Title:
Benefits of Structured After-School Literacy Tutoring by University Students for Struggling Elementary Readers


**Endia J. Lindo**
Texas Christian University

**Beverly Weiser**
Southern Methodist University

**Jennifer P. Cheatham**
Southern Methodist University

**Jill H. Allor**
Southern Methodist University

**Acknowledgements:**
Support for carrying out this research was provided by grant R324A160132 from the Institute of Education Sciences. The views expressed herein are those of the authors and have not been reviewed or approved by the granting agency.
Benefits of Structured After-School Literacy Tutoring by University Students for Struggling Elementary Readers

Endia J. Lindo¹, Beverly Weiser², Jennifer P. Cheatham², and Jill H. Allor²

¹Texas Christian University, Fort Worth, Texas, USA; ²Southern Methodist University, Dallas, Texas, USA

ABSTRACT
This study examines the effectiveness of minimally trained tutors providing a highly structured tutoring intervention for struggling readers. We screened students in Grades K–6 for participation in an after-school tutoring program. We randomly assigned those students not meeting the benchmark on a reading screening measure to either a tutoring group or a control group. Students in the tutoring group met twice per week across one school year to receive tutoring from non–education major college students participating in a service-learning course. The goal of this study was to determine whether tutors without prior teaching experience or instruction could improve student reading outcomes with minimal training, a structured reading curriculum, and access to ongoing coaching. Tutored students displayed significantly more growth than control students in letter-word identification, decoding, and passage comprehension, with robust effect sizes of 0.99, 1.02, and 0.78, respectively. We discuss the implications and limitations of these findings.

According to the 2015 National Assessment of Educational Progress report, only 36% of fourth graders are considered proficient readers and 31% lack basic reading skills (National Center for Education Statistics, 2015). These numbers are even more discouraging when one includes additional risk factors such as economic disadvantage (44% of students who qualify for free lunch scored below basic) or limited language proficiency (68% of English language learners scored below basic; National Center for Education Statistics, 2015; Slavin, Lake, Davis, & Madden, 2011). Despite a variety of government initiatives designed to address the relatively large proportion of students who struggle with reading in the United States, there has been minimal improvement in the reading trends for these students (National Center for Education Statistics, 2015). In the face of these disappointing overall reading trends there remains cause for hope. Considerable research has been conducted in the past few decades in an effort to advance understanding of reading development, instruction, and intervention (Slavin et al., 2011). From this work, it has been established that difficulties in phonological awareness and word-level reading often serve as a significant hindrance to early reading outcomes (Adams, 1990, 2013; Smith, Simmons, & Kame’enui, 1998; Snow, Burns, & Griffin, 1998; Torgesen, 2004).

For many students who are at risk for reading failure, school success can be increased if students receive systematic and direct instruction in achieving the basic skills necessary for reading (Mathes et al., 2005; Scanlon, Vellutino, Small, Fanuele, & Sweeney, 2005; Torgesen, 2004). Substantial evidence exists demonstrating that the provision of early intervention addressing phonological and alphabetic skills for those at risk for reading difficulties fosters reading development (Torgesen, 2004; Vadasy & Sanders, 2008). If these early literacy skills are not addressed, students struggle to
build fluency, which in turn negatively impacts comprehension (Torgesen, 2004). Unfortunately, addressing these deficits in struggling readers requires instruction that is more explicit and intensive than is typically provided in classroom settings (Torgesen, Rashotte, Alexander, Alexander, & MacPhee, 2003). Two challenges teachers face in providing struggling readers with the necessary interventions during typical class time are their ability to (a) adequately differentiate instruction to meet individual needs and (b) provide sufficient practice with immediate feedback, particularly in reading text that is at the student’s instructional level. One means of providing individualized instruction and practice opportunities is through supplemental tutoring.

**Benefits of tutoring**

Tutoring instruction, specifically one-to-one instruction supplementing core classroom instruction, is generally considered the gold-standard model for intervening with struggling readers to increase their reading achievement (Elbaum, Vaughn, Hughes, & Moody, 2000; Lauer et al., 2004; Ritter, Barnett, Denny, & Albin, 2009; Slavin et al., 2011; Snow et al., 1998; Torgesen et al., 2001). An extensive body of work by researchers has repeatedly supported the efficacy of one-to-one phonics-based tutoring for improving outcomes for at-risk (based on performance on literacy screeners) elementary students (Vadasy, Jenkins, Antil, Wayne, & O’Connor, 1997a, 1997b; Vadasy, Jenkins, & Pool, 2000; Vadasy, Sanders, & Abbott, 2008; Vadasy, Sanders, & Peyton, 2005, 2006; Vadasy, Sanders, Peyton, & Jenkins, 2002).

One major benefit of tutoring is that low achievers can spend additional time engaged in instructional activities and receive frequent and immediate corrective feedback on their progress (Miller, 2003; Morris, Tyner, & Perney, 2000; Smartt & Glaser, 2010). Another benefit of tutoring is that instruction can target individual student needs (Ehri, Dreyer, Flugman, & Gross, 2007; Justice, 2006; Lane, Pullen, Hudson, & Konold, 2009). In addition, tutoring fosters a high degree of interaction between teachers/tutors and students, affording more opportunities for student response and practice (Baker, Gersten, & Keating, 2000; Carnine, Silbert, Kame’enui, & Jungiohann, 2005; Vadasy, Sanders, & Tudor, 2007). Furthermore, these benefits of tutoring provide an environment for improving students’ confidence and attitude toward reading (Ehri et al., 2007; Moore-Hart & Karabenick, 2009; Parker, Hasbrouck, & Denton, 2003).

