An Evaluation of Close Reading With At-Risk Fourth-Grade Students in Science Content

Marcy Boudreaux-Johnson, Paul Mooney, and Renée E. Lastrapes

Abstract: The study’s primary purpose was to evaluate the effectiveness of a widely promoted close reading instructional routine for elementary grades students at risk for reading failure. Close reading is designed to help students read complex text independently and proficiently. Participants were six fourth-grade students receiving supplemental instruction in a rural public school. A single subject alternating treatments design was implemented to compare the close reading instructional routine to a validated reading comprehension strategy instruction intervention over a six-week intervention time frame. Results determined through visual inspection of a general outcome reading comprehension measure were mixed, seemingly favoring the validated intervention and not close reading. Limitations of the research and implications for use of close reading with students at risk are discussed.

In the most functional United States public school systems today, students at risk for academic failure receive preventive supports in what have been called responsiveness-to-intervention (RTI) systems (Fuchs, Mock, Morgan, & Young, 2003). The stated goal of RTI is to provide appropriate academic services to the most students through a tiered support system. Most students receive the core instructional program (i.e., Tier 1), with periodic assessment included, to identify objective success or failure in meeting core expectations and allowances for classroom-level differentiation of instruction if warranted. Generally core instructional programs are based on research, rather than validated through a true experiment (e.g., randomized control trial; Fuchs, Fuchs, & Compton, 2012). In reading curricula, for example, program components likely include phonemic awareness, phonics, vocabulary, fluency, and comprehension development.

In functional RTI systems, students identified as failing in the core program are provided additional academic intervention(s) in the area of identified need, and their progress is monitored more frequently than Tier 1 only students in order to determine if the supplemental (i.e., Tier 2) efforts are working (i.e., moving students closer to being successful in the core curriculum without supports). Supplemental intervention efforts for at-risk students in Tier 2 programs are designed to be different from the core curriculum by incorporating empirically validated interventions in the area of need (e.g., phonemic awareness or comprehension) and delivered to small-group formats for a specified time frame. Tier 2 progress monitoring is more frequent, with the assessment tools validated to serve the function of determining whether the combination of Tier 1 and 2 services is preventing additional academic and/or behavioral failure (Fuchs et al., 2012).

The implementation of RTI in schools has coincided with employment of the Common Core State Standards (CCSS) for English Language Arts-Literacy (ELA-L) in history/social studies, science, and technical subjects (Coleman & Pimentel, 2012; National Governors Association Center for Best Practices/Council of Chief State School Officers, 2010). From the outset of CCSS, proponents of the newly designed standards wanted to ensure that students graduating from high school were prepared to take college courses or enter the workforce. Changes to the ELA-L curriculum have included emphases on (a) a deeper understanding of the content of challenging or complex texts; (b) literary (i.e., reading, writing, and speaking) tasks that demand documentation of text-based evidence; and (c) a greater emphasis on the reading of informational texts in ELA classrooms.

Close Reading

One literary practice that has been featured prominently in the promotion and implementation of the CCSS and the accompanying national assessments (e.g., Partnership for Assessment of Readiness for College and Careers) is close reading. Also described as analytical reading, close reading has been defined as an in-depth analysis of a short piece of complex (i.e., at or above grade level) text conducted over multiple readings or lessons that stress attention to multiple textual aspects (Brown & Kappes, 2012). The instructional routine generally consists of teacher selection of short, complex texts that students read multiple times, with each subsequent reading having a different emphasis. Along with the multiple readings, there are discussions of teacher-developed questions that also serve to develop deeper understandings of the text than would likely be gained through single readings. Moreover, during the reading and discussion process, students annotate texts as another way to engage with what can be complex material (Brown & Kappes, 2012; Fisher & Frey, 2014b; Shanahan, 2014).

Close reading has been promoted as an essential element of the implementation of the CCSS ELA-L for all students. However, the public school evidence base, particularly in the younger grades, remains scant. Fisher and Frey (2012) conducted an observational study of close reading implementation in secondary school settings. The stated goals of the inquiry were to evaluate the appropriateness of the close reading instructional routine for use with elementary school students and to determine modifications that would be useful in implementing close reading with students in elementary school. Fisher and Frey (2012) indicated that findings supported the assertion that close reading practices were appropriate for elementary school settings. Use of short, varied, complex (at or above grade level) texts was described as an appropriate instructional
routine in elementary classrooms as long as the texts were read multiple times and students were afforded opportunities to provide text-based responses to teacher or student inquiry. Modifications for elementary grade classrooms included the teacher engaging in reading of text in certain incidences and the use of limited preteaching activities, such as with difficult vocabulary.

