The relative effectiveness of two approaches to early literacy intervention in grades K–2

Barbara Foorman
Sarah Herrera
Jennifer Dombek
Florida Center for Reading Research at Florida State University

This randomized controlled trial in 55 low-performing schools across Florida implemented two literacy interventions—one using a standalone intervention and one using materials embedded in the core reading program. The interventions were delivered daily for 45 minutes for 27 weeks in 2013/14 and 2014/15 in small groups of K–2 students at risk of literacy failure. The standalone intervention significantly improved grade 2 spelling outcomes relative to the embedded intervention, but impacts on other student outcomes were similar for the two interventions. On average, students in schools in both intervention groups showed improvements relative to each other in reading and language outcomes. The two interventions also had a similar impact on reading and language outcomes among English learner students and non–English learner students, except for some reading outcomes in kindergarten.

Why this study?

Understanding written language is crucial to academic success in all content areas. Ensuring a strong foundation in the components of written language—the literacy skills of reading, writing, and oral language (Mehta, Foorman, Branum-Martin, & Taylor, 2005)—is essential if students are to read with understanding and, thus, is a primary goal of early literacy instruction and of the Regional Educational Laboratory Southeast Improving Literacy Research Alliance. When students fall behind in developing literacy skills, early intervention can reduce the numbers of students failing to reach grade-level expectations (Foorman & Al Otaiba, 2009; Foorman, Breier, & Fletcher, 2003; Foorman & Torgesen, 2001). Findings from studies on early literacy interventions are likely of interest to district and state reading coordinators in the alliance and nationwide.

There is a strong research base on the skills targeted by effective early literacy intervention (Foorman et al., 2016). Effective early literacy intervention includes explicit instruction in phonological awareness, links from letters to sounds, decoding, and word study as well as practice in reading text for accuracy, fluency, and comprehension (Foorman et al., 2016; Foorman & Connor, 2011; National Institute of Child Health and Human Development, 2000; Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001; Snow, Burns, & Griffin, 1998). These components are often delivered in multiple tiers of instruction that include the classroom at tier 1, supplementary, small-group intervention at tier 2, and intensive intervention at tier 3 for students who do not progress after a reasonable amount of time with tier 2 intervention (Gersten et al., 2009).

To improve comprehension of content area text, students must also learn the vocabulary and discourse elements—the academic language—of the texts. Research is increasingly demonstrating the efficacy of directly teaching academic language to students in order to improve their comprehension (Baker et al., 2014; Foorman et al., 2016). Thus, literacy intervention that aims to improve comprehension must include instruction in both reading and language skills.

An important consideration for schools is which early literacy intervention to select (Dombek, Foorman, Garcia, & Smith, 2016). One approach is to use the intervention materials embedded in the core reading program employed for classroom instruction. That approach is appealing because the embedded materials are aligned with core classroom instruction and generally do not require buying additional materials. But even though embedded materials may claim to be research based, they are rarely evaluated empirically.

Another approach is to use standalone instructional materials and strategies that are outside the core curriculum. It is reasonable to expect that a standalone intervention with a strong evidence base (as indicated by the What Works Clearinghouse, for example) will lead to better reading and language outcomes for small group tier 2 intervention than will an embedded intervention that has not been empirically evaluated.

The current study compared these two approaches to early literacy intervention—a standalone intervention and an intervention embedded in the core curriculum. The standalone intervention combined a reading component and two oral language components: Sound Partners (reading component; Vadasy & Sanders, 2012; Vadasy, Sanders, & Abbott, 2008; Vadasy, Sanders, & Peyton, 2006; Vadasy et al., 2004), a What Works Clearinghouse–reviewed intervention that had strong levels of evidence in alphabets, fluency, and comprehension (taught daily); Bridge of Vocabulary (oral language component; Montgomery, 2007), which focuses on building oral vocabulary and concepts using manipulatives and discussion (taught three times a week); and Language in Motion (oral language component; Phillips, 2014), an inferential language program that uses science-based manipulatives to build oral language components of syntax, inferential language, and listening comprehension (taught twice a week). The intervention embedded in the core curriculum combined a reading component and an oral language component that were both included within Houghton Mifflin Harcourt Journeys (the core curriculum followed in all the study schools): the tier 2 Strategic
Intervention (reading component) and Curious about Words (a supplementary vocabulary piece that made up the oral language component); both were taught daily.