**Tutoring by trained volunteers**

A meta-analysis conducted by Elbaum et al. (2000) revealed that supplemental one-to-one reading intervention presented by trained volunteers (including teachers and paraprofessionals) or college students was effective (weighted effect size = 0.41) at improving reading outcomes for elementary students at risk for reading failure. Ritter et al. (2009) reported similar findings in their meta-analytic review of studies from 1989 to 2004, noting that elementary and middle school students who received tutoring from volunteers (i.e., preservice teachers, college students, community members, and parents) made significantly greater gains in decoding and fluency (i.e., effect sizes = 0.26–0.45) over students who did not receive any tutoring. More recently, Slavin et al. (2011) conducted a best evidence synthesis that found a small weighted mean effect size for interventions using paraprofessionals and volunteers as tutors (0.16); these authors noted that when two specific volunteer tutor interventions providing tutoring only 1 or 2 times a week were removed from the analyses, the impact of the remaining programs was substantially higher (effect size = 0.50). These reviews demonstrate the potential benefit of volunteers serving as literacy tutors, but each also included studies in the synthesis in which the volunteers were trained education professionals.

Although individual tutoring is considered an ideal means of intervening, it is often expensive, and it can be difficult for schools to provide this type of tutoring by highly trained teachers (Slavin et al., 2011). One alternative is to supplement instruction with tutoring provided by trained volunteers. Though most of the research on the effectiveness of tutoring interventions has used certified teachers.
or paraprofessionals to present the intervention (e.g., Slavin et al., 2011; Vadasy et al., 2008), several studies have examined the use of volunteers as tutors with positive results (e.g., Allor & McCathren, 2004; Baker et al., 2000; Fitzgerald, 2001; Meier & Invernizzi, 2001; Pullen, Lane, & Monaghan, 2004; Walker, Kronick, & Diambra, 2007). The majority of these studies using volunteer tutors involved first-grade students identified as at risk based on their performance on a reading screening assessment. The exception, Walker et al. (2007), served students in Grades 3–5 using an assignment assistance model in which college student volunteers helped students complete assignments sent by the classroom teacher. Studies using volunteer tutors most often engaged tutors who were education majors, had substantial prior tutoring experience, or received a minimum of 4 hr of training in preparation for tutoring. Baker et al. (2000) was the rare study that used volunteer tutors who were not education professionals or majors and who had minimal training (<4 hr) or experience in tutoring. These authors recruited community businesspeople to serve as tutors for first-grade students attending a Title I school. These students were selected and screened based on teacher referral, and tutors received minimal training in providing the scripted tutoring program, which was provided during the students’ school day. The outcome of this tutoring was moderate effects across reading skills (0.42–0.53) at the end of tutoring and small to moderate effects at the 1-year follow-up (0.32–0.53).

**Unanswered questions regarding effective tutoring**

Although the aforementioned studies and reviews summarized the findings from many tutoring programs, these programs varied on important variables. Some programs were more intensive than others (i.e., in total hours of instruction) and used curricula that varied in structure and the amount of training/expertise required of the tutor. Some programs provided tutors with scripted lessons and specific materials, whereas others only provided general guidelines, requiring tutors to create lesson plans and select materials. Across studies tutors were provided with varying amounts of training (3–33 hr) and guidance in lesson structure (i.e., from unscripted to highly scripted). Furthermore, tutoring amounts varied from 8 to 36.5 hr. Despite this variance, each of these studies reported positive outcomes, with moderate to large effects on students’ reading skills. Slavin et al.’s (2011) best evidence synthesis found that those tutoring interventions that emphasized phonics showed the greatest improvement in outcomes. They also noted that although small-group tutoring in phonics was effective ($d = 0.31$), presenting tutoring one to one had a slightly larger impact ($d = 0.39$). Questions remain regarding factors such as how much structure and training volunteers need to be effective and how much tutoring is required to realize benefits, highlighting the need to better understand the conditions in which minimally trained tutors are likely to make the most impact.

**Overview of the study**

The current study adds to the body of knowledge on tutoring for struggling readers in Grades K–6 by tutors who are minimally trained. This study expands on the current literature by (a) examining a model of tutoring intervention using unpaid tutors who are not professionals or education majors, (b) serving a broader range of student grade levels than in many studies, and (c) serving these students outside of the school day and building. Specifically, we addressed the following research question: Can a highly structured after-school reading intervention program have a positive impact on the reading outcomes of students in Grades K–6 who are at risk for reading failure when implemented by minimally trained tutors under the supervision of reading professionals? This study contributes to the research literature because it analyzes whether minimally trained tutors can (a) be effective tutors for struggling elementary students even though they receive very little training; (b) follow a highly structured reading curriculum without previous teaching training and instruction; and (c) be successful in improving the reading performance of students in kindergarten through sixth grade who are at risk because of low income, language difference, and poor reading performance.
Method

Participants

Site selection
The tutoring intervention took place in a free, privately funded after-school program that was housed in two low-income apartment complexes in a large urban city in the Southwest. This program was designed to provide a safe after-school environment and academic support to students from low-income families. These locations were chosen because (a) they included many high-need elementary school-age students; (b) they were close enough to the university to make the study feasible for student volunteers; and (c) the instructors at the university had a prior relationship with the program, which increased the likelihood that the program would support research efforts such as collecting permission slips and scheduling.

