size of the treatment-benchmark group but also partly due to the comparatively smaller differences between the two compared groups. In other words, even though significant differences were noted between these three groups in the winter and spring, the effect sizes were much smaller in the spring than in the winter (.26 vs. .67 for PSF and .24 vs. .51 for NWF), indicating that the gap between the treatment-intensive/strategic students and the other two groups became less salient after the intervention. Considering the limited length of intervention periods (5–14 weeks), the improvements for the treatment group students were noteworthy.

PSF and NWF Progress Monitoring Scores

First round of analyses. The estimates for the unconditional models are presented in Table 3, whereas those for the conditional models are presented in Table 4.

When comparing the growth rates for all treatment group students before and during the intervention, the results showed that both rates were significantly different from zero, $p < .001$. The growth rate increased from 1.02 to 5.30 for PSF and from 1.34 to 4.98 for NWF. This means that on average, all treatment group students gained 1.02 in PSF scores every week before the intervention but gained 5.30 in PSF scores every week during the intervention. Similarly, all treatment group students gained 1.34 in NWF scores every week before the intervention but gained 4.98 in NWF scores every week during the intervention (Table 3).

Regarding the length of treatment, the results suggested that the amount of time students received the intervention significantly influenced their growth rates before the intervention, $t(20) = 3.76$, $p = .001$, and during the intervention, $t(20) = 3.12$, $p = .006$, for PSF. On average, the weekly PSF growth rate for students receiving 5 weeks of treatment was 0.79 higher than that for students receiving 10 weeks of treatment, while the students with 10 weeks of treatment had a growth rate that was 0.79 higher than that for students receiving 14 weeks of treatment before the intervention. For the growth rates during the intervention, the weekly PSF growth rate for students receiving 5 weeks of treatment was 2.13 higher than that for students receiving 10 weeks of treatment, while the students with 10 weeks of treatment had a growth rate that was 2.13 higher than that for students receiving 14 weeks of treatment. Similar patterns were noted for NWF with the difference of the weekly

Table 3
Estimates for the Unconditional Model of Growth in PSF and NWF for Treatment Groups Only

	Fixed Effects				Random Effects		
Estimated Parameters	Coefficient	*SE*	*t*	*p*	Variance	χ^2	*p*
PSF							
Intercept	17.00	2.87	5.92	<.001	177.61	282.82	<.001
Slope 1	1.02	0.20	5.10	<.001	0.67	60.47	<.001
Slope 2	5.30	0.79	6.74	<.001	12.66	307.74	<.001
Residual					63.79		
NWF							
Intercept	23.40	3.17	7.39	<.001	219.55	376.20	<.001
Slope 1	1.34	0.14	9.71	<.001	0.14	30.04	.091
Slope 2	4.98	0.47	10.66	<.001	3.66	149.35	<.001
Residual					58.51		

Note: PSF = Phonemic Segmentation Fluency subtest; NWF = Nonsense Word Fluency subtest; Slope 1 = increase rate before intervention; Slope 2 = increase rate during intervention; *df* = 21.

Table 4
Estimates for the Conditional Model of Growth in PSF and NWF for Treatment Groups Only

	Simple Conditional Model				Complete Conditional Model			
Estimated Parameters	Coefficient	*SE*	*t*	*p*	Coefficient	*SE*	*t*	*p*
PSF intercept								
Group 1	16.52	5.49	3.01	.007	12.97	5.83	2.23	.038
Group 2	8.93	2.79	3.20	.005	6.58	2.78	2.37	.029
PSF Slope 1								
Group1	0.01	0.45	0.02	.988				
Group 2	0.79	0.21	3.76	.001	0.86	0.23	3.73	.02
PSF Slope 2								
Group 1	–0.46	1.73	–0.27	.792				
Group 2	2.13	0.68	3.12	.006	2.46	0.78	3.15	.006
NWF intercept								
Group 1	14.57	6.10	2.39	.027	7.88	5.82	1.35	.192
Group 2	13.70	3.02	4.53	<.001	12.26	2.98	4.12	.001
NWF Slope 1								
Group 1	0.74	0.26	2.79	.012	0.63	0.27	2.31	.032
Group 2	0.75	0.33	2.28	.034	0.66	0.32	2.07	.052
NWF Slope 2								
Group 1	1.27	0.84	1.52	.144				
Group 2	1.56	0.36	4.33	<.001	1.48	0.46	3.22	.005