The study examined two cohorts of schools (referred to as cohort 1 and cohort 2), one in 2013/14 and one in 2014/15. Each school had three to four interventionists who taught the lessons associated with each intervention, serving four to six small groups daily. Interventionists had experience working with young children in education settings and received two days of training in late September. Some interventionists were school-based paraprofessionals assigned by the schools, and others were hired by Regional Educational Laboratory (REL) Southeast. For cohort 1 REL Southeast provided 66 interventionists, schools provided 17 paraprofessionals, and together they served 370 small groups. For cohort 2 REL Southeast provided 64 interventionists (42 percent of whom were interventionists for cohort 1 schools), schools provided 25 paraprofessionals, and together they served 424 small groups.

**What the study examined**

To evaluate the effectiveness of an intervention, it should be compared with logical alternatives, preferably in a random assignment design using appropriate outcome measures. This study used a randomized controlled trial conducted across the 2013/14 and 2014/15 school years in 55 low-performing Florida schools, as identified by the state’s school grading system.

The study addressed three research questions separately for students in kindergarten, grade 1, and grade 2 who were at risk of literacy failure:

- What are the improvements in percentile rank on reading and language measures in the standalone and embedded early literacy interventions?
- What are the impacts of a standalone early literacy intervention relative to an embedded early literacy intervention on reading and language outcomes? Does the impact differ by cohort or baseline performance?
- Are there differences in reading and language outcomes between the standalone and embedded early literacy interventions depending on English learner status? Are there differences in reading and language outcomes within interventions between English learner students and non–English learner students?

Box 1 summarizes the data and methods used in the study.

**Box 1. Data and methods**

**Data**

The study used data provided by schools in a large urban district in south Florida, a medium-size urban district in central Florida, and three small rural districts in north Florida. There were two nonoverlapping cohorts of schools: cohort 1 included 27 schools and 1,598 students that participated in the 2013/14 school year, and cohort 2 included 28 schools and 1,870 students that participated in the 2014/15 school year. All participating schools were low performing, as identified by the state’s school grading system. Participating students were in grades K–2, were at risk of literacy failure, and had parent consent to participate. The average percentage of students who qualified for the federal school lunch program (a proxy for low-income status) was 72–78 percent across the two interventions. Approximately 30–42 percent of participating students in cohorts 1 and 2 combined across interventions and grades were English learner students.

Several reading and language measures were included at baseline and outcome. Reading baseline measures were the Letter Sounds (kindergarten only), Phonological Awareness (kindergarten only), Word Reading (grades 1 (continued)
Box 1. Data and methods (continued)

and 2), and Spelling (grade 2 only) subtests from the Florida Center for Reading Research Reading Assessment (FRA; Foorman, Petscher, & Schatschneider, 2015; see the appendix for a description). Language baseline measures were the Vocabulary Pairs, Following Directions, and Sentence Comprehension (kindergarten and grade 1) subtests from the FRA. Reading outcomes were the Phonological Awareness (kindergarten only), Word Reading, and Spelling (grade 2 only) subtests from the FRA and the Word Reading subtest from the Stanford Early Scholastic Achievement Test (SESAT) in kindergarten. Language outcomes were the Vocabulary Pairs, Following Directions, and Sentence Comprehension subtests from the FRA; the Sentence Reading subtest from the SESAT in kindergarten; and the Reading Comprehension subtest from the Stanford Achievement Test, 10th edition, in grades 1 and 2.