Student participants
All study participants attended the after-school program and a public elementary school and were being taught to read in English. Personnel in the after-school program distributed consent letters to parents of all students in the after-school program, and parental consent for 66 students was obtained. All students with consent in kindergarten through sixth grade were administered reading measures to determine their eligibility for the tutoring program. At the time of the screening, all students also gave their assent to participate. Kindergarten and first-grade students were screened with measures of phonemic awareness and the ability to sound out nonsense words. Second through sixth graders were screened with measures of their ability to sound out nonsense words and their oral reading fluency (see descriptions below). A total of 39 students who did not meet published reading benchmarks (see Good, Gruba, & Kaminski, 2002) qualified for the study. Students were ranked based on their screening scores, paired with the student with the most similar scores, and then randomly assigned to either the treatment or control group. Ten students left the study because they stopped attending the after-school program, which resulted in 16 treatment students and 13 control students the first semester. At the beginning of the second semester, eight students who were new to the program were screened. Five students qualified for the intervention and were randomly assigned to the treatment or control group. The final sample included 19 students in the treatment group and 15 students in the control group. The treatment group was composed of 53% males (n = 10); 74% of the students were Hispanic (n = 14) and the other 26% were African American (n = 5). Similarly, slightly more than half of the control group students were male (53%, n = 8); 67% of the students were Hispanic (n = 10), 27% were African American (n = 4), and 7% were other (n = 1). Chi-square analyses revealed no statistically significant differences between groups on sex ($\chi^2 = 0.002$, $p = 0.97$), race ($\chi^2 = 1.33$, $p = 0.52$), or grade level ($\chi^2 = 6.49$, $p = 0.37$), which indicates that the treatment and control groups had similar distributions of sex, race, and grade level.

Tutor participants
A total of 19 undergraduate college student tutors participated in the study during the first semester, and 20 undergraduate college tutors participated during the second semester. Students volunteered by enrolling in an elective course on the impact of illiteracy on society that included an academic service-learning component. The purpose of service-learning is to enhance learning through providing meaningful community service. In this case, tutoring enhanced the college students’ understanding of the challenges and consequences of illiteracy. Each tutor was assigned a student from the treatment group. In order to serve the students in the after-school program while maintaining the integrity of the study, tutors also provided math tutoring or assistance with nonliteracy homework to students who were not a part of the study or were in the control group. None of the tutors had completed undergraduate courses related to reading instruction, and none were education majors. Tutors attended four 1-hr training sessions as part of their coursework (details described in “Procedures”),
completed projects related to literacy, wrote reflection reports on literacy articles, and tutored their students twice each week. Each tutor received 3 hr of course credit for participation.

**Measures**

**Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Good & Kaminski, 2002)**

Three DIBELS subtests, Phoneme Segmentation Fluency (PSF), Nonsense Word Fluency (NWF), and Oral Reading Fluency (ORF), were used in this study. Students were screened using the first DIBELS benchmark assessment for their grade level. In addition, students in second through sixth grades were also administered the NWF measure to assist in matching the students prior to random assignment to treatment or control. All subtests were individually administered following standard procedures.

PSF measures a student’s ability to segment two-, three-, and four-phoneme words into their individual phonemes. A student’s PSF score is the total number of phonemes segmented correctly within 1 min. PSF was also administered to kindergarten students as a screening measure. The reliability coefficient (alternate form) is 0.88 for a single probe and 0.96 for the mean of five probes. Concurrent and predictive validity with other reading assessments range from 0.45 to 0.68. To minimize the possibility of practice effects, alternate versions of the assessment were given at pretest and posttest.

NWF measures a student’s ability to sound out three-phoneme nonwords (Good et al., 2002). The student is shown a list of nonsense words containing one short vowel and one or two consonants (e.g., sim, lut) and is asked to say the sound for each letter or read the entire word. The total NWF score is the number of letter sounds pronounced correctly in 1 min. The reliability coefficient (alternate form) is 0.92 for a single probe and 0.98 for the mean of five probes. Predictive and concurrent validity with a variety of reading assessments range from 0.59 to 0.82. NWF was also given to all students as a screener. Alternate forms were given at screening, pretest, and posttest.

ORF is a test of reading accuracy and fluency with grade-level-appropriate connected text. The student’s score is the number of words read correctly in 1 min. Words omitted, words substituted, and hesitations of more than 3s are scored as errors, whereas self-corrected words are scored as accurate. Each grade level has different reading passages with grade-level-appropriate text. Test–retest reliabilities for elementary students range from 0.92 to 0.97, whereas alternate-forms reliability of different reading passages drawn from the same level ranges from 0.89 to 0.94. ORF was also administered to students in Grades 2–6 as a screening measure and to students in Grades 1–6 as a pretest and posttest measure. Alternate forms of ORF were used for the screener, pretest, and posttest.

**Woodcock Language Proficiency Battery–Revised (WLPB–R; Woodcock, 1991)**

Three subtests of the WLPB–R were also administered as pretest and posttest measures: Letter-Word Identification, Word Attack, and Passage Comprehension. Students in the treatment and control groups were given the pretest upon entry into the study during the same week. Students were administered an alternate form of the test at the end of the school year.

The Letter-Word Identification subtest requires the student to pronounce words presented in isolation. The Word Attack subtest requires the student to pronounce nonsense words. The Passage Comprehension subtest has students read a short sentence or passage and identify one word that is missing. For all three subtests, testing procedures were followed as outlined in the testing manual. Reliability coefficient alphas range from 0.90 to 0.92. Test–retest coefficients range from 0.82 to 0.96. More information on these subtests can be found in the technical manual.

