Abstract: The purpose of this study was to compare the effect of two theoretically different reading programs on the reading outcomes of first-grade students (N = 107). Two elementary schools were chosen for participation based on the first-grade reading program currently being implemented in the schools. One school used the Horizons Fast Track A-B reading program and the other used a Guided Reading approach. Students were assessed on Phoneme Segmentation Fluency, Nonsense Word Fluency, and Oral Reading Fluency using the AIMSweb (Edformation, 2006) progress monitoring system. Results showed differing responses to the interventions. Students in both conditions significantly increased across time on Phoneme Segmentation Fluency, Nonsense Word Fluency, and Oral Reading Fluency. However, the students in the Guided Reading condition significantly outperformed the students in the Horizons condition on Phoneme Segmentation Fluency, while the students in the Horizons condition made significantly greater gains than students in the Guided Reading condition on Oral Reading Fluency. Results are discussed in regards to the differences between the two reading programs.

One of the most pressing societal issues in our country is that of teaching our children to read. The first graders of today will be adults in a global world where the literacy demands placed on them will determine, even more than today, their access to knowledge and economic success (Adams, 1990; Snow, Burns, & Griffin, 1998). Because of this concern, over 20 years of research has been conducted by the National Institute of Child Health and Human Development, the U.S. Office of Education, and many private foundations and institutes to examine how best to teach reading to ensure that all children acquire adequate progress in reading (Lyon, Alexander, & Yaffee, 1997; Torgesen et al., 2001). This goal is repeated in the No Child Left Behind Act (2001) requiring that all children should receive evidence-based reading instruction (Snow et al.).

Even though research has shown the importance of explicit and systematic teaching of phonological awareness and phonics (Adams, 1990; Cunningham, 1990; Iverson & Tunmer, 1993; National Reading Panel (NRP), 2000), many teachers and administrators take for granted that a published reading program has research evidence supporting efficacy and benefit. Unfortunately, this is not always the case. Many programs and educational methods are developed around a set of philosophical beliefs on student learning or an ideal that is identified by a marketing survey; however, these programs rarely, if ever, have controlled evidence-based research evaluating their effectiveness (Brown-Chidsey & Steege, 2005).

With research showing that waiting until middle elementary school years to implement
remedial reading programs is more costly and less effective than early intervention and prevention, it is not surprising that a major change in reading pedagogy is emerging that emphasizes early identification and prevention of reading failure (Torgeson, 2002). Cunningham and Stanovich (1997) studied the reading comprehension, vocabulary, general knowledge, and print exposure of students in 11th grade who completed a reading battery 10 years earlier in first grade and found that first-grade reading ability predicted all 11th-grade outcomes. There is compelling evidence that children who rapidly acquire reading proficiency can enhance their verbal intelligence by increasing their vocabulary and general knowledge (Cunningham & Stanovich, 1998). Further, the Connecticut Longitudinal Study (Francis, Shaywitz, Steubling, Shaywitz, & Fletcher, 1996) reported that children who are poor readers at the end of first grade do not obtain adequate reading proficiency by the end of elementary school. The prevention of reading difficulties is best accomplished through systematic early reading instruction (Foorman, 2003; Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2002). Therefore, it is important to identify those programs that maximize the development of reading skills and that lead to greater reading fluency as early as possible.

Currently, federal law promotes the adoption of educational practices based on research; in the future, widespread pedagogical practices may gradually be influenced by this legislative agenda. The Horizons program (Engelmann, Engelmann, & Seitz-Davis, 1997) is a Direct Instruction (DI) reading program. This program incorporates research findings concerning optimal decoding and comprehension strategies. For example, the Horizons Fast Track A-B program presents a diversity of story reading formats to increase understanding of characters, plot, and sequence and allows students to analyze details from the story. The Horizons reading program has four levels (A, B, C, & D) and two fast-track options (A-B & C-D).