Note: PSF = Phonemic Segmentation Fluency subtest; NWF = Nonsense Word Fluency subtest; Group 1 = students classified according to treatment condition (intensive/strategic vs. benchmark); Group 2 = students classified according to the length of treatment; Slope 1 = increase rate before intervention; Slope 2 = increase rate during intervention; *df* = 20 for the simple conditional model, and *df* = 19 for the complete conditional model.

growth rate being 0.75 before the intervention and 1.56 during the intervention (Table 4). In other words, the shorter the time the students received the intervention, the higher the increase rates.

The weekly growth rates for PSF between the treatment-intensive/strategic students and treatment-benchmark students were not significantly different either before the intervention, $t(20) = 0.02$, $p = .988$, or during the intervention, $t(20) = -0.27$, $p = .792$. The weekly growth rates for NWF between these two groups were significantly different before the intervention, $t(20) = 2.79$, $p = .012$, but not significantly different during the intervention, $t(20) = 1.52$, $p = .144$. On average, the treatment-benchmark students' weekly growth rate in NWF scores before the intervention was 0.74 higher than that of the treatment-intensive/strategic students.

In the simple conditional models, both treatment condition and the length of treatment were significant correlates of the intercept for PSF and NWF. Treatment condition was not a significant correlate of the slope either before or during the intervention for PSF, but a significant correlate of the slope before the intervention for NWF. In the complete conditional models, both treatment condition and the length of treatment were still significant correlates of the intercept for PSF, but only the length of treatment was a significant correlate of the intercept for NWF. These variables accounted for 34.04% of the explainable variance in the initial PSF scores before the intervention. The length of treatment accounted for 46.78% of the explainable variance in the initial NWF scores before the intervention. Both treatment condition and the length of treatment were significant correlates for the slopes before the intervention for NWF, accounting for 78.57% of the explainable variance in the NWF before-intervention weekly growth rates. On the other hand, the length of the treatment was the only significant correlate of the slope before the intervention for PSF, accounting for 23.88% of the explainable variance in this slope. The length of treatment was also the only significant correlate of the slopes during the intervention for both PSF and NWF, accounting for 13.82% and 27.60% of the explainable variance in the weekly growth rate during the intervention for PSF and NWF, respectively.

Second round of analyses. When both treatment and nontreatment groups were included in the second round of analysis, the overall weekly growth rates for all participants were significantly different from zero: 1.37 for PSF, $t(45) = 4.06$, $p < .001$, and 1.38 for NWF, $t(45) = 6.25$, $p < .001$ (Table 5). The length of treatment did not have any statistically significant impact on the weekly growth rates for PSF, $t(43) = 0.82$, $p = .418$, but significantly influenced the weekly growth rates for NWF, $t(43) = 2.11$, $p = .040$. On average, the weekly NWF growth rate for students receiving 5 weeks of treatment was 0.54 higher than that for students receiving 10 weeks of treatment, while the students with 10 weeks of treatment had a growth rate that was 0.54 higher than that for students with 14 weeks of treatment (Table 6).

When the weekly growth rates were compared between the three groups, statistically significant differences were noted for both PSF, $t(43) = -2.99$, $p = .005$, and NWF, $t(43) = -5.28$, $p < .001$. The nontreatment-benchmark students' weekly growth rate in PSF was 2.32 lower than that of the treatment-benchmark students, and the treatment-benchmark students' weekly growth rate in PSF was 2.32 lower than that of the treatment-intensive/strategic students. Similarly, the nontreatment-benchmark students' weekly growth rate in NWF was 1.94 lower than that of the treatment-benchmark students, and the treatment-benchmark students' weekly growth rate in NWF was 1.94 lower than that of the treatment-intensive/strategic students (Table 6).