**Procedures**

**Data collection**

A senior researcher, the course professor, and two research assistants administered all assessments and conducted all fidelity observations. All were experienced reading teachers with at least a master’s
degree in education and expertise in reading. All measures were administered in a quiet room at the
tutoring location during a 2-week time period at the beginning of each semester. Prior to testing,
the assessors agreed that they would discontinue testing if students displayed any signs of fatigue;
however, none of the students required more than one testing session.

**Tutor training and supervision**
Each semester, tutors participated in four 1-hr classroom instructional sessions that included
intervention modeling, practice, and feedback led by one of the researchers, who also served as the
professor of the class. Detailed procedures for each routine were demonstrated, and students prac-
ticed following lesson scripts with one another. The professor also modeled correct pronunciation
of sounds and how to model blending sounds. Three of the training sessions occurred prior to the
beginning of treatment and one was later in the semester to discuss questions, concerns, and
implementation issues. During tutoring sessions, tutors were assisted as needed by the professor of
the class or one of two graduate research assistants (one for each semester), all of whom had expertise
in reading instruction; therefore, in each semester we had a ratio of two supervisors to approximately
20 tutors. Typical assistance included brief reminders to follow the scripted procedures, modeling
correct pronunciation of sounds and blending, and brief reminders to provide corrective feedback.
Lessons were followed in a predetermined sequence, so the tutors did not need to choose which lesson
or activity to do each day. They simply moved on to the next lesson. Occasionally tutors were
instructed to repeat lessons if a student was observed to be having difficulty. Tutors kept daily field
notes. These notes included lessons completed, book titles and page numbers read, and comments
about student progress. Tutors also kept reflective journals of their daily experiences that were read
by the professor and research assistants, who provided comments and suggestions.

**Overview of tutoring intervention and implementation**
Tutoring sessions were scheduled to occur twice each week (except for weeks that included one of
the four tutor training sessions) across one school year for approximately 25–45 min per session
(approximately 16 weeks per semester). We selected two times per week because this fit within a
reasonable schedule for a college course and because we had found positive effects in previous studies
when tutoring occurred approximately two times per week (Allor & McCathren, 2004). Each session
included phonics practice, fluency building, comprehension, and supplemental activities. The phonics
portion of the lesson was composed of instructional routines from Bookshop Phonics (Allor &
Minden-Cupp, 2007). At the beginning of the first semester, placement tests were administered by
the research assistants to place students in the appropriate phonics lesson and level of book reading
routines (see descriptions in the next sections). These placement tests and oral reading fluency scores
were also used to match students to appropriate texts for book reading routines. Table 1 provides
intervention details for each student, including where the student was placed in the curriculum
and the number of sessions completed. If students successfully completed the phonics lessons or
demonstrated mastery of the skills, then the phonics portion of the lessons was discontinued, leaving
more time for book reading routines. As Table 1 specifies, students in Grades K–3 typically began at
the beginning of the phonics lessons and participated in the early book reading routines (described in
the next section), whereas students in Grades 4 and 5 began a little later in the lessons. One fifth-grade
student needed to start early enough in the lessons to review short vowels; only one student, a sixth
grader, did not need any of the phonics lessons.

The professor and a research assistant (one per semester) monitored the selection of text to ensure
that students were reading on an instructional level (approximately 90% accuracy) during tutoring
sessions. The research assistant monitored student performance during the lessons and ensured
that phonics lessons were repeated if needed. The research assistant occasionally directed tutors to
skip some phonics lessons if the students did not require as much repetition as was provided (lessons
provided extensive cumulative review that was not necessary for all students).
Phonics activities. The first 15–20 min of the session were devoted to phonics activities. These were composed of sequenced charts and practice pages that included letter-sound practice, reading lists of words, and connected text. For example, in an early lesson the students practiced sounding out simple words (e.g., pan, nap, fast), reading high-frequency words (e.g., look, see, little), and reading simple sentences (e.g., The little cat is fast.) In later lessons, more complex words (e.g., smooth, point, frown) were sounded out, and short passages were read that included taught words. All items were provided, so tutors did not need to make decisions about which activities they should conduct or which words or passages to practice. Tutors were provided with scripted directions to use with the sequenced charts and practice pages. Each lesson typically included brief scripted activities or routines for introducing new sounds, reviewing previously taught letter/sound combinations, reading multisyllabic words or words with inflectional endings, reading sentences, and reading passages. Each activity was short and fast paced to increase student motivation and participation.

Although lessons increased in difficulty as students progressed through the program, the basic structure included several consistent routines that allowed students to become familiar with the curriculum and work at a fast pace, with a minimum amount of time spent on providing directions. The repetitive nature of the routines was an advantage for tutors as well, as tutors quickly became familiar and confident with the routines. Lessons began with the chart, which included letters and words for students to read. The first activity or routine on the chart was the introduction of a new sound. The second activity was sounding out words using the decoding strategy (described below). The third activity was practicing reading sight words. The last activity on the chart presented words in parts, separated by syllables or base word and inflectional ending. Students read the parts separately and then read the whole word.