A follow-up study of students who were instructed in a systematic and explicit phonics program in first grade showed students who received Horizons Fast Track A-B maintained and even increased their superior performance in third and fourth grades on state-mandated tests of reading and English language arts (Tobin, 2004). Gersten, Keating, and Becker (1988) documented the long-term positive impact of DI in longitudinal studies of Project Follow Through students. Adding to this, Tobin (2003) found that students in first grade who received instruction in Horizons Fast Track A-B had significantly higher reading fluency, reading accuracy, and nonsense word fluency scores, and scored higher on the Woodcock Diagnostic Reading Test (Woodcock, 1997) when compared to students who received reading instruction with a conventional basal program. The students were followed for several years to assess the longitudinal effects of the initial explicit-phonics instruction. Students who were instructed with Horizons Fast Track A-B in first grade scored significantly higher on the Massachusetts Comprehensive Assessment System (MCAS) in reading at the end of third grade and in English language arts at the end of fourth grade compared to students who received instruction in a typical basal reading program (Tobin, 2004).

Another approach to early reading is Guided Reading. It is not a formal reading program; rather, it is designed to expose students to a wide range of literature (Pinnell & Fountas, 1996). Swartz, Shook, and Klein (2004) refer to Guided Reading as a pedagogical approach in which a teacher instructs small groups of students or individual students to provide support to help them comprehend text. Reitsma (1988) compared the effects of Guided Reading, reading while listening, and reading with computer-based speech feedback on the ability of first-grade beginning readers to read a target word list. The students read a passage containing 20 target words each day for 5 consecutive days under one of the three experi-

Winter 2009
mental conditions. Students in Guided Reading and in reading with computer-based speech feedback performed better than students in the reading while listening condition. Swartz and Shook (2003) reported that Guided Reading is an effective teaching method to develop proficient readers. They indicated that long-term use of Guided Reading led to an increase in achievement test scores.

The benefits of Guided Reading to engage students in thinking about the meaning of text have been noted. However, this approach does not provide systematic instruction, practice in basic reading skills, sufficient opportunities to master basic skills, or systematic and judicious review of previously taught skills (Torgesen, 2006). Furthermore, it has been suggested that leveled books do not necessarily provide good practice for essential phonics skills (Torgesen).

Both the Horizons program and the Guided Reading approach are widely used in school systems across the country. However, it appears a school-based comparison of the two programs has never been conducted. With this in mind, this evaluation of the two different approaches to teaching beginning reading skills was conducted by the school district in an effort to meet the expectation of state and federal departments of education for improved student outcomes. First grade was chosen for this evaluation because prior research by Good III, Simmons, and Smith (1998) suggested that variations in initial reading skill acquisition determine a large proportion of subsequent reading proficiency and the trajectory of reading improvement.

Therefore, the purpose of this exploratory study was to examine the effectiveness of two theoretically different reading programs on the Phonemic Segmentation Fluency (PSF), Nonsense Word Fluency (NWF), and Oral Reading Fluency (ORF) of first-grade students. It was hypothesized that a more explicit and systematic initial reading program would result in greater PSF, NWF, and ORF. This study is a preliminary attempt at program evaluation conducted within a public school system, rather than a research study with random assignment of participants.

Methods

Participants and Settings

Students were 107 first-grade students (53 girls and 54 boys) enrolled in two elementary schools in a northeastern state. Two schools participated; one was a Title I school (n=47) using the Horizons Fast Track A-B program with a student composition of 17% African American, 2.3% Asian, 6.1% Hispanic, 70.4% Caucasian, 4.2% multiracial, 3.8% limited English speaking, and 57.6% low income and receiving free or reduced-cost lunch. The second school (n=60) used the Guided Reading approach (Fountas & Pinnell, 2000; Pinnell & Fountas, 1996) and had a student composition of 4.2% African American, 2.4% Asian, 2.4% Hispanic, 89.5% Caucasian, 1.5% multiracial, 2.6% limited English speaking, and 27.2% low income and receiving free or reduced-cost lunch.

Teachers. The teachers in this study were Caucasian females with an average of 17.1 years of teaching experience. Four of the participating teachers had master’s degrees and two had bachelor’s degrees. All teachers were certified as highly qualified in reading under NCLB legislative requirements. Three teachers in the same school taught the Horizons Fast Track A-B program and three teachers, in the other school, taught the Guided Reading approach.