Fisher and Frey (2014c) used teacher and student interviews of close reading practices to better understand close reading implementation. Teachers reportedly questioned their own teaching abilities, struggled to locate texts, and wondered about how the lack of preteaching impacted their students, yet voiced optimism that the practice could help their students over time. Students, meanwhile, were said to express the view that the multiple readings of complex texts, complete with discussions and annotations, were “mentally exhausting” (Fisher & Frey, 2014c, p. 33). Both parties reportedly noted that the texts used were interesting; there was a focus on the right answers to questions; and the practice required strong effort on the part of all.

The lone experimental study to date (Fisher & Frey, 2014a) was directly aimed at middle school students at risk for academic failure who were being served in an afterschool program. Fisher and Frey (2014a) targeted 438 students in grades 7–8, with 100 of the total randomly assigned to one of five intervention classrooms. The treatment involved close reading instruction; independent reading; and small-group, teacher-led instruction in areas of instructional need related to vocabulary, comprehension, or fluency, whereas control group activities involved a combination of computerized interventions; teacher-led, small-group instruction; and independent reading. Findings noted increases in student attendance, better reader perception survey scores, and improved statewide reading test scores following the yearround intervention.

Problem Statement

In spite of the advocacy by CCSS and/or close reading proponents, there is reason to proceed cautiously (Hinichman & Moore, 2013). In the present context of school-based personnel who are responsible for implementing CCSS and RTI, two major concerns emanate. First, the present body of evidence supporting implementation of close reading practices is lacking. Second, and equally concerning, is the fact that an implementation protocol for close reading has yet to be settled upon (Fisher & Frey, 2015). In the research described herein, there were no checklists used and processes described to ascertain the degree to which all of the components of close reading (e.g., multiple readings, annotation, text-dependent questions) were implemented or not, as well as whether or not the implementation was effective. These concerns were weighed against the reality that the CCSS ELA were being implemented in public schools and that new national assessments addressing CCSS ELA-L (e.g., PARCC) was administered nationally for the first time in the spring of 2015. These concerns were balanced by the positive potential of close, analytical reading practices (e.g., Fisher & Frey, 2014a) to provide students at risk with ever deeper understandings of complex informational and narrative text, as well as any potential accompanying positive impact on students’ college and career readiness. Utilizing a close reading protocol adapted from Shanahan’s (2014) description, the present study addressed the following research questions:

1. What were the effects on fourth-grade student reading comprehension achievement of implementation of close reading and Collaborative Strategic Reading (CSR; Klingner, Vaughn, Dimino, Schumm, & Bryant, 2001), a validated intervention, in RT1 Tier 2 programming utilizing science content instruction?
2. What did students like and dislike about close reading and CSR following implementation?

Method

Participants

Participants were five fourth-grade boys and one girl who were recommended by their classroom teacher due to risk of academic failure. Risk status was determined by teacher recommendation and was based on previous poor performance on state accountability tests and at-risk scores on the fall benchmarking assessment (i.e., less than 70 words correct per minute on grade-level oral reading fluency [ORF] probes for the fall benchmarking period). All African American students were receiving Tier 2 supplemental reading intervention at the time of the study.

Interventions

Two interventions were compared in the present study, close reading and CSR. The researchers chose to make the comparison because at the time of the study, close reading was not an empirically validated intervention for elementary grades students. That, therefore, made close reading an inappropriate choice for a Tier 2 intervention program, which is designed to utilize small group formats and research-validated interventions (Fuchs et al., 2012). In order to ensure that students received a validated intervention as part of the program, a decision was made to include CSR as part of the Tier 2 programming and compare it against a close reading instructional routine that was based on the description of Shanahan (2014) and the qualitative research of Fisher and Frey (2012).