The primary decoding strategy, which was used in every lesson, was referred to as sound and say. Words were listed on a chart, with letter combinations underlined to scaffold sounding out the word (e.g., smooth). The tutor modeled sounding out and saying the first word. Next students said the underlined sounds and pointed to the underlined letters. Then students said the word. Students repeated the process with all of the words on the chart for the sound-and-say activity. Students used this strategy when reading the words in the sound-and-say section of the chart and also when reading connected text. Students were encouraged to use this sound-and-say strategy any time they came

### Table 1. Intervention details by student participant.

<table>
<thead>
<tr>
<th>Student</th>
<th>Grade</th>
<th>Placement</th>
<th>Number of sessions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>K</td>
<td>Beginning phonics/early book reading activities</td>
<td>31</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>Beginning phonics/early book reading activities</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>Beginning phonics/early book reading activities</td>
<td>32</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>Beginning phonics/early book reading activities</td>
<td>29</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>Beginning phonics/early book reading activities</td>
<td>29</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>Beginning phonics/early book reading activities</td>
<td>30</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>Beginning phonics/early book reading activities</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>Beginning phonics/early book reading activities</td>
<td>29</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>Beginning phonics/early book reading activities</td>
<td>27</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>Beginning phonics/late book reading activities</td>
<td>15</td>
</tr>
<tr>
<td>11</td>
<td>3</td>
<td>Beginning phonics/late book reading activities</td>
<td>28</td>
</tr>
<tr>
<td>12</td>
<td>3</td>
<td>Beginning phonics/late book reading activities</td>
<td>32</td>
</tr>
<tr>
<td>13</td>
<td>3</td>
<td>Lesson 19 phonics/late book reading activities</td>
<td>29</td>
</tr>
<tr>
<td>14</td>
<td>4</td>
<td>Lesson 48 phonics/late book reading activities</td>
<td>29</td>
</tr>
<tr>
<td>15</td>
<td>4</td>
<td>Lesson 48 phonics/late book reading activities</td>
<td>32</td>
</tr>
<tr>
<td>16</td>
<td>5</td>
<td>Lesson 48 phonics/late book reading activities</td>
<td>31</td>
</tr>
<tr>
<td>17</td>
<td>5</td>
<td>Lesson 19 phonics/late book reading activities</td>
<td>26</td>
</tr>
<tr>
<td>18</td>
<td>6</td>
<td>Lesson 34 phonics/late book reading activities</td>
<td>28</td>
</tr>
<tr>
<td>19</td>
<td>6</td>
<td>No phonics/late book reading activities</td>
<td>29</td>
</tr>
</tbody>
</table>

Note. K = kindergarten.

*In the second semester, phonics discontinued spending an entire session on book reading activities. *\(^a\)Chart 19 begins with a review of short vowel sounds. *\(^b\)Chart 48 begins with advanced vowel patterns. *\(^c\)Chart 34 begins with an introduction to long vowel patterns.
across unknown words during reading in order to help them transfer skills to reading beyond the tutoring sessions.

After completing the scripted routines with the charts, students read from a corresponding practice page (Allor & Minden-Cupp, 2007) that included cumulative practice of taught words (in a list) and a few sentences or brief paragraphs of connected text. No new letter sounds or skills were introduced during this portion of the lesson. The practice pages were for cumulative review and application of previously taught skills. The first activity on the practice page was reading a list of words carefully. Word lists were designed to include discriminant pairs (e.g., *hand* and *had* or *sit* and *sat*). Students were encouraged to look at each letter in the word to be sure it was read correctly. Next students read connected text in the form of either separate sentences or a story.

Another important component of Bookshop Phonics is immediate corrective feedback, which was used during all parts of the sessions. When a student made an error, the tutor immediately stopped the student and modeled the correct response. Next the student said the correct response with the tutor and then individually to ensure that the student could say the correct response. The activity was repeated until either the student could do the activity with no mistakes or the sentence with the misread word was reread, depending on the type of activity. This corrective feedback routine was used for any mistake a student made, even during book reading or game playing.

**Early book reading routines.** The next 10–20 min of the lesson focused on book reading. Students read texts at their reading level (instructional level of approximately 90% accuracy, which means they required assistance on no more than 1 out of every 10 words in a text). A variety of trade books were used in addition to leveled readers. Students in Grades K–3 typically began with the early book reading routines (see Table 1). Before reading each book, the tutor would lead the student through the book, encouraging the student to make predictions about what was happening on each page. Then the tutor read two pages and the student read the same two pages to support fluency and comprehension. Students received corrective feedback when needed, with tutors helping students to decode unknown words.

**Later book reading routines.** Students in Grades 4–6 typically began with the later (i.e., more advanced) book reading routines (see Table 1). Also, as students in earlier grades made progress, they began completing later book reading routines during each lesson (instead of the early book reading routines). These more advanced activities consisted of making predictions, reading a book, stopping periodically to check predictions, and practicing a passage to increase fluency. Additional books that already belonged to the after-school program were coded for readability level by one of the researchers. Students chose books that appealed to their interests and were within their reading level. Book reading activities were based on activities successful in another tutoring study (Allor & McCathren, 2004) and similar to peer tutoring procedures (Fuchs, Fuchs, Mathes, & Simmons, 1997). Tutors used scripted prompts to guide students as they made and checked predictions and constructed main idea statements in 10 words or fewer. As with all activities, tutors provided corrective feedback when students misread a word.

**Additional supplemental learning games.** Depending on the amount of time available, the last 5–15 min of each lesson included a variety of games. Several supplemental materials were used, including a previously validated and published phonemic awareness game called the *Stop and Go Phonemic Awareness Game* (Allor, Gansle, & Denny, 2006; Allor & Mathes, 2012; Allor & McCathren, 2004) and other games created by the researchers. The supplemental materials were used at the end of each lesson. In the beginning lessons, students played the *Stop and Go Game* at the conclusion of each lesson. The *Stop and Go Game* was a fun, motivating way for students to practice phonemic awareness (blending and segmenting) and letter-sound knowledge (see Allor et al., 2006, for details). During the game, the student and tutor took turns drawing a letter card and saying the taught sound. If
the sound was a stop sound (i.e., a sound that must be said quickly; e.g., /t/, /d/), the student pronounced the sound quickly. If the sound was a go sound (i.e., a continuous sound that can be held without distortion; e.g., /mmm/ or /sss/), the student stretched the sound for about 2 s. Once enough letters had been identified to make a word, the student practiced either blending the letter sounds to say a word or segmenting a spoken word into its individual sounds. For blending, the tutor moved the necessary letters to a word mat and the student blended the sounds to pronounce the word. For segmenting, the tutor pronounced the word and the student moved the letters to build the word, saying each sound separately (segmenting).