Methods
Participating schools were randomly assigned to use a standalone or embedded approach to early literacy intervention. Students received daily intervention for 45 minutes from mid-October through May, about 27 weeks, in small groups of four (kindergarten and grade 1) or five (grade 2). About 30 minutes were devoted to the reading component, and about 15 minutes to the oral language component.

The study team observed interventionists once in the fall and once in the spring to rate fidelity of implementation. Fidelity is defined as the percentage of the lesson in which instruction followed the lesson sequence and script for the skills taught in each component. For both interventions, 72–91 percent of small groups demonstrated at least 80 percent fidelity on the reading and oral language components. The median overall fidelity across interventions was 96 percent in kindergarten, 94 percent in grade 1, and 96 percent in grade 2.

Across grades K–2, interventionists covered an average of 55–80 percent of the reading component and 77–79 percent of the oral language component in the standalone intervention and 86–88 percent of the reading and oral language components in the embedded intervention. Out of 134 days of instruction, students in standalone intervention schools attended 92–95 days of intervention on average, and students in embedded intervention schools attended 96–98 days.

Prior to analyses, baseline equivalence was assessed by comparing differences between the interventions on all reading and language baseline measures by grade at the school and student levels. Most of the differences in baseline scores by grade at the school and student levels between students in standalone intervention schools and students in embedded intervention schools were not statistically significant. One exception was the FRA Word Reading subtest for grade 1, where scores were significantly higher for students in embedded intervention schools than for students in standalone intervention schools.

Multilevel analyses of student outcomes were conducted by grade, with students nested in small groups, nested within schools. All analyses included student, small group, and school-level baseline measures as covariates. Student-level covariates in kindergarten were FRA Phonological Awareness, Letter Sounds, Vocabulary Pairs, Following Directions, and Sentence Comprehension scores, and student-level covariates in grades 1 and 2 were FRA Vocabulary Pairs, Following Directions, Sentence Comprehension (grade 1 only), Word Reading, and Spelling (grade 2 only) scores. Baseline scores were aggregated by small group and then by school and were used as covariates at their respective levels. Cohort and region were also included as school-level covariates. Cohort was included as an analytic variable because different schools participated each year.

The differences in outcomes between the standalone and embedded interventions are reported as effect sizes (Hedges’ g), which describes the magnitude of the difference as the proportion of a standard deviation. For example, an effect size of 0.25 means that the average student in one intervention group would be expected to have scored 0.25 standard deviation higher had he or she been in the other intervention group. The What Works Clearinghouse criterion of an effect size of 0.25 or greater (U.S. Department of Education, 2014) is used to identify substantively important differences.

Note
1. One of the standalone intervention schools in cohort 2 was excluded from the grade 2 analyses because scheduling conflicts resulted in the withdrawal of the 21 participating grade 2 students at that school. The cohort total includes these 21 students.
What the study found

This section discusses the findings of the study, starting with baseline and outcome percentile ranks on the reading and language measures by grade and intervention. It then reports differences in reading and language outcomes between the standalone intervention and embedded interventions for all students and by cohort and baseline performance. Finally, it reports differences in reading and language outcomes by English learner status.

Students in both intervention groups made gains on reading and language outcomes

In grades K–2, students in schools in both intervention groups started, on average, at or below the 10th percentile on Florida Center for Reading Research Reading Assessment (FRA) reading measures (Phonological Awareness in kindergarten, Word Reading in grades 1 and 2, and Spelling in grade 2) and ended the year 13–25 percentile points higher across FRA reading subtests and grades, based on Florida mid-year norms.

In kindergarten, students in both intervention groups started, on average, at or below the 10th percentile on two of the FRA language measures (Following Directions and Sentence Comprehension) and ended the year above the 25th percentile. Students started at about the 25th percentile on the FRA Vocabulary Pairs measure and ended the year at about the 34th percentile, a 9 percentile point gain.