Close reading. Table 1 outlines the close reading instructional routine utilized, which included 10 elements of teacher and/or student action. An active routine was prescribed, consisting of three 30-min sessions over three days in one week. The teacher opened the lesson with introductory remarks about her expectations and a purpose statement. The teacher and students then shared responsibility over the three days for conducting multiple reads of the same text. Following the suggestion of Shanahan (2014), the first reading involved determining what the text said; the second, figuring out how the text worked; and the third, analyzing the text and making connections to
<table>
<thead>
<tr>
<th>Elements of Close Reading</th>
<th>Possible Behaviors Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher opens with introductory remarks and purpose statement</td>
<td>Teacher introducing close reading text, lesson purpose and/or students’ responsibilities</td>
</tr>
<tr>
<td>Teacher conducts multiple reads of the same text</td>
<td>Close read lesson plans including previously developed text questions</td>
</tr>
<tr>
<td></td>
<td>Instructing students to reread the text (or sections of it) with a different focus in mind</td>
</tr>
<tr>
<td>Students actively engaged with complex text</td>
<td>Students annotating complex text while reading</td>
</tr>
<tr>
<td></td>
<td>Students referring to their annotations during discussions</td>
</tr>
<tr>
<td>Students engage in meaningful discussion</td>
<td>Students agreeing or disagreeing with each other and/or teacher during group discussions, citing textual evidence to support their claims</td>
</tr>
<tr>
<td>Students annotate complex text</td>
<td>Students writing on the actual text to communicate/record/capture their thinking</td>
</tr>
<tr>
<td>Lesson experiences connect to the close read text</td>
<td>Teacher redirecting “off text” discussions</td>
</tr>
<tr>
<td></td>
<td>Students giving textual evidence/support when adding to the discussion</td>
</tr>
<tr>
<td>Teacher comments/actions support students’ learning</td>
<td>Teacher leading “Think Aloud” on a section of complex text to demonstrate how to annotate</td>
</tr>
<tr>
<td></td>
<td>Teacher explicitly praising and/or expanding student(s’) comments/writing</td>
</tr>
<tr>
<td>Students use text during close read activities</td>
<td>Students revisiting/relying on annotated text to successfully complete an activity</td>
</tr>
<tr>
<td>Teacher formally or informally checks for student understanding</td>
<td>Teacher keeping a checklist of who’s participating</td>
</tr>
<tr>
<td></td>
<td>Teacher recording accuracy of student’s response for further reflection</td>
</tr>
<tr>
<td></td>
<td>Teacher administering online vocabulary check</td>
</tr>
<tr>
<td>Students engage in collaboration</td>
<td>Students discussing and making decisions about their thinking while working with other students (not directly with the teacher)</td>
</tr>
</tbody>
</table>
science classroom reading and students’ own background knowledge. Reading on day two also involved teacher-led attention to critical vocabulary. The selected science passage was chosen to align with the current topic of study in class. Students were actively engaged in the reading of the text by reading, listening, and/or annotating text. They were also part of teacher-directed discussions that were based on text-dependent questions that addressed different reading purposes.

**CSR.** A reading comprehension strategy instruction intervention for students with, or at risk for, academic disabilities, CSR is an evidence-based practice for students with learning disabilities (Vaughn et al., 2011). Students are taught a soft-scripted body of 17 lessons relating to before-, during-, and after-reading strategies (i.e., predicting, monitoring, developing main ideas, and summarizing) in a scaffolded format, beginning with interventionist modeling and then proceeding through guided and independent practice. Students are also taught and expected to carry out collaborative roles.

As with close reading, when CSR was implemented, students read a single informational passage three times per week with one reading taking place per session. During the first week of CSR, preview, brainstorming, and “clicks and clunks” were introduced. Students were given a science text to read and discuss. Prior to reading, students brainstormed what they may already know about the topic and predicted what they may learn after reading the article. Students used pictures and graphs as part of the process. Students took turns reading aloud and were encouraged to highlight words or terms that were “clunks.” Clunks are words that are unfamiliar to student readers and interrupt comprehension. During the second week of CSR, the “fix-up strategies” of clunks were continued, which included reading sentences around the clunk sentence looking for comprehension clues, as well as examining the clunk for affixes that might facilitate understanding. Students were taught to annotate or take notes when they come to clunks to help them understand what they read. The final week of CSR, students were introduced to the “get the gist” strategy.