All treatment students were proficient in the Stop and Go Game by the end of the first semester. When students mastered the game, which many older readers did very quickly, they spent this portion of each session either reading a book or playing a different researcher-created reading game. Games were kept in file folders or envelopes in a file box, with difficulty levels labeled so students could select games on their own levels. These games were highly engaging and motivating for the students and addressed a variety of reading skills, including building words, word sorts, high-frequency word recognition, and reading fluency. Games included traditional games such as Old Maid and Memory as well as simple trail games.

Student motivation

Throughout each session, tutors kept a point chart on which they or the student could mark off points earned for each activity. Once the point chart was filled (100 points earned), the student could play an extra reading game or receive a small tangible reward (e.g., pencil, eraser, small rubber ball). Students were awarded points for effort and participation and could earn points even when activities were completed with mistakes. The structure of the curriculum itself was motivating because each activity was short but engaging. In addition, the opportunity to play a reading game at the end of the session proved to be motivating for students. This sequencing of activities encouraged students to do their best during the structured tutoring activities so they would have time to play a game.

Context and control condition

The tutors in this study provided students in the control condition with math tutoring or assistance with nonliteracy homework. Other activities provided by the after-school program continued as usual. These included assistance with homework and nonacademic activities. The program and training provided to the tutors in this study were not provided to after-school staff. All students attended public schools and were being taught to read in English.

Results

Dosage

Tutoring sessions lasted 25–45 min depending on student motivation, ability, and availability. Tutors were required to make up missed tutoring sessions. The average number of sessions was 27.26 (SD = 5.41). The 16 students who were tutored for both semesters received 26–32 sessions (averaging 29 sessions), approximately 15–19 hr of instruction across one school year. The additional three students who were added to the treatment group in January received 15 or 16 sessions, approximately 9 hr of instruction. One student demonstrated mastery of the phonics skills on the placement test and therefore spent the entire tutoring session participating in the advanced book reading routines. In the second semester, seven students discontinued the phonics portion of the instruction because they had either completed the lessons or demonstrated mastery of the skills; these students spent the entire session participating in advanced book reading routines and supplemental games (see Table 1).
Fidelity

Twice each semester tutoring sessions were observed by one of two research assistants using fidelity observation forms to determine the level of implementation fidelity (13.5% of sessions). Components observed were (a) appropriate corrective feedback and scaffolding; (b) organizational skills and time management; (c) following of procedures outlined in the script, including correct modeling of sounds and decoding strategies; and (d) inclusion of verbal praise and positive reinforcement. Implementation of each component was evaluated on a 3-point scale (the behavior occurred during most of the lesson, some of the lesson, or little/none of the lesson). After each observation, each tutor was given feedback and suggestions for future tutoring sessions. Fidelity of implementation over both semesters ranged from 84% to 100%, with the average being 96.3%. The observation instrument was relatively straightforward. Research assistants were trained in the observation instrument by the professor; however, reliability observations were not conducted because of limited resources.

Pretreatment equivalences

Pretreatment differences were analyzed using separate t tests. There were no statistically significant differences between groups for any of the measures. Means and standard deviations for the treatment and control groups on the pretest measures are presented in Table 2.

Posttreatment achievement changes

Separate t tests were used to analyze student gain scores from pretest to posttest on each measure. Means, standard deviations, and difference scores for both groups on all measures can be found in Table 2. Cohen’s d effect sizes were computed to measure how much the mean gain of the treatment group exceeded the mean gain of the control group at posttest in standard deviation units (Glass, McGraw, & Smith, 1981).

Differences between gain scores for the treatment group and control group were statistically significant in favor of the treatment group for all three subtests of the WLPB–R (Letter-Word Identification, Word Attack, and Passage Comprehension), with robust effect sizes of 0.99, 1.02, and 0.78, respectively. On all three posttests, treatment students experienced greater gains in these reading areas than students in the control group who did not receive tutoring.

Growth on DIBELS progress monitoring measures

Because of the range of grades in the sample, summarizing data on DIBELS measures would not be appropriate, as no single measure would best represent growth across kindergarten through sixth grade. Note that students were administered the ORF subtest appropriate for their grade level. Although phoneme segmentation fluency and nonsense word fluency did not vary across grades, they were not expected to grow once mastery was achieved, which typically occurs in kindergarten and first grade. As can be seen in Table 3, growth occurred across measures and across grade levels, including oral reading fluency for students in first through sixth grades. Twelve out of the

<table>
<thead>
<tr>
<th>Measure</th>
<th>Treatment (n = 19)</th>
<th>Control (n = 15)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pretest</td>
<td>Posttest</td>
</tr>
<tr>
<td>Word Identification</td>
<td>91.74</td>
<td>112.05</td>
</tr>
<tr>
<td>Word Attack</td>
<td>91.11</td>
<td>111.58</td>
</tr>
<tr>
<td>Passage Comprehension</td>
<td>85.63</td>
<td>99.32</td>
</tr>
</tbody>
</table>

Note. Data are standard scores. ES = effect size.
*p < .05. **p < .01.
18 (67%) first- through sixth-grade students in the treatment group grew at least 32 words per minute (one word per week of intervention), whereas six of the 12 (50%) first- through sixth-grade students in the control group grew at least 32 words across the same period of time (see bold scores in Table 3).