Finally, in the simple conditional models, both treatment condition and the length of treatment were significant correlates of the intercept and the slopes for PSF and NWF. In the complete conditional models, only the length of treatment remained to be the significant correlate of the intercept for PSF, accounting for 52.45% of the explainable variance in the initial PSF scores, and the intercept for NWF, accounting for 43.03% of the explainable variance in the initial NWF scores. Treatment condition was the only significant correlate of the slope for PSF, accounting for 65.99% of the explainable variance in the weekly growth rates for PSF. Both treatment condition and the length of the treatment were significant correlates of the slope for NWF, accounting for 89.66% of the explainable variance in this slope.

Discussion

This study evaluated the effects of the ERI program, as a supplemental reading intervention, on the phoneme awareness and letter-sound correspondence skills of 22 treatment group students in comparison with 25 control group students. The results of this study support the efficacy of the ERI program in three aspects: (a) the performance gap between the treatment-intensive/strategic students and the benchmark students was reduced (i.e., pretest and posttest analysis on benchmark assessments), (b) the treatment-intensive/strategic students greatly improved their PSF and NWF growth rates with statistical significance after the intervention was implemented (i.e., before-during treatment analysis), and (c) the treatment-intensive/strategic students produced the highest growth rates on both PSF and NWF progress monitoring measures, followed by the treatment-benchmark students, and then the nontreatment-benchmark students.

Table 5
Estimates for the Unconditional Model of Growth in PSF and NWF for All Participants

	Fixed Effects				Random Effects		
Estimated Parameters	Coefficient	*SE*	*t*	*p*	Variance	χ^2	*p*
PSF							
Intercept	33.48	2.48	13.49	<.001	263.98	895.50	<.001
Slope	1.37	0.34	4.06	<.001	4341	345.94	<.001
Residual					104.55		
NWF							
Intercept	38.31	2.56	14.98	<.001	278.85	784.92	<.001
Slope	1.38	0.22	6.25	<.001	1.45	160.15	<.001
Residual					118.40		

Note: PSF = Phonemic Segmentation Fluency subtest; NWF = Nonsense Word Fluency subtest; slope = increase rate; *df* = 45.

Table 6
Estimates for the Conditional Model of Growth in PSF and NWF for All Participants

	Simple Conditional Model				Complete Conditional Model			
Estimated Parameters	Coefficient	*SE*	*t*	*p*	Coefficient	*SE*	*t*	*p*
PSF intercept								
Group 1	12.01	2.15	5.59	<.001	5.17	3.69	1.40	.168
Group 2	13.28	1.82	7.28	<.001	8.81	3.24	2.72	.010
PSF slope								
Group 1	–1.83	0.32	–5.76	<.001	–2.32	0.78	–2.99	.005
Group 2	–1.44	0.36	–4.02	<.001	0.53	0.65	0.82	.418
NWF intercept								
Group 1	10.21	2.33	4.38	<.001	0.84	3.20	0.26	.794
Group 2	12.49	1.80	6.94	<.001	12.06	2.84	4.25	<.001
NWF slope								
Group 1	–1.41	0.16	–9.04	<.001	–1.94	0.37	–5.28	<.001
Group 2	–0.99	0.17	–5.65	<.001	0.54	0.25	2.11	.040

Note: PSF = Phonemic Segmentation Fluency subtest; NWF = Nonsense Word Fluency subtest; Group 1 = students classified according to treatment conditions (intensive/strategic, treatment benchmark, and nontreatment benchmark); Group 2 = students classified according to the length of treatment; slope = increase rate; *df* = 44 for the simple conditional model, and *df* = 43 for the complete conditional model.