In grades 1 and 2, students in schools in both intervention groups started, on average, between the 10th and 15th percentiles on two of the FRA language measures (Following Directions and Vocabulary Pairs) and ended the year between the 18th and 30th percentiles; the average gain was 6–15 percentile points.

The largest average difference between baseline and outcome percentile ranks for any FRA measure was for Sentence Comprehension in grade 1. Students in schools in both intervention groups began just below the 30th percentile and ended the year above the 60th percentile. This reflects an average difference of 35–39 percentile points between baseline and outcome percentile ranks across interventions. However, the norms for FRA Sentence Comprehension are based on kindergarten students, which means that the percentile ranks for all grades reflect ability on a kindergarten scale.

The standalone intervention group had better average spelling outcomes than the embedded intervention group, but the two groups had similar reading and language outcomes

For grade 2 students in cohorts 1 and 2 the standalone intervention led to significantly better FRA Spelling outcomes compared with the embedded intervention. The effect size was 0.18 (figure 1). Overall, there were no other significant or substantively important relative impacts on reading or language outcomes for grades K–2.

Among students with low baseline scores, those in the standalone intervention groups experienced the largest gains. Relative to the embedded intervention the standalone intervention significantly improved FRA Spelling outcomes for grade 2 students in cohorts 1 and 2 combined who had a baseline FRA Spelling score one standard deviation below the mean (effect size of 0.27) and FRA Sentence Comprehension outcomes for grade 2 students in cohort 1 who had a baseline FRA Vocabulary Pairs score one standard deviation below the mean (effect size of 0.38; figure 1).

The SESAT Word Reading outcome in kindergarten and the FRA Word Reading outcome in grade 1 showed substantively important (effect size greater than 0.25) but not statistically significant differences. In kindergarten the SESAT Word Reading outcome was higher in standalone intervention schools than in embedded intervention schools among students in cohort 1 with a baseline FRA Sentence Comprehension score one standard deviation above the mean (effect size of 0.37) and students in cohort 2 with a baseline FRA Sentence...
Comprehension score one standard deviation below the mean (effect size of 0.28; figure 1). In grade 1 the FRA Word Reading outcome among students in cohort 1 was higher in embedded intervention schools than in standalone intervention schools (effect size of 0.29; see figure 1). There were no other differences in any other reading or language outcome in kindergarten, grade 1, or grade 2 for subgroups of students between interventions.

For reading outcomes in kindergarten, English learner students appeared to benefit more from embedded intervention while non-English learner students appeared to benefit more from the standalone intervention.

Among English learner and non-English learner students there were no differences in language outcomes in kindergarten or in reading and language outcomes in grades 1 and 2 between and within interventions. However, two differences in reading outcomes between interventions in kindergarten were substantively important. The FRA Phonological Awareness outcome among English learner students was higher in embedded intervention schools than in standalone intervention schools (effect size of 0.32; see figure 1). Among non–English learner students the standalone intervention had a substantively important effect on the SESAT Word Reading outcome relative to the embedded intervention (effect size of 0.31; see figure 1).

Also in kindergarten, the SESAT Word Reading outcome was higher among English learner students in embedded intervention schools than among non–English learner students in embedded intervention schools (effect size of 0.27).

Figure 1. Significant and substantively important differences in outcomes between standalone and embedded interventions for the full sample, by cohort and baseline performance, and for English learner and non–English learner students for grades K–2

Adjusted outcome based on estimated model

$g$ is Hedges’ $g$ effect size estimate. FRA is Florida Center for Reading Research Reading Assessment. SESAT is Stanford Early Scholastic Achievement Test. SC is Sentence Comprehension. VP is Vocabulary Pairs.

* $p$-values are significant after applying the Benjamini–Hochberg correction procedure (1995): .009 for FRA Spelling full sample (cutoff of $p \leq .025$), .001 for FRA Spelling subset of full sample (cutoff of $p \leq .0025$), and .001 for FRA Sentence Comprehension (cutoff of $p \leq .00125$).