<table>
<thead>
<tr>
<th>Group</th>
<th>Student/grade</th>
<th>Oral Reading Fluency</th>
<th>Phoneme Segmentation Fluency</th>
<th>Nonsense Word Fluency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1/K</td>
<td>0</td>
<td>32</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>2/first</td>
<td>14</td>
<td>26</td>
<td>39</td>
</tr>
<tr>
<td>1</td>
<td>3/first</td>
<td>35</td>
<td>38</td>
<td>42</td>
</tr>
<tr>
<td>1</td>
<td>4/first</td>
<td>11</td>
<td>0</td>
<td>49</td>
</tr>
<tr>
<td>1</td>
<td>5/first</td>
<td>72</td>
<td>45</td>
<td>39</td>
</tr>
<tr>
<td>1</td>
<td>6/first</td>
<td>40</td>
<td>24</td>
<td>49</td>
</tr>
<tr>
<td>1</td>
<td>7/second</td>
<td>68</td>
<td>54</td>
<td>43</td>
</tr>
<tr>
<td>1</td>
<td>8/second</td>
<td>53</td>
<td>8</td>
<td>23</td>
</tr>
<tr>
<td>1</td>
<td>9/second</td>
<td>69</td>
<td>32</td>
<td>24</td>
</tr>
<tr>
<td>1</td>
<td>10/third</td>
<td>87</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>11/third</td>
<td>105</td>
<td>57</td>
<td>51</td>
</tr>
<tr>
<td>1</td>
<td>12/third</td>
<td>109</td>
<td>14</td>
<td>36</td>
</tr>
<tr>
<td>1</td>
<td>13/third</td>
<td>112</td>
<td>21</td>
<td>56</td>
</tr>
<tr>
<td>1</td>
<td>14/fourth</td>
<td>107</td>
<td>28</td>
<td>118</td>
</tr>
<tr>
<td>1</td>
<td>15/fourth</td>
<td>128</td>
<td>37</td>
<td>36</td>
</tr>
<tr>
<td>1</td>
<td>16/fifth</td>
<td>152</td>
<td>43</td>
<td>53</td>
</tr>
<tr>
<td>1</td>
<td>17/fifth</td>
<td>113</td>
<td>56</td>
<td>43</td>
</tr>
<tr>
<td>1</td>
<td>18/sixth</td>
<td>108</td>
<td>20</td>
<td>39</td>
</tr>
<tr>
<td>1</td>
<td>19/sixth</td>
<td>132</td>
<td>53</td>
<td>111</td>
</tr>
<tr>
<td>0</td>
<td>20/K</td>
<td>15</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>21/K</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>22/K</td>
<td>0</td>
<td>40</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>23/first</td>
<td>46</td>
<td>25</td>
<td>37</td>
</tr>
<tr>
<td>0</td>
<td>24/second</td>
<td>85</td>
<td>35</td>
<td>24</td>
</tr>
<tr>
<td>0</td>
<td>25/second</td>
<td>91</td>
<td>25</td>
<td>46</td>
</tr>
<tr>
<td>0</td>
<td>26/third</td>
<td>126</td>
<td>15</td>
<td>28</td>
</tr>
<tr>
<td>0</td>
<td>27/third</td>
<td>73</td>
<td>50</td>
<td>36</td>
</tr>
<tr>
<td>0</td>
<td>28/third</td>
<td>95</td>
<td>21</td>
<td>41</td>
</tr>
<tr>
<td>0</td>
<td>29/third</td>
<td>120</td>
<td>40</td>
<td>27</td>
</tr>
<tr>
<td>0</td>
<td>30/first</td>
<td>84</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>0</td>
<td>31/fifth</td>
<td>72</td>
<td>49</td>
<td>59</td>
</tr>
<tr>
<td>0</td>
<td>32/fifth</td>
<td>98</td>
<td>41</td>
<td>47</td>
</tr>
<tr>
<td>0</td>
<td>33/fifth</td>
<td>108</td>
<td>49</td>
<td>20</td>
</tr>
<tr>
<td>0</td>
<td>34/fifth</td>
<td>94</td>
<td>41</td>
<td>45</td>
</tr>
</tbody>
</table>

Note. K = kindergarten.  
1 = treatment, 0 = control. Bold scores indicate growth of 32 words or more. Bold scores indicate mastery met (35). Bold scores indicate mastery met (50).

**Discussion**

In this small randomized controlled trial, we examined the effectiveness of minimally trained tutors for struggling readers in Grades K–6 when tutors were provided with supervision and a structured phonics-based and book reading tutoring program. Specifically, we found that college students with only minimal training and supervision from a research assistant effectively implemented the program. Students in Grades K–6 who participated in the tutoring significantly outperformed students who did not receive the intervention. This study extends our prior research (Allor & McCathren, 2004) in which we implemented these same routines with struggling first graders by demonstrating the effectiveness of this program with students ranging from kindergarten to sixth grade.

Although interrater reliability observations were not performed, reading experts did conduct observations to measure the fidelity of implementation. These experts reported that the tutors in our study implemented the tutoring program with fidelity despite receiving minimal reading instruction training. Tutors could use the scripted prompts to provide structured and systematic instruction.
with relatively little training and no previous education classes. However, the professor and research assistants provided guidance and support as needed, which may be an important aspect of after-school tutoring with minimally trained tutors. Of particular note is that the professor and research assistants were responsible for initial placement in the program and monitored regularly to ensure that individual students advanced appropriately in the phonics lessons and read books on their instructional level.