**Treatment Fidelity**

Treatment implementation fidelity observations were scheduled to be conducted once weekly over the course of the six-week study, for a total of 33% of the 18 sessions. Each intervention was to be observed three times by the author familiar with both interventions. Researcher-developed checklists were used during the observations. Each checklist contained a specific number of intervention components that were marked in terms of whether or not the component was implemented during the intervention session. The close reading checklist consisted of the 10 components listed in Table 1. The CSR checklist included different forms to account for the four comprehension strategies that were implemented over the course of the experiment. Data were reported as the proportion of components observed.

**Dependent Variables**

There were two measures used as dependent variables in this study: (a) Daze (Dynamic Measurement Group, 2011); and (b) critical content monitoring (Mooney, McCarter, Russo, & Blackwood, 2013). Daze was administered weekly, while critical content monitoring was administered at pre- and posttesting.

**Daze.** Daze is a group-administered measure of reading comprehension in which students were asked to read a passage silently. In the passage, approximately every seventh word is removed and replaced with three choices, one of which is correct. Daze requires students to choose the correct word as they read the passage. Students were given 3 min to work on this task. The score was the number of correct words circled minus one-half of the number of incorrect words the student circled. Daze was chosen because it assesses reading comprehension, is group administered, and has adequate reliability and validity (Wayman, Wallace, Wiley, Ticha, & Espin, 2007).

**Critical content monitoring.** Critical content monitoring is an online, timed general outcome measure of content knowledge (Mooney et al., 2013). Students were expected to read a stem/question and then select the best answer by placing a mark beside the correct choice. Scores consisted of the number of correct choices in the time frame. Correct score totals were provided to students in the form of a fraction (such as 6 out of 20) at the completion of testing. Mooney et al. (2013) reported moderately strong criterion validity findings for the measure when compared with statewide accountability test results. Critical content monitoring was included as a measure of content learning.

**Social Validity**

Researcher-created social validity surveys were administered to all students at posttesting. Students rated their experiences with the interventions from 1 (not at all) to 9 (liked very much) for 10 items. For item 11, students were asked to circle the strategy they preferred: close reading or CSR. For item 12, students were asked to circle the strategy in which they learned the most science content: close reading, CSR, or both.

**Procedures**

Parent consent was secured prior to initiation of the intervention. Students provided assent to participate. The first experimental activity was administration of the dependent measures for pretesting. Then, the intervention order of presentation was determined. Each treatment was implemented once every two weeks over the six-week time frame with a coin flip determining which intervention was first.

The school’s Tier 2 program included 60 min of supplemental services daily and was taught by a certified teacher who held a master’s degree. Students received supplemental instruction in place of their regular physical education class. The Tier 2 curriculum used was the Harcourt Storytown Strategic Intervention Program (Harcourt School Publishers, 2009), supplemented by a computer-aided reading intervention. The Harcourt program was used and lessons were
selected to reinforce what was being taught in students’ reading class. The four components that were targeted during intervention were phonics/phonemic awareness, comprehension, vocabulary, and fluency. In a typical Tier 2 intervention session, students began by engaging in vocabulary word exercises. Students read a story and completed fill-in-the-blank activities using the vocabulary words. They answered two questions at the end of the exercise and then shared their answers with the teacher. Next, students were guided through the comprehension task of drawing conclusions. A student volunteer read the passage and the teacher asked questions. The teacher guided students through the use of a graphic organizer to promote comprehension. Students targeted decoding/spelling skills using words with suffixes. Before wrapping up the lesson, students participated in fluency practice where they were instructed to read words in a column from their text aloud to each other. They then practiced reading the column of words to their partners. Two days out of the week, students were given the opportunity to substitute a computer-aided reading intervention for the Harcourt program.

Prior to the intervention, the students were administered the CCM pretest, which was a timed assessment. When the intervention began, students were pulled for the first 30 min of their Tier 2 intervention time three days a week. The intervention took place in an empty classroom. On the first day of interventions, students selected their seating positions at a large table, remaining in those seats for all of the subsequent sessions. At the end of each week of implementation, students completed the Daze measure. Once the intervention component of the experiment was complete, students completed posttesting and the social validity survey.