Note: The standard error around the adjusted outcome score is reflected at the top of each bar. High baseline scores are one standard deviation above the mean. Low baseline scores are one standard deviation below the mean.

Source: Authors’ analysis based on data from participating districts in Florida.
Implications of the study findings

On average, students in standalone and embedded intervention schools made gains in percentile ranks and had similar relative impacts on all reading and language outcomes. There was, however, a statistically significant improvement in spelling outcome in grade 2 for students in standalone schools relative to students in embedded schools. It is important to note that reading instruction in the standalone intervention integrated the decoding and encoding (spelling) of words, whereas reading instruction in the embedded intervention included only the decoding of words. Many effective early reading interventions integrate instruction in encoding and decoding (Foorman et al., 2016).

Additionally, for reading outcomes in kindergarten, English learner students appeared to benefit more from the embedded intervention, whereas non–English learner students appeared to benefit more from the standalone intervention. Both interventions included instruction in phonological awareness, but the addition of comprehension activities in the embedded intervention may have helped scaffold English learner students’ ability to segment sounds in speech. This finding is consistent with studies showing an advantage for bilingual students in phonological awareness tasks (for example, Bialystok, Majumder, & Martin, 2003).

Within embedded intervention schools, SESAT Word Reading outcomes in kindergarten were higher among English learner students than among non–English learner students. These results underscore the value of emphasizing comprehension when building on English learner students’ sensitivity to sounds in speech in order to connect to the sound-spelling patterns fundamental to reading.

The study also has implications for future directions in research on early literacy interventions. Experiments could modify the standalone intervention in ways that might make it easier to implement. First, it was challenging for interventionists to decide how to remediate students on different skills and what to do with students who did not need remediation. A version of the reading component of the standalone intervention that eliminates remediation could be contrasted with the current version to see whether student reading outcomes differed. Second, interventionists had to remember which day of the week to teach vocabulary and which day to teach inferential language. This was challenging because of the disruptions in school schedules that required interventionists to remember which language piece had to be rescheduled. An integrated version of the language component in the standalone intervention in which vocabulary and inferential language are taught each day could be contrasted with the current version to see whether student language outcomes differed.

An area of investigation for the embedded intervention is to verify its alignment to core classroom (tier 1) instruction and then to manipulate enhancements to both core classroom (tier 1) instruction and small group (tier 2) instruction. To enhance and thereby achieve high implementation fidelity in the embedded intervention in the current study, REL Southeast staff developed an implementation manual that revealed the scope and sequence and established procedures for well-trained interventionists to deliver daily small-group intervention in a consistent fashion to a diverse population of students. Once this enhanced implementation of the tier 2 embedded intervention is developed, the next step in studying modifications is to compare the current version of enhanced tier 2 and typical tier 1 with a version where both are enhanced. Smith et al. (2016) found higher reading outcomes for at-risk students in the primary grades when they received enhanced tier 1 and 2 instruction compared with when they received the typical, nonenhanced tier 1 and 2 instruction. Tier 1 might be enhanced by making evidence-based elements more explicit and providing more scaffolding so that instruction is accessible to a broad range of students (for example, Smith et al., 2016). Through aligning curricula and enhancing the quality of instruction, multtiered intervention for at-risk readers in the early grades has the potential to ensure literacy success for many students.
Appendix. Description of the K–2 Florida Center for Reading Research Reading Assessment screening tasks

The Florida Center for Reading Research Reading Assessment (FRA) is a computer-adaptive screening assessment of reading and language for students in grades K–2 (table A1). The FRA was developed under federal grants to Florida State University (Foorman et al., 2015) and normed on Florida students. In all of the FRA subtests, students receive five items at grade level and then the system adapts up or down based on performance to reach a precise estimate of a student’s ability. The marginal reliability (Sireci, Thissen, & Wainer, 1991) for the FRA subtests based on the normative sample ranges from .85 to .96 across grades K–2. Students are given a developmental ability score on each subtest; subtest scores have a mean of 500 and a standard deviation of 100.