Statistically significant differences were found in favor of the students receiving tutoring on measures of word identification, word attack (i.e., phonics), and comprehension. Pre- and posttest scores on progress monitoring measures of phonemic awareness (DIBELS PSF), alphabetic skills (DIBELS NWF), and oral reading fluency (DIBELS ORF) demonstrated educationally relevant growth for many students in the treatment group and fewer students in the control group. These scores were not averaged, as many students in the higher grades, as expected, had already mastered DIBELS PSF and NWF at pretest, and ORF passages varied according to grade level. In sum, students who received tutoring (approximately 27 sessions across one school year) made gains in reading ability that were statistically greater than those made by untutored students. Furthermore, these differences appear to be educationally meaningful.

These findings are consistent with other research demonstrating that tutoring has positive academic benefits for low-achieving and at-risk students (e.g., Elbaum et al., 2000; Slavin et al., 2011). Our tutoring program included features that have also been found to be effective in other studies, including immediate corrective feedback on performance, high levels of opportunities to respond, and instruction at the student’s current instructional level, all of which contribute to a positive environment and increased self-confidence (Allor & McCathren, 2004; Baker et al., 2000; Miller, 2003; Morris et al., 2000; Pullen et al., 2004; Vadasy et al., 2007). Supervision of tutoring by research assistants with expertise in reading was also likely critical to ensuring modeling of correct sound pronunciation and blending, appropriate placement and pacing through the curriculum, as well as selection of texts at each student’s instructional level.

Limitations

Several limitations should be considered when interpreting these findings. First, the overall number of participants in the study was relatively small, and participants were from a convenience sample; however, effect sizes were large enough that even with limited power statistically significant differences were detected. Furthermore, random assignment to condition improved internal validity, and careful screening increased the likelihood that the findings would be generalizable to similar students. Second, the number of students in each grade was small, which makes it difficult to determine conclusively that the program is effective across all grade levels. However, student-level data indicate that meaningful growth occurred across grade levels. Third, the program was not coordinated with instruction taking place in each child’s school, and no social validity data were collected. It is possible that tutoring programs that are closely aligned with classroom instruction would be more effective and perceived as more beneficial. Fourth, in our study we did not compare different amounts of tutoring received or the level of supervision required. Fifth, because of limited resources, a limited number of sessions were observed for fidelity, and interrater reliability was not examined; however, the observations conducted showed high levels of treatment fidelity. Formal observations in the control condition were not conducted; training and lessons were not provided to the after-school program staff to minimize the possibility that students in the control group would receive tutoring components. All students attended public school and received reading instruction in English, but additional information about typical reading instruction was not gathered. Finally, data collectors were not blind to condition.

Educational implications

Although this after-school tutoring study was small in scale, we are encouraged that with only 27 sessions (on average) provided by minimally trained tutors we found statistically significant
differences between the treatment and control groups, with robust effect sizes of 0.78, 0.99, and 1.02, for word identification, word attack, and passage comprehension, respectively. These differences were seen with a relatively small amount of tutoring (i.e., 16–19 hr) over one school year. The lessons were very structured, which means that they followed a systematic phonics scope and sequence, implemented routine book reading procedures (e.g., making predictions, summarizing), and included explicit instruction with corrective feedback. Lessons focused on major components of literacy instruction. Book reading allowed students to apply previously learned letter/sound combinations in connected text and also afforded students time to enhance their fluency, vocabulary, and comprehension skills. Most important, the minimally trained tutors quickly learned the repetitive teaching formats, book reading questioning techniques, and error correction prompts and could paraphrase directions slightly without lowering treatment fidelity. In addition to the support provided by the structure of the program, the college instructor or research assistant was available to provide ongoing support and assistance. It is likely that this ongoing support was a key component of the effective use of these minimally trained tutors; however, at any given time we had only one or two supervisors to approximately 20 tutors. That said, tutoring by minimally trained tutors who are provided with supervision by a certified teacher or other trained professionals would likely be more feasible and definitely more cost effective than hiring all highly trained or certified teachers to provide this one-to-one instruction.

**Future research**

Feasible methods for helping struggling readers are in high demand. The current study used a combination of an existing after-school program; a structured, scripted curriculum; minimally trained tutors; and tutor support to provide targeted reading instruction. However, more research is needed to investigate which aspect or aspects of the intervention contributed the most to student gains. Knowing which aspects, such as the curriculum or tutor support, are key can help educators know how to allocate funds and other resources. Research with larger samples is needed, particularly for students in upper grades. Further research is also needed on how to develop larger scale after-school tutoring programs.

**Conclusion**

This study contributes to the research base on the effects of after-school tutoring on struggling readers with multiple risk factors (i.e., low income and racial minority status). These findings should encourage policymakers, educators, and researchers to further explore using minimally trained tutors as an effective and feasible strategy for improving the reading skills of elementary students at risk for reading failure. This after-school program had significant positive effects on reading performance favoring the treatment group over the control group in areas of (a) decoding real words and pseudowords and (b) comprehension, even though the tutors received very little training in how to tutor students struggling with reading problems. Some factors that we believe contributed to the success of the program were focusing on high-priority word recognition skills and providing the opportunity to apply those skills to text. Most likely, aligning tutoring content and routines with primary literacy instruction provided by a child’s school would yield even stronger results. Ideally, the amount of tutoring would be adjusted to fit the needs of students.

**References**


BENEFITS OF STRUCTURED TUTORING