The results on the DIBELS benchmark assessment scores indicated that the treatment-intensive/strategic students scored significantly lower on the PSF and NWF subtests in the winter (pretest), when compared to the other two benchmark groups with or without treatment. This was an expected finding based on the fact that the treatment-intensive/strategic students were purposefully selected to receive intervention due to their skill deficits. The posttest data collected in the spring showed that the treatment-intensive/strategic students continued to score the lowest among the three groups on both PSF and NWF assessments; however, the effect sizes for the differences were much smaller. All but 1 of the 17 treatment-intensive/strategic students achieved at the benchmark category on the Spring Benchmark Assessment. It is also noteworthy, though not previously reported, that 10 of the 17 treatment-intensive/strategic students did not change their risk status from the Fall to the Winter DIBELS Benchmark Assessments with the core literacy program. It is possible that the combination of the core literacy program and the supplemental reading intervention during the spring semester is responsible for the treatment-intensive/strategic students' performance in the following three areas: (a) the improvements from winter to spring in benchmark assessment scores, (b) the reduced performance gap with their counterparts, and (c) the reduced risk status for reading failure.

These findings are consistent with previous research (Bursuck et al., 2004; Kamps & Greenwood, 2005;

Yurick, 2006) in that secondary or supplemental reading instruction with a strong emphasis on letter-sound relationships and phoneme awareness is beneficial to at-risk students who are nonresponsive to a schoolwide core literacy program in improving their basic literacy skills and reducing their risk status for reading failure. An implication points to the need for formulating secondary or supplemental early reading interventions in order for at-risk learners to catch up with their peers and to bridge the achievement gaps of disadvantaged children as mandated in the NCLB. Nevertheless, it is important to note that these data do not offer evidence to suggest that the participants at the intensive or strategic levels will no longer be considered at risk or in need of supplemental instruction. In fact, the continued gaps (although reduced) suggest that an enduring early reading intervention may be essential (Cavanaugh et al., 2004) and that prevention and intervention should start early to ensure reading success in young children (Simmons et al., 2002).

The PSF and NWF growth rate comparisons among the participants in the three groups revealed three important findings. First, when aggregating all participants' data, we found that the overall growth rates for both PSF and NWF progress monitoring assessments were significant ($p < .001$). Although various potential explanations may be possible for the significance, two were most relevant and apparent: (a) the positive overall growth rates for all participants may indicate a natural growing of young children in their basic literacy skills as they interacted with their learning environment, and (b) the results may imply the potential positive impacts of the regular literacy instruction (i.e., Open Court Reading Program) as an attributor to the participants' performance gains. To prevent reading failure in the majority of young children, schools can provide systematic and effective interventions at the schoolwide or global level to provide student support as the first-tier instruction along a continuum of intervention intensities (Bursuck et al., 2004; Lane & Menzies, 2002).

Second, the treatment-intensive/strategic students' weekly growth rates in both PSF and NWF were significantly higher than those of the treatment-benchmark group and the nontreatment-benchmark students, suggesting that the ERI program had a significant impact on the students' weekly growth rates and that students at higher risks might have benefited more from an explicit supplemental reading program than students at lower risks.

Third, the before- and during-treatment comparisons indicated that the treatment group students had statistically significant growth rates ($p < .001$) during the supplemental early reading intervention. In other words, the slope of the growth rate during treatment was steeper than the slope of the growth rate before treatment, indicating the additive effects of the supplemental reading program. Most importantly, when evaluating the trajectory growth rates for individual participants based on the before-treatment data, we found that the treatment-intensive/strategic students would not have reached the benchmark level on the Spring Benchmark Assessment without a supplemental reading intervention. This lends further support to previous literature (Musti-Rao, 2005; Yurick, 2006), suggesting that without additional instruction or a second-tier intervention, students at risk may continue to be victims of reading failure.

The analyses of treatment duration suggested that the amount of treatment (5, 10, or 14 weeks) was a significant factor affecting the level of performance growth rates. Specifically, students who received the longest amount of treatment had the lowest growth rate, and those who received the shortest treatment produced the highest growth rate. This finding may not be surprising. The lowest performing students were selected to receive the treatment first with the assumption that they would require a longer period of time to catch up with their peers or to reach the benchmark level. Students with a relatively higher performance or a lower level of risk status (e.g., strategic) may need only a short period of booster instruction to remediate their deficits. This is consistent with the existing research in two ways: (a) the unique needs of individual students in school require a multitiered program that can provide different levels of support for students with varying levels of skills and competencies (Bursuck et al., 2004; Lane & Menzies, 2003), and (b) high-risk children can reach a similar level of achievement to that of higher performing students but will need more sustained teaching to do so (Hindson et al., 2005). Teachers working with at-risk children will be well advised to seek a comprehensive reading model that provides different levels of intensity of instruction for reading achievement to all students.