Program Treatment

Horizons Fast Track A-B. This program provides structured teaching of reading skills. It has a track design which presents four or five skills per lesson that are gradually upgraded in complexity. Horizons Fast Track A-B teaches phonics explicitly and systematically. It has procedures
for dealing with phonologically irregular words and includes many opportunities to read decodable text. Spelling exercises reinforce the relationship between sounds and spelling patterns, and activities are included to help students increase reading comprehension. Each lesson begins with letter sounds, followed by word-attack skills and oral reading of a story. After oral reading, story-based activities such as workbook activities are conducted, followed by letter writing, sentence writing, and spelling. Horizons Fast Track A-B has scripted lessons in a teacher-presentation book. The first 55 lessons are highly prompted. The following is an example from lesson 44: “Next word. Touch and say the sounds. Get ready. (Tap three times.) rrr ... aaa... mmm What word? (Signal) Ram” (Engelmann et al., 1997). Throughout, this program will be referred to simply as Horizons.

Guided Reading. According to Torgesen (2006), Guided Reading helps students integrate new skills and concepts and read text for meaning. A book is introduced to the class. The class reads the book and then discusses the book in a book club-type group. Comprehension strategies are taught in the process of reading and discussion. In this approach, children read from multiple copies of trade books and with their teacher from a big book. This approach divides instruction into blocks. The first block usually begins with a teacher-led discussion to establish background knowledge needed to understand the story that will be read. During the guided-reading block, students are divided into groups and the teacher explains how they will work. There is no specific set of books, nor is there a definite scope and sequence in this approach.

Teacher Experience and Training
All teachers taught their particular programs based on the training designated by the specific guidelines and directions of the different programs. The teachers who taught Horizons were using the standard curriculum for their school. Two of the three teachers were experienced (3 and 5 years) teachers of the Horizons program, and one was a first-year instructor of this program. Teachers were given one full day of training with follow-up consultation (verbal and written) twice during the year by the initial trainer. The training consisted of oral explanations of the curriculum, consultant modeling, and guided teacher practice. Teachers were trained to follow the scripts, do correction procedures, and maintain the pace of instruction.

The teachers who used the Guided Reading approach were instructed in this approach by the school principal. All teachers had used the Guided Reading approach for 5 years before the evaluation was conducted. The initial training occurred over a two-day period as part of the district’s in-service schedule. Training involved a description and overview of the program and a description of the use of leveled books in combination with big-book reading. Each year the program was reviewed during in-service training days, and a whole-faculty study group was created to improve the consistency and quality of reading instruction using this approach.

Treatment Fidelity
The principal in each school completed a Qualitative Monitoring of Instruction Survey on each teacher by observing each teacher twice during the year—once in the fall and once in the spring. The survey consisted of the following nine true-false questions that were completed during each observation: (a) materials are ready and the class is set up and prepared; (b) the lesson was covered faithfully and presented accurately; (c) the teacher covered all the key concepts in her lesson plan; (d) the teacher followed the suggested format in the curriculum guide accurately; (e) the teacher knows the program expectations and uses them to enhance instruction; (f) the lesson was covered at a reasonable pace; (g) the teacher checked to see if students were
engaged; (h) students appeared engaged and responded to questions; and (i) the teacher gave homework when applicable.

For this survey, scores between 7 and 9 indicate the teacher gave very organized and reliable instruction, scores between 4 and 6 indicate the teacher gave somewhat organized and reliable instruction, and scores below 4 indicate the teacher is at risk for not being reliable. On the Qualitative Monitoring of Instruction Survey, the Horizons teachers achieved an average score of 8.0 out of 9.0 and the Guided Reading approach teachers achieved a perfect score of 9.0.

Measures
Hintze (2002) suggested that progress monitoring is a useful form of evaluating response to intervention and program evaluation. Therefore, this technique was used to assess the reading programs in this study. The measures included Phonemic Segmentation Fluency (PSF), Nonsense Word Fluency (NWF), and Oral Reading Fluency (ORF). PSF is an individually administered test of phonemic awareness. NWF is an individually administered test of the alphabetic principle that includes letter-sound correspondence and blending sounds into words. ORF is an individually administered measure designed to assess the fluency and accuracy of reading connected text. All measures were assessed using AIMSweb (Education, 2006) and were scored on a Palm computer at the time of administration.

