Replicating the impact of a supplemental beginning reading intervention: The role of instructional context

Coyne, M. D., Little, M. E., Rawlinson, D. M., Simmons, D. C., Kwok, O., Kim, M, Simmons L., E., Hagan-Burke, S., & Civetelli, C.

University of Connecticut

Context / Purpose

The purpose of this study was to conduct a rigorous varied replication evaluation of a beginning reading intervention in schools characterized by a different instructional context than in the initial study. The Early Reading Intervention (ERI) is a commercially available kindergarten beginning reading intervention designed to supplement classroom instruction for students at risk of experiencing reading difficulties (Early Reading Intervention; Pearson/Scott Foresman, 2004). In our initial study (Simmons, Coyne, Hagan-Burke, Kwok, Simmons, Johnson et al., 2011), students (N = 206) were assigned randomly at the classroom level to the experimental group or a comparison group (i.e., a school-designed intervention). Both interventions were taught for 30 min per day in small groups of three to five students for approximately 100 sessions. Multilevel hierarchical linear analyses revealed statistically significant effects favoring the ERI on foundational alphabetic, phonemic, and untimed decoding skills with effect sizes ranging from .40 to .51.

The initial study took place in six different school districts in Connecticut and Texas characterized by a distinctive instructional context. In these schools, core reading instruction varied widely both within and across classrooms and schools. Although some schools used basal programs, most schools used a combination of published materials and less structured guided-reading strategies. In addition, a majority of these schools did not routinely provide systematic supplemental kindergarten reading intervention; therefore implementing ERI represented a significant extension of their current reading practices. Overall, these districts included schools with a less coordinated and more individual approach to providing kindergarten reading instruction and intervention.

In the current study, we were interested in determining whether the impact of ERI would replicate in a school district in a different geographical region of the country that had a very different instructional context. Therefore, we evaluated ERI in a Florida school district that had a more coordinated, systematic, and consistent approach to providing beginning reading instruction and intervention. Kindergarten teachers in schools in this district had recently received professional development related to evidence-based reading instruction and delivering common core instruction using a published commercial program. This district also had experience providing small-group kindergarten reading intervention to students at risk for reading difficulties. We believed that evaluating ERI in schools in this district would extend the findings from our initial study and provide information about the efficacy of ERI in schools with a very different instructional context.

In summary, we were interested in whether the impact of ERI in our initial study would replicate in the Florida school district. We also wanted to know
whether the effects of ERI on the absolute level of student performance measured at posttest would be similar across studies. And whether the posttest performance of the comparison groups, who received different school-designed interventions, would be comparable across studies.

**Setting / Participants**

In the replication study, kindergarten students (n = 162) identified as at risk of reading difficulty from 48 classrooms were assigned randomly at the classroom level either to a commercial program (i.e., Early Reading Intervention; Pearson/Scott Foresman, 2004) that included explicit/systematic instruction (experimental group) or school-designed typical practice intervention (comparison group). Both interventions were taught by classroom teachers for 30 min per day in small groups for approximately 100 sessions.

In the 2nd month of kindergarten, researchers consulted with school personnel to identify 6-7 children who were considered in need of supplemental, small-group reading intervention. Nominated students with parental consent were screened using the (a) DIBELS Letter Naming Fluency (LNF; Good & Kaminski, 2002) and (b) Sound Matching subtest from the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999). Students were determined to be at risk and qualified to participate in the study if they met either of the following criteria: (a) a score at or below the 33rd percentile on the DIBELS LNF measure (i.e., fewer than six letters correctly named in 1 min) or (b) a score below the 37th percentile on the CTOPP Sound-Matching subtest.

**Intervention**

The ERI curriculum includes 126 daily lessons. A 30-min lesson consists of seven activities, each designed to last 3 to 5 min. The first 15 min of the lesson focus on phonological awareness and alphabetic understanding; the second 15 min integrate writing and spelling with previously taught phonemic and alphabetic skills. The program is organized into four major components. Part I: Learning Letters and Sounds consists of 42 lessons and introduces 11 letter names and sounds and the phonemic skills of first and last sound isolation. Part II: Segmenting, Blending, and Integrating includes 30 lessons and continues with the introduction of five new letter names and sounds while introducing phonemic blending and segmenting using letter tiles. Part III: Reading Words completes the introduction of the six remaining letter names and sounds with primary instruction focusing on word decoding in vowel-consonant and consonant-vowel-consonant words. This part consists of 24 lessons in which instruction integrates oral segmenting and blending with real-word decoding and the introduction of irregular word reading. Part IV: Reading Sentences and Storybooks consists of 30 lessons. This final section focuses on combining alphabetic skills with irregular word reading to read sentences and short storybooks.