Limitations and Future Research

The current study has several limitations that warrant future research. First, this study did not include a true control group with random selection and assignment to allow for comparisons of students at risk with and without treatment. The study design and constraints derived from the larger study solely with four kindergarten classrooms restricted the availability of sufficient students at the strategic or intensive levels to serve as the control group. Future research should consider including a larger scale of student population, randomly selecting students from all levels of literacy proficiency and then randomly assigning them into treatment and control groups to establish comparable groups in order to identify contributing factors (including maturation) to students' growth rates in literacy. In addition, although we had benchmark students in both treatment and control groups, the sample size for the treatment-benchmark group was small ($n = 5$) for statistical comparisons.

Second, the current study utilized PSF and NWF subtests as the sole and primary dependent variables. The predictive outcomes of PSF and NWF to early reading have been mixed (e.g., Fuchs et al., 2001; Ritchey & Speece, 2006). Some researchers found that PSF uniquely predicted spelling but not word reading in kindergarten (Ritchey & Speece, 2006). Others have shown that the word identification fluency measure may be a more accurate indicator, over NWF measure, on the first-grade end-of-year reading outcomes (Fuchs et al., 2004). As a result, solely depending on the use of phoneme segmentation and nonsense word measures to identify the need for supplemental reading instruction or to monitor at-risk readers' progress may bear limited relationship to students' overall reading performance beyond kindergarten. Multiple inclusion of valid sublexical fluency measures may be necessary.

Third, a limitation of this study concerns the use of staggered, delayed delivery of the treatment at three varying waves in time for the treatment group students. Although theoretically sound and practically parsimonious, this arrangement with various lengths of intervention requires a careful interpretation of the results when applying statistical methodology for group comparisons. Due to the design of this study, the length of intervention was not independent of students' baseline performance. As a result, the statistically significant differences between the groups classified according to the length of intervention should be interpreted with caution. The result that students who received shorter periods of intervention had higher weekly growth rates could be interpreted that students with better performance at baseline might have had sufficient skills to allow them progress at a higher rate with minimum to moderate support or instruction. To produce more accurate comparisons, future research is warranted to explore the relationships between treatment length and growth rates of students performing at comparable baseline levels.

Finally, one important benefit of incorporating HLM as a statistical method is that it does not require the same number of data sets obtained at the same period in time, nor does it require the same length of intervention for groups of students to yield meaningful comparisons. These features support the use of HLM as a useful method to statistically analyze the individual growth rates using repeated performance measures of the participants from a single-subject research study (Horner, 2006). Despite the unique features of HLM, we were unable to determine the most adequate length of treatment that may be required for individual students based on their risk status and growth rate pattern. In other words, we were not certain whether the length of treatment provided to the participants was too long, too short, or just right. Future researchers should examine the fluctuation of growth rates during treatment in order to identify the proper treatment length to enhance its cost-effectiveness. Investigation in this area will provide implications for researchers and practitioners on appropriate implementation of treatment duration based on students' needs.

Implications for Practice

Despite these limitations, this study has important implications. First, urban at-risk students often are in jeopardy of developing reading deficits and falling behind their more affluent peers without systematic, well-rounded reading intervention programs. To reduce their achievement gaps with other average- or high-achieving students as mandated in the NCLB, it is important that schools adopt explicit early literacy programs and monitor the amount and intensity of support (e.g., supplemental instruction) that at-risk children will need to remediate their underachievement. Second, careful study of growth rates in specific skills may provide useful information to teachers and literacy specialists regarding at-risk students' progress over time and may suggest an estimate of intervention duration for instructional planning purpose. Last, this study demonstrates that classroom teachers can be effective supplemental reading instructors for at-risk urban students. On many occasions, classroom teachers may find it very difficult to allocate additional class time to work with a small number of students who will need additional support. Some studies have shown that teacher assistants or paraprofessionals (Gunn et al., 2000; Musti-Rao, 2005; Yurick, 2006) are great resources to serve as instructors to deliver supplemental reading programs in order to ease classroom teachers' roles. In the current study, the four classroom teachers delivered the supplemental reading instruction with high fidelity while the teacher assistants worked with the remaining students in seatwork or review of skills. All of the teachers reported high acceptability of the program and expressed that they planned to continue the program with their low-achieving students the following year. This may be related to the availability of the highly prescribed curriculum and the fact that the participants were grouped according to their entry performance, therefore minimizing the need for teachers to differentiate instruction for individual students in each group. In either case, it is clear that well-trained instructors can lead to effective implementation of a reading program that contributes to students' learning. Depending on the number of students who need additional interventions, schools may designate specific roles to classroom teachers, teacher assistants, or other staff members who are well trained in delivering supplemental reading instruction to at-risk students.