Table A1. Florida Center for Reading Research Reading Assessment subtests, by grade and assessment period

<table>
<thead>
<tr>
<th>Subtest</th>
<th>Kindergarten Baseline</th>
<th>Kindergarten Outcome</th>
<th>Grades 1 and 2 Baseline</th>
<th>Grades 1 and 2 Outcome</th>
<th>Subtest description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phonological Awareness</td>
<td>✔</td>
<td>✔</td>
<td></td>
<td></td>
<td>Students listen to a word that has been broken into parts and then blend them back together to reproduce the word.</td>
</tr>
<tr>
<td>Letter Sounds</td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
<td>A letter is presented on the monitor in upper and lower case and students provide the sound it makes.</td>
</tr>
<tr>
<td>Vocabulary Pairs</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>Three words appear on the monitor and are pronounced by the computer. The student selects the two words that go together best (for example, dark, night, swim).</td>
</tr>
<tr>
<td>Following Directions</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>Students listen and then click and drag objects in response to the computer’s directions (for example, put the square in front of the chair and then put the circle behind the chair).</td>
</tr>
<tr>
<td>Sentence Comprehensiona</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>Students listen to a sentence given by the computer (for example, click on the picture of the bird flying towards the nest) and then select the one picture out of the four presented on the monitor that depicts the sentence.</td>
</tr>
<tr>
<td>Word Reading</td>
<td></td>
<td>✔</td>
<td>✔</td>
<td></td>
<td>Words of varying difficulty are presented on the monitor one at a time and students read them aloud.</td>
</tr>
<tr>
<td>Spellingb</td>
<td></td>
<td></td>
<td>✔</td>
<td></td>
<td>The computer provides each word and uses it in a sentence. Students respond by using the computer keyboard to spell the word.</td>
</tr>
</tbody>
</table>

Note: Subtests were administered to individual students. Baseline testing occurred in September or October; outcome testing occurred in April or May.

a. Administered at baseline only to kindergarten and grade 1 students.

b. Administered only to grade 2 students.

Source: Authors’ compilation based on tasks included in the computer-adaptive K–2 Florida Center for Reading Research Reading Assessment.
References


The National Center for Education Evaluation and Regional Assistance (NCEE) conducts unbiased large-scale evaluations of education programs and practices supported by federal funds; provides research-based technical assistance to educators and policymakers; and supports the synthesis and the widespread dissemination of the results of research and evaluation throughout the United States.

April 2017

This report was prepared for the Institute of Education Sciences (IES) under Contract ED-IES-12-C-0011 by Regional Educational Laboratory Southeast administered by the Florida Center for Reading Research, Florida State University. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

This REL report is in the public domain. While permission to reprint this publication is not necessary, it should be cited as:


This report is available on the Regional Educational Laboratory website at http://ies.ed.gov/ncee/edlabs.

The Regional Educational Laboratory Program produces 7 types of reports

<table>
<thead>
<tr>
<th>Type of Report</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Making Connections</strong></td>
<td>Studies of correlational relationships</td>
</tr>
<tr>
<td><strong>Making an Impact</strong></td>
<td>Studies of cause and effect</td>
</tr>
<tr>
<td><strong>What’s Happening</strong></td>
<td>Descriptions of policies, programs, implementation status, or data trends</td>
</tr>
<tr>
<td><strong>What’s Known</strong></td>
<td>Summaries of previous research</td>
</tr>
<tr>
<td><strong>Stated Briefly</strong></td>
<td>Summaries of research findings for specific audiences</td>
</tr>
<tr>
<td><strong>Applied Research Methods</strong></td>
<td>Research methods for educational settings</td>
</tr>
<tr>
<td><strong>Tools</strong></td>
<td>Help for planning, gathering, analyzing, or reporting data or research</td>
</tr>
</tbody>
</table>