School designed intervention (SDI) was provided to control students for a similar amount of time across a similar number of lessons. All kindergarten teachers in this school district were experienced in providing supplemental beginning reading intervention and had previously received extensive professional
development and resources in evidence-based reading instruction methods, both at the district and school site, through the use of literacy coaches. Typical kindergarten intervention supports in this school district include strategically integrated phoneme awareness and alphabetic understanding instruction.

**Data Collection**

The assessment battery included screening, pretest, and posttest assessments. Assessments measured a range of early literacy and beginning reading constructs. Screening and pretest assessments were completed in September prior to the start of the intervention. Posttesting occurred within 2 weeks after the completion of the intervention in May. Measures assessed Letter Knowledge (Letter Naming Fluency subtest of the DIBELS, The Supplementary Letter Checklist of the Woodcock Reading Mastery Tests–Revised/Normative Update); Phonological Awareness (The Sound-Matching subtest of the CTOPP, Phoneme Segmentation Fluency subtest of the DIBELS); Decoding (The Word Attack subtest of the WRMT-R/NU, The Nonsense Word Fluency (NWF) subtest of the DIBELS); Word Identification (The Word Identification subtest of the WRMT-R/NU); and Vocabulary (Peabody Picture Vocabulary Test).

**Research Design**

To evaluate the impact of the ERI intervention in the varied replication study and compare the effects of ERI across the initial and replication studies, we included data from both studies in one multiyear model. We looked at four different comparisons across four different groups of students: (a) students who received the ERI intervention in the initial study compared to students who received the comparison intervention in the initial study, (b) students who received the ERI intervention in the replication study compared to students who received the comparison intervention in the replication study, (c) students who received the ERI intervention in the initial study compared to students who received the ERI intervention in the replication study, and (d) students who received the comparison intervention in the initial study compared to students who received the comparison intervention in the replication study. It is important to note that only Comparisons a and b are experimental (i.e., groups randomized). Comparisons c and d were included to examine whether performance of students in the ERI and comparison groups were similar or different across the initial and replication studies. Furthermore, we use effect size terminology to interpret the magnitude of differences among all group comparisons.

Because of the nested structure of our data (i.e., students nested within intervention groups nested within schools), multilevel modeling (Hox, 2002) was chosen to analyze the data using Hierarchical Linear Model (HLM; V6.08; Raudenbush, Bryk, Cheong, & Congdon, 2004). The three-level analysis included students (i.e., Level 1) nested within interventionists (i.e., Level 2) and further nested within schools (i.e., Level 3).

**Findings / Results**

Procedural fidelity items for both the initial and the replication study were
evaluated on a 1-to-4 scale, with 4 indicating excellent, 3 indicating good, 2 indicating fair, and 1 indicating poor levels of fidelity. Mean procedural fidelity for the ERI interventionists was 3.07 for the initial study and 2.84 for the replication study, suggesting that interventionists implemented ERI with generally satisfactory levels of integrity. Items related to the quality of implementation of both ERI and SDI was also evaluated on a 1-to-4 scale. Mean quality scores for interventionists across studies and instructional conditions ranged from 2.52 to 3.01, indicating that the quality of instruction provided to students was fair to good.

Similar to findings reported in Simmons, Coyne, Hagan-Burke, Kwok, Simmons, Johnson et al. (2011), results from combined HLM models that included students from both the initial and varied replication studies indicated that during the initial study conducted in Texas and Connecticut, there was a statistically significant impact of ERI compared to the school-designed comparison condition on students’ timed and untimed phonemic awareness outcomes, letter-sound knowledge, and untimed word attack outcomes, with effects sizes ranging between .40 and .51. (See Table 1)

In the varied replication study conducted in Florida, however, no statistically significant differences between the ERI and SDI conditions were found on any posttest measure. In other words, the impact of ERI in the initial study did not replicate in the Florida study; that is, students in the treatment condition in the latter study did not benefit more than students in the comparison condition who received SDI. In fact, a trend favored the SDI condition on a number of the measures.

To unpack these findings, we next compared the absolute level of posttest performance of the students who received ERI during the initial and varied replication studies. Except for one statistically significant difference favoring the ERI students in the Florida replication on phonemic segmentation fluency (g = .50), there were no differences across studies. These findings suggest that the response of students to ERI was similar across the two studies, with students experiencing comparable achievement outcomes even though the ERI group outperformed the SDI comparison group in the initial study but not in the replication study.