References

Burns, M. K., & Senesac, B. V. (2005). Comparison of dual discrepancy criteria to assess response to intervention. *Journal of School Psychology*, *43*, 393–406.

Bursuck, W. D., Smith, T., Munk, D., Damer, M., Mehlig, L., & Perry, J. (2004). Evaluating the impact of a prevention-based model of reading on children who are at-risk. *Remedial and Special Education*, *25*, 303–313.

Byrne, B., & Fielding-Barnsley, R. (1989). Phonemic awareness and letter knowledge in the child's acquisition of the alphabetic principle. *Journal of Educational Psychology*, *81*, 313–321.

Cartledge, G., & Lo, Y. (2006). Teaching urban learners: Culturally responsive strategies for developing academic and behavioral competence. Champaign, IL: Research Press.

Cavanaugh, C. L., Kim, A., Wanzek, J., & Vaughn, S. (2004). Kindergarten reading interventions for at-risk students: Twenty years of research. *Learning Disabilities: A Contemporary Journal*, *2*(1), 9–21.

Compton, D. L. (2000). Modeling the growth of decoding skills in first-grade children. *Scientific Studies in Reading*, *4*, 219–259.

Fuchs, L. S. (2003). Assessing intervention responsiveness: Conceptual and technical issues. *Learning Disabilities Research & Practice*, *81*, 172–186.

Fuchs, L. S., Fuchs, D., & Compton, D. L. (2004). Monitoring early reading development in first grade: Word identification fluency versus nonsense word fluency. *Exceptional Children*, *71*, 7–21.

Fuchs, D., Fuchs, L. S., Thompson, A., Al Otaiba, S., Yen, L., Yang, N. J., et al. (2001). Is reading important in reading-readiness programs? A randomized field trial with teachers as program implementers. *Journal of Educational Psychology*, *93*, 251–267.

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. Retrieved May 15, 2005, from http://dibels.uoregon.edu

Good, R. H., III, Kaminski, R. A., Simmons, D., & Kame'enui, E. J. (2001). *Using Dynamic Indicators of Basic Early Literacy Skills (DIBELS) in an outcomes-driven model: Steps to reading outcomes*. Unpublished manuscript, University of Oregon at Eugene.

Good, R. H., Simmons, D. C., Kame'enui, E., Kaminski, R. A., & Wallin, J. (2002). DIBELS instructional recommendations: Intensive, strategic, and benchmark. In R. H. Good & R. A. Kaminski (Eds.), *Dynamic Indicators of Basic Early Literacy Skills* (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement.

Good, R. H., Simmons, D. C., & Smith, S. B. (1998). Effective academic intervention in the United States: Evaluating and enhancing the acquisition of early reading skills. *Educational and Child Psychology*, *15*(1), 56–70.

Gunn, B., Biglan, A., Smolkowski, K., & Ary, D. (2000). The efficacy of supplemental instruction in decoding skills for Hispanic and Non-Hispanic students in early elementary school. *The Journal of Special Education*, *34*, 90–103.

Gunn, B., Smolkowski, K., Biglan, A., & Black, C. (2002). Supplemental instruction in decoding skills for Hispanic and non-Hispanic students in early elementary school: A follow-up. *The Journal of Special Education*, *36*, 69–79.