Design

A single subject research alternating treatments design was utilized to answer the first research question. Use of an alternating treatments design allowed for a direct comparison of the effectiveness of close reading and CSR on students’ reading and writing performance. Visual analysis was the interpretation approach utilized (Alberto & Troutman, 2009). A descriptive summary was utilized to address the third (social validity) research question.

Results

Interscorer Reliability and Treatment Fidelity

The Daze student response protocols were scored by independent parties and their respective total scores compared. For Daze, agreement was 100%. As score agreement with an independent scorer was greater than 80%, scores of the first author were graphed. Treatment fidelity was completed by a graduate student who was trained in both procedures. Collaborative Strategic Reading received three observations and close reading received two. For CSR, fidelity ranges on separate checklists related to three independent reading comprehension strategies. Fidelity ranged from 60% for both of the reading strategies (i.e., get the gist; click and clunk) to 100% for the preview (before reading) strategy. The overall average across three observations was 75.6%. For close reading, the average rating was 95% (range 90%–100%).

Content Learning

Comparison of pre- and posttest scores for the content-focused general outcome measure critical content monitoring showed performance improvements in four of six cases. Figure 1 showed that Students 1, 3, 4, and 5 had larger scores at posttest, with Student 6’s performance showing no change and Student 2’s evidencing score decreases. Due to the nature of the design, it was impossible to determine if achievement gains were impacted differentially by intervention.

Research Question 1: Effects on Reading Comprehension Achievement

Visual analysis indicated a pattern of no clear separation between the two interventions across the six students (see Figure 2). Only in the case of Student 6 was there no overlap across conditions from start to finish. Four of the six students indicated increasing Daze achievement trends across conditions, with growth from first to last testing for both interventions. While overall results were mixed, with no intervention evidencing clear separation across conditions, there was one uniform finding. In all six cases, the Daze score for CSR was higher than that for close reading by intervention’s end. Related to that, in four of six cases—Students 1, 2, 5, and 6—the largest differences between intervention data points were at intervention’s end. That is, the difference between Daze scores for CSR and close reading was greatest for the last two data points.

Research Question 2: Likes and Dislikes of the Interventions

Participants responded favorably to both interventions. In terms of rating each intervention as a whole, close reading received a slightly higher mean score than CSR, with all participants rating the intervention in the highest third of the nine-item Likert scale. However, the
Figure 2. Daze scores for close reading and collaborative strategic reading.
close reading components were viewed less favorably than were CSR elements. The highest mean score of 9.0 came for the CSR click and clunk comprehension monitoring strategy. The lowest mean score (i.e., 5.5) related to the annotating/note taking component of the close reading intervention. In fact, the three lowest-rated items were related to close reading components. If students were given the opportunity to select the intervention, four out of six selected CSR over close reading. However, when students were asked which intervention helped them learn more science content, five students felt that both interventions prepared them equally well.

Discussion
Close, analytical reading is a promoted and potentially promising instructional practice that, to date, lacks an evidence base of empirical support to match the professional promise. For purposes of the present study, the evidence for close reading includes contexts with elementary school students, students at risk for academic failure, and informational text. The evidence supporting the efficacy of close reading in relation to these areas thus far has been scant. What follows is a summary of the findings from the study and how they fit within the larger literature, a description of the study’s limitations, and a presentation of implications and future research needs.

Summary of Research Findings
The predominantly nonexperimental close reading literature previously described has generally been supportive of the instructional routine advocated for CCSS ELA-L use. The lone experimental study (Fisher & Frey, 2014a), implemented by close reading advocates in an afterschool program for at-risk middle school students, itself evidenced supportive findings. Yet, while qualitative inquiry with secondary close reading teachers has indicated that the practice can be successful if adapted in elementary classrooms (Fisher & Frey, 2014c), the results from our experimental inquiry suggest that the caution proposed by Hinchman and Moore (2013) is indeed warranted.