Finally, we compared the outcomes of students who received the comparison SDI during the initial study with students who received SDI during the Florida varied replication study. In this case, analyses revealed a number of differences favoring the SDI students in the replication study. Specifically, we found statistically significant differences supporting Florida SDI students on both timed and untimed measures of phonemic awareness, letter-sound knowledge, nonsense word fluency, and word identification, with effect sizes ranging from .24 to 1.06. In addition, there was a trend favoring the Florida SDI students on all other measures as well. In other words, even though the comparison students in both studies had very similar pretest scores, the students who received SDI in the Florida replication study substantially outperformed the students who received SDI in the initial study.

In summary, results from combined analyses that included students from both the initial and varied replication studies suggest that differences in the impact of the ERI intervention across the initial and the replication study were largely explained by differences in students’ response to the SDI in the two studies. In the initial study, compared to SDI, ERI had a substantively important impact on the majority of
reading and reading-related measures. In the replication study, however, even though students who received ERI performed similarly to students who received ERI in the initial study, there was no differential impact of ERI because of the strong response of students to the SDI.

Conclusions

In this study, we found that the impact of supplemental beginning reading intervention observed in an initial randomized control study did not replicate in an experimental study conducted a year later in a different geographical location and a different instructional context. Multiyear analyses revealed that the differences in impact were largely explained by differences in students’ response to the SDI that comprised the comparison condition in each study, rather than differential response to the experimental intervention. These findings suggest that factors related to instructional context may play an important role in the impact of beginning reading interventions. For example, the Florida schools’ coordinated approach to beginning reading instruction provided a strong foundation for delivering supplemental intervention to students who required additional support. In these schools, the provision of school-designed kindergarten intervention was part of accepted practice and implemented as a matter of course and policy (Just Read Florida, 2005). Unlike the consistent framework for supporting beginning reading in the Florida replication schools, the Texas and Connecticut schools in the initial study took a more eclectic and individualized approach to reading instruction and intervention and implementation was less coordinated. Overall, findings reinforce the value of varied replication studies for providing evidence about the effects of practices in different settings and under different conditions.

References


**Update.** Bloomington, MN: Pearson Assessments.

### Table 1

**Effect Sizes (Hedge’s g) for Different Models**

<table>
<thead>
<tr>
<th>Measure</th>
<th>ERI Initial Study vs. SDI Initial Study*</th>
<th>ERI Replication Study vs. SDI Replication Study*</th>
<th>ERI Replication Study vs. ERI Initial Study*</th>
<th>SDI Replication Study vs. SDI Initial Study*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Alphabet Knowledge</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WRMT-R/NU Supplementary Letter Checklist-Name</td>
<td>0.19</td>
<td>-0.11</td>
<td>-0.17</td>
<td>0.24</td>
</tr>
<tr>
<td><strong>Letter Sound Knowledge</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WRMT-R/NU Supplementary Letter Checklist-Sounds</td>
<td><strong>0.44</strong></td>
<td>-0.04</td>
<td>-0.02</td>
<td><strong>0.45</strong></td>
</tr>
<tr>
<td><strong>Phonemic Awareness</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTOPP Sound Matching</td>
<td>0.43</td>
<td>0.15</td>
<td>-0.09</td>
<td>0.30</td>
</tr>
<tr>
<td>CTOPP Blending Words</td>
<td><strong>0.40</strong></td>
<td>-0.07</td>
<td>0.19</td>
<td><strong>0.80</strong></td>
</tr>
<tr>
<td>DIBELS Phonemic Segmentation Fluency</td>
<td><strong>0.46</strong></td>
<td>-0.08</td>
<td><strong>0.50</strong></td>
<td><strong>1.06</strong></td>
</tr>
<tr>
<td><strong>Word Attack</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIBELS Nonsense Word Fluency</td>
<td>0.36</td>
<td>-0.21</td>
<td>0.25</td>
<td><strong>0.77</strong></td>
</tr>
<tr>
<td>WRMT-R/NU Word Attack</td>
<td><strong>0.51</strong></td>
<td>0.00</td>
<td>-0.05</td>
<td>0.51</td>
</tr>
<tr>
<td><strong>Word Identification</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WRMT-R/NU Word ID</td>
<td>0.25</td>
<td>-0.18</td>
<td>0.22</td>
<td><strong>0.77</strong></td>
</tr>
</tbody>
</table>

*Note.* *reference group; Bolded:* significant effect.