Phonemic Segmentation Fluency (PSF). PSF involved the teacher saying a list of up to 24 words, one at a time, and asking the student to say all the sounds in each word. For example, if the teacher said “frog” the student said /f/ /r/ /o/ /g/. A running time was kept for 1 minute. The measure was the total number of phoneme segments verbalized in one minute. Kaminski and Good III (1996) found that the concurrent criterion validity is .54 with the Woodcock Johnson Psycho-Educational Battery (Woodcock & Johnson, 1989) readiness cluster score.

Nonsense Word Fluency (NWF). In NWF, students were presented with a sheet of 80 one-syllable, randomly arranged, VC or CVC nonsense words (e.g., lut). Students were trained on practice items and then were told, “When I say ‘begin,’ read the words as best you can.” The number of words read correctly in 1 minute was reported as NWF. NWF has a one-month alternate form reliability of .83. The concurrent criterion validity of NWF with the Woodcock-Johnson Psycho-Educational Battery (Woodcock & Johnson, 1989) readiness cluster is .59.

Oral Reading Fluency (ORF). ORF is a 1-minute oral reading test that provides the number of words correct per minute (WCPM). For this study, the alternate forms of the AIMSweb (Education, 2006) first-grade reading passages were used. These probes were used because they were grade leveled and were neutral with respect to the narrative text taught in the different programs. Students were given three passages to read, and the score from the median passage was put into the computer to track the level of performance of each student. The alternate form reliability of this measure ranges from .89 to .94. Criterion-related validity ranges from .52 to .91 with the Stanford Diagnostic Reading Test (Karlsen, Madden, & Gardner, 1975), the Peabody Individual Achievement Test (PIAT) (Dun & Markwardt, 1989), the Stanford Achievement Test (Psychological Corporation, 1988), and the Woodcock Reading Mastery Test-Revised (Woodcock, 1998).

Data Collection
All measures were assessed using AIMSweb (Education, 2006). Teams of trained teachers and paraprofessionals conducted all the assessments. The first data collection occurred during the first week in October (Fall) and...
included Letter Naming Fluency (LNF), Phoneme Segmentation Fluency (PSF) and Nonsense Word Fluency (NWF). Good III, Simmons, and Kame’enui (2001) generally suggest that LNF only be administered in first grade during the Fall assessment time due to ceiling effects. The second data collection time occurred in mid-January (Winter) and included, PSF, NWF, and Oral Reading Fluency (ORF). The final assessment was conducted in mid-May (Spring) and included only NWF and ORF.

Results

Pretreatment Analysis
To establish if there were any pretreatment differences between students in the Horizons and the Guided Reading conditions, one-way analyses of variance (ANOVAs) were conducted. No significant differences were found between the Horizons and Guided Reading conditions in reading performance at the Fall administration of NWF, F (1, 105) = .74, p > .05 and at the Winter administration of ORF, F (1, 105) = .56, p > .05. There was, however, a significant difference favoring the Horizons condition on PSF, F (1, 105) = 6.13, p < .01 at the Fall administration. Means and standard deviations for all testing administrations (Fall, Winter, and Spring) are provided in Table 1.

Post-treatment Analysis
Effect of reading programs on Phoneme Segmentation Fluency. To explore the response to treatment, a repeated measures ANOVA was conducted with Time (Fall, Winter, Spring) and Condition (Horizons vs. Guided Reading) for PSF. A statistically significant main effect of Time was shown for both conditions—Horizons and Guided Reading—for NWF, F (1,105) = 288.66, p < .000. However, no significant interaction was found between conditions, F (1,105) = .91, p = .38. Additionally, a trivial effect size was shown favoring the Guided Reading condition (.06) (See Table 1). Figure 2 shows means and growth for NWF.

Effect of reading programs on Oral Reading Fluency. A repeated measures ANOVA was run comparing Time (Winter, Spring) and Condition (Horizons vs. Guided Reading) for ORF. The effect for Time was significant, F (1, 105) = 462.97, p < .000 for both conditions. Furthermore, a significant Time X Condition interaction was present, F (1, 105) = 22.85, p < .000. For ORF, a large effect size of .92 was found favoring students in the Horizons condition over students in the Guided Reading condition (see Table 1 for means and SD). Means and growth for ORF are shown in Figure 3.