Hecht, S. A., & Close, L. (2002). Emergent literacy skills and training time uniquely predict variability in responses to phonemic awareness training in disadvantaged kindergartners. *Journal of Experimental Child Psychology*, *82*, 93–115.

Hindson, B., Byrne, B., Fielding-Barnsley, R., Newman, C., Hine, D. W., & Shankweiler, D. (2005). Assessment and early instruction of preschool children at risk for reading disability. *Journal of Educational Psychology*, *97*, 687–704.

Horner, R. H. (2006, May). *Using single case designs to establish evidence-based practice*. Tutorial presented at the 32nd annual convention of the Association for Behavior Analysis, Atlanta, GA.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, *71*, 165–179.

Jenson, W. R., Clark, E., Kircher, J. C., & Kristjansson, S. D. (2007). Statistical reform: Evidence-based practice, meta-analyses, and single subject designs. *Psychology in the Schools*, *44*, 483–493.

Kame'enui, E. J., Simmons, D. C., Good, R. H., & Harn, B. A. (2001). The use of fluency-based measures in early identification and evaluation of intervention efficacy in schools. In M. Wolf (Ed.), *Dyslexia, fluency, and the brain* (pp. 307–331). Baltimore: York Press.

Kaminski, R. A., & Good, R. H. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology*, *25*, 215–227.

Kamps, D. M., & Greenwood, C. R. (2005). Formulating secondary-level reading interventions. *Journal of Learning Disabilities*, *38*, 500–509.

Lane, K. L., & Menzies, H. M. (2002). The effects of a school-based primary intervention program: Preliminary outcomes. *Preventing School Failure*, *47*, 26–32.

Lane, K. L., & Menzies, H. M. (2003). A school-wide intervention with primary and secondary levels of support for elementary students: Outcomes and considerations. *Education and Treatment of Children*, *26*, 431–451.

Lane, K. L., Menzies, H. M., Munton, S. M., von Duering, R. M., & English, G. L. (2005). The effects of a supplemental early literacy program for a student at risk: A case study. *Preventing School Failure*, *50*, 21–28.

Lane, K. L., Wehby, J. H., Menzies, H. M., Gregg, R. M., Doukas, G. L., & Munton, S. M. (2002). Early literacy instruction for first-grade students at-risk for antisocial behavior. *Education and Treatment of Children*, *25*, 438–458.

Lutkus, A. D., Rampey, B. D., & Donahue, P. (2005). *The nation's report card: Trial Urban District Assessment: Reading 2005* (NCES 2006-455). Washington, DC: Government Printing Office, U.S. Department of Education, National Center for Education Statistics.

Mahoney, J. L., Lord, H., & Carryl, E. (2005). An ecological analysis of after-school program participation and the development of academic performance and motivational attributes for disadvantaged children. *Child Development*, *76*, 811–825.

Mann, V. A., & Foy, J. G. (2003). Phonological awareness, speech development, and letter knowledge in preschool children. *Annals of Dyslexia*, *53*, 149–173.

McMaster, K. L., Fuchs, D., Fuchs, L. S., & Compton, D. L. (2005). Responding to nonresponders: An experimental field trial of identification and intervention methods. *Exceptional Children*, *71*, 445–463.

Musti-Rao, S. (2005). *The effects of a supplemental early reading intervention with urban kindergarten and first-grade students: A preventative approach*. Unpublished doctoral dissertation, Ohio State University, Columbus.

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4754). Washington, DC: Government Printing Office.

Nichols, W. D., Rupley, W. H., Rickelman, R. J., & Algozzine, B. (2004). Examining phonemic awareness and concepts of print patterns of kindergarten students. *Reading Research and Instruction*, *43*(3), 56–81.

Qi, S., & O'Connor, R. (2000). Comparison of phonological training procedures in kindergarten classrooms. *The Journal of Educational Research*, *93*, 226–234.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage.

Ritchey, K. D., & Speece, D. L. (2006). From letter names to word reading: The nascent role of sublexical fluency. *Contemporary Educational Psychology*, *31*, 301–327.