Discussion
The purpose of this exploratory study was to examine the effect of two theoretically different first-grade reading programs on phonological and reading fluency skills. This study asked
Table 1
Means, Standard Deviations, Growth, and Effect Size for AIMSweb Reading Subskills

<table>
<thead>
<tr>
<th>Variable</th>
<th>Horizons n = 47</th>
<th>Guided Reading n = 60</th>
<th>Interaction F</th>
<th>ES</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Letter Naming Fluency</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall</td>
<td>M (SD)</td>
<td>36.30 (16.85)</td>
<td>41.30 (17.20)</td>
<td></td>
</tr>
<tr>
<td><strong>Phoneme Segmentation Fluency</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall</td>
<td>M (SD)</td>
<td>39.49 (14.02)</td>
<td>31.27 (19.09)</td>
<td></td>
</tr>
<tr>
<td>Winter</td>
<td>M (SD)</td>
<td>47.15 (7.42)</td>
<td>57.22 (10.60)</td>
<td>32.64***</td>
</tr>
<tr>
<td>Growth</td>
<td>M (SD)</td>
<td>7.66 (13.79)</td>
<td>25.95 (18.58)</td>
<td>-1.10</td>
</tr>
<tr>
<td><strong>Nonsense Word Fluency</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fall</td>
<td>M (SD)</td>
<td>27.66 (20.55)</td>
<td>24.35 (19.13)</td>
<td></td>
</tr>
<tr>
<td>Winter</td>
<td>M (SD)</td>
<td>50.89 (14.55)</td>
<td>52.88 (24.53)</td>
<td></td>
</tr>
<tr>
<td>Spring</td>
<td>M (SD)</td>
<td>78.78 (26.09)</td>
<td>77.18 (34.96)</td>
<td>.91</td>
</tr>
<tr>
<td>Growth</td>
<td>M (SD)</td>
<td>51.13 (25.14)</td>
<td>52.83 (28.28)</td>
<td>-.06</td>
</tr>
<tr>
<td><strong>Oral Reading Fluency</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Winter</td>
<td>M (SD)</td>
<td>28.19 (22.49)</td>
<td>31.92 (27.89)</td>
<td></td>
</tr>
<tr>
<td>Spring</td>
<td>M (SD)</td>
<td>65.75 (27.86)</td>
<td>55.82 (33.75)</td>
<td>22.85***</td>
</tr>
<tr>
<td>Growth</td>
<td>M (SD)</td>
<td>37.55 (15.54)</td>
<td>23.90 (13.94)</td>
<td>.92</td>
</tr>
</tbody>
</table>

***p<.000
Figure 1
Means for Assessments and Growth for Phoneme Segmentation Fluency

Phoneme Segmentation Fluency

0 10 20 30 40 50 60 70
Sounds Correct Per Minute

Assessment Times

Fall Winter Growth

Horizons
Guided Reading

Figure 2
Means for Assessments and Growth for Nonsense Word Fluency

Nonsense Word Fluency

0 10 20 30 40 50 60 70 80 90
Words Correct Per Minute

Assessment Times

Fall Spring Winter Growth

Horizons
Guided Reading
one question: What is the effect of an explicit and structured reading program on the acquisition of phonological and reading fluency skills? Results will be discussed in terms of each reading fluency skill.

Phoneme Segmentation Fluency
Results demonstrate that students in both conditions increased significantly across time on PSF; however, students in the Guided Reading condition made significantly greater gains in this skill than students in the Horizons condition. To further substantiate this finding, a large effect size (1.10) was found favoring the Guided Reading condition. It is interesting to note at pretest that although students in the Horizons condition began with a significantly higher mean score than the students in the Guided Reading condition, these students made up the difference and exceeded gains made by students in the Horizons condition. Research has suggested that the ability to process the phonological or sound structure of oral language is a vital part of explaining the rate or fluency of reading acquisition (Torgesen, Wagner, Rashotte, Burgess, & Hecht, 1997). The results shown in this study for PSF were assessed at the beginning and middle of the year. It would be interesting to see if this effect held longitudinally through to the end of the year given that at the end of first grade, students in the Horizons condition performed significantly better than the students in the Guided Reading condition on ORF. If PSF had been measured again at the end of the year, the students in the Horizons condition may have shown larger gains in their PSF scores.

Nonsense Word Fluency
Impressively, students in both conditions significantly increased their NWF scores across the school year. Although the students in the
Horizons condition did not outperform the students in the Guided Reading condition, they did make substantial growth. A new meta-analysis reanalyzing the results of the National Reading Panel findings has demonstrated that phonics instruction is only marginally more effective than nonphonics instruction for instructing first-grade reading (Hammill & Swanson, 2006). The results from this study support the findings of Hammill and Swanson. However, many other studies do not support the results demonstrated here (Adams, 1990; Chall, 1967; Snow et al., 1998). Additional research is needed to substantiate these findings.

Oral Reading Fluency
Oral reading fluency significantly increased across time for both conditions; however, students in the Horizons condition made substantially greater gains than students in the Guided Reading condition. According to Kame‘enui and Simmons (2001), ORF is an indicator of efficient word recognition, a well-developed vocabulary, and meaningful comprehension of reading selection. Good III et al. (2001) proposed that fluency is a causal factor, which results in higher-order skills. Torgesen (2006) describes it as “a bridge to comprehension” (p. 336). When one can read a text fluently, one can recognize the words and simultaneously understand what one is reading. Hence, a reading program that maximizes reading fluency by the end of first grade is likely to result in improved reading outcomes, thus meeting a requirement of NCLB legislation. These findings suggest ORF develops best when reading instruction involves systematic explicit phonics instruction, built-in correction procedures, and judicious practice of skills. Students receiving instruction with Horizons showed the better rate of improvement in ORF.

The findings of this study are consistent with prior studies that demonstrate enhanced reading outcomes with more explicit and systematic phonics in first grade. Shinn, Good III, Knutson, Tilly, and Collins (1992) used a confirmatory factor analysis to determine the relationship between ORF and reading comprehension. Cates, Thomason, and Hvey (2007) found that interventions that increased reading fluency also increased reading comprehension. Thus, research suggests that ORF is an excellent measure of reading comprehension at the elementary school level.

Conclusions
Results of this study show a differing response to interventions across conditions. Students in both conditions increased substantially on phonological and oral reading fluency skills, demonstrating that both programs were effective in increasing these skills for first-grade students. However, an important aspect of beginning reading is not just attaining the skills but maintaining the skills. Of interest is whether these skills will be maintained for all students in both conditions. Many studies show that providing students with early and systematic phonics instruction leads to better reading achievement and maintenance of those reading skills than when less systematic phonics instruction is provided (Adams 1990; Chall, 1967; National Reading Panel, 2000; Snow et al., 1998). Results from this study, however, are inconclusive in determining which program produces the better overall reading outcomes for first-grade students. On the other hand, the Horizons condition produced superior outcomes for ORF, and research suggests that with this skill students will be better readers at the end of third, fifth, and 11th grades (Cunningham & Stanovich, 1998; Foorman, 2003; Francis et al., 1996; Rayner et al., 2002).

Limitations of This Study
As with any study there are limitations. In this study the schools were not randomly assigned, there was a small sample of students in each condition, and fidelity of implementation was not quantitatively examined, but rather imple-
mented based on each school’s current policies. Future studies could eliminate these limitations by randomly assigning schools and students to conditions, using a larger sample size, conducting teacher fidelity of implementation, and reporting a level of testers’ interrater and interscorer reliability.

Implications

The findings from this exploratory comparison of two theoretically different reading programs resulted in mixed findings across programs for beginning reading skills. These findings contradict Hammill and Swanson (2006) who showed that for whole-class instruction and for most typically achieving students, phonics instruction and nonphonics instruction are not substantially different in their effectiveness when the goal is oral text reading. Results from this comparison suggest that an explicit and structured reading program helps in producing higher rates of oral reading fluency than a more implicit reading program. However, additional research is needed to further substantiate these findings.

References


Karlsen, B., Madden, R., & Gardner, E. (1975) Stanford