Santi, K. L., Menchetti, B. M., & Edwards, B. J. (2004). A comparison of eight kindergarten phonemic awareness programs based on empirically validated instructional principles. *Remedial and Special Education*, *25*, 189–196.

Schacter, J., & Jo, B. (2005). Learning when school is not in session: A reading summer day-camp intervention to improve the achievement of exiting first-grade students who are economically disadvantaged. *Journal of Research in Reading*, *28*, 158–169.

Shadish, W. R., & Rindskopf, D. M. (2007). Methods for evidence-based practice: Quantitative syntheses of single-subject designs. *New Directions for Evaluation*, *113*, 95–109.

Simmons, D. C., & Kame'enui, E. J. (Eds.). (1998). *What reading research tells us about children with diverse learning needs: Bases and basics*. Mahwah, NJ: Lawrence Erlbaum.

Simmons, D. C., & Kame'enui, E. J. (2003). *Scott Foresman Early Reading Intervention*. Glenview, IL: Scott Foresman.

Simmons, D. C., Kame'enui, E. J., Coyne, M. D., & Chard, D. J. (2002). Effective strategies for teaching beginning reading. In E. J. Kame'enui, D. W. Carnine, R. C. Dixon, D. C. Simmons, & M. D. Coyne (Eds.), *Effective teaching strategies that accommodate diverse learners* (2nd ed., pp. 53–92). Upper Saddle River, NJ: Merrill/Prentice Hall.

Speece, D. L. (2005). Hitting the moving target known as reading development: Some thoughts on screening children for secondary interventions. *Journal of Learning Disabilities*, *38*, 487–493.

Speece, D. L., Mills, C., Ritchey, K. D., & Hillman, E. (2003). Initial evidence that letter fluency tasks are valid indicators of early reading skill. *The Journal of Special Education*, *36*, 223–233.

Speece, D. L., Ritchey, K. D, Cooper, D. H., Roth, F. P., & Schatschneider, C. (2004). Growth in early reading skills from kindergarten to third grade. *Contemporary Educational Psychology*, *29*, 312–332.

Spira, E. G., Bracken, S. S., & Fischel, J. E. (2005). Predicting improvement after first-grade reading difficulties: The effects of oral language, emergent literacy, and behavior skills. *Developmental Psychology*, *41*, 225–234.

SRA/McGraw-Hill. (2005). *SRA Open Court Reading–Grade K*. Columbus, OH: Author.

Tawney, J. W., & Gast, D. L. (1984). *Single-subject research in special education*. Columbus, OH: Merrill.

Taylor, B. M., Pearson, P. D., & Peterson, D. S. (2005). The CIERA school change framework: An evidence-based approach to professional development and school reading improvement. *Reading Research Quarterly*, *40*, 40–69.

Van den Noorgate, W., & Onghena, P. (2003a). Combining single-case experimental data using hierarchical linear models. *School Psychology Quarterly*, *18*, 325–346.

Van den Noorgate, W., & Onghena, P. (2003b). Hierarchical linear models for the quantitative integration of effect sizes in single-case research. *Behavior Research Methods, Instruments, & Computers*, *35*, 1–10.

Vellutino, F. R. (1991). Introduction to three studies on reading acquisition: Convergent findings on theoretical foundations of code-oriented versus whole-language approaches to reading instruction. *Journal of Educational Psychology*, *83*, 437–443.

Yurick, A. L. (2006). *The effectiveness of an instructional assistant led supplemental early reading intervention with urban kindergarten students*. Unpublished doctoral dissertation, Ohio State University, Columbus.

Ya-yu Lo, PhD, is an assistant professor of special education at the University of North Carolina at Charlotte. Her current interests include effective academic instruction, positive behavior interventions, challenging behavior, social skill instruction, and urban at-risk learners.

Chuang Wang, PhD, is an assistant professor of educational research at the University of North Carolina at Charlotte. His current interests include research design, program evaluation, and longitudinal data analysis.

Sherry Haskell, BA, is an early intervention therapist at Carolina Developmental Therapy Services in Union County, North Carolina. Her current interests include early childhood development, positive behavior interventions, and family collaboration.