Academically at-risk students in today’s public schools are likely served in some form of RTI framework. In effectively implemented systems, at-risk students receive a research-based core curriculum plus supplemental research-validated small-group intervention that is regularly evaluated to determine if the combination is successfully advancing students academically in comparison to same-age peers. Tier 2 (supplemental) programming needs a valid research base to support its implementation. That is why a nonvalidated close reading instructional routine (see Table 1) was compared to the validated CSR reading comprehension in an alternating treatments experimental design. The close reading instructional routine implemented herein in an elementary RTI Tier 2 program did not measure up to CSR over time. Scores for all six students in the study favored the CSR condition, with the largest differences between interventions recorded by study’s end in two thirds of the cases. What is more, the complete CSR intervention was not implemented in the experimental time given, so that its impact over time may have been greater if full implementation had occurred.

Potentially positive findings were threefold. First, students expressed satisfaction with the close reading instructional routine as implemented, which might facilitate student buy-in over time, with increased engagement leading to better results. Second, in four of six cases, student trend lines for the close reading routine were positively leaning in the short amount of time that the intervention was implemented. Moreover, those findings came about through implementation by a practicing teacher who was not an expert in close reading. Experts had been implementers in the lone other experimental study (Fisher & Frey, 2014c). Finally, student knowledge in science content utilized as part of the Tier 2 program grew as measured by a general outcome measure of content knowledge. That is, posttest critical content monitoring scores for four of six students were higher than pretest scores after six weeks of implementation.

Limitations
The results of this study should be viewed with caution. First, the length of the study was short and may not have allowed for a clear pattern to develop in terms of separation between the two interventions. Second, interpretations for alternating treatments designs are considered weak until experimental control is determined following a successful implementation of the treatment of choice. Third, the fact that critical content monitoring was not implemented weekly did not allow for an evaluation of whether close reading or CSR contributed more to the pre-to posttest growth in content achievement scores. Finally, the low treatment implementation proportions for CSR may have influenced the weekly Daze scores that were collected from participants. With only three weeks and 90 min total of intervention time devoted to CSR implementation, not all of the components of the 17 lessons were adequately addressed.

Implications
With the study results and limitations noted, some takeaway points revolve around a central theme, which is that research rather than advocacy needs to drive the discussion of how close reading practices are to be implemented in elementary schools and content courses and with students at risk for academic failure. The need for caution in moving forward in the implementation of close reading that was expressed by Hinchman and Moore (2013) and Fisher and Frey (2015) is warranted, particularly as it relates to the elementary grades, elementary students, and informational texts. At this point, there are two experimental studies targeting close reading, with only the presenting study involving elementary students.

A starting point for systematic inquiry has to be development of a detailed operational checklist that clearly outlines a testable close reading instructional routine. First, the descriptions of how it can be implemented seemingly fit well within an evidence-based effective instruction paradigm (Rosenshine, 2012). Research-based practices include
presenting new material in small steps and practicing after each step, asking a large number of questions and checking for all students' understanding, providing models, guiding student practice, and obtaining high student success rates. Such practices are particularly pertinent to students at risk for academic failure. Second, a product in the form of a validated intervention is necessary if a close reading instructional routine is to be included in the prototypical RTI Tier 2 system. Tier 2 programming incorporates research-validated practices that presumably target areas of student weakness and are delivered in small group settings in addition to the core instructional program. Experimental studies need to be conducted which include findings indicating positive effects for an accurate implementation of the clearly described treatment in comparison to a similar intervention or business-as-usual condition. Therefore, in order for what could be considered a research-informed close reading instructional practice to be implemented in Tier 2 (and 3) settings with students at risk for academic failure as well as students with disabilities, there needs to be a clearly outlined and validated instructional practice or protocol that has been validated for the setting and students served in Tier 2 (or 3) programs.

References
Shanahan, T. (2014). This is not close reading but we’ll tell you what is. Scholastic Instructor, Winter 2014, 28–29.
Authors
Marcy Boudreaux-Johnson, PhD, Louisiana State University. Dr. Boudreaux-Johnson’s areas of research interests are sensory processing disorder and its effects on education, at-risk readers in elementary grades, and inclusive learning environments.

Paul Mooney, PhD, is an Associate Professor in the School of Education at Louisiana State University. Dr. Mooney’s present research interests focus on structured formative assessment and academic interventions for students with or at risk for academic or behavioral disabilities.

Renée E. Lastrapes, PhD, serves as an Assistant Professor in the Department of Educational Research & Assessment at the University of Houston – Clear Lake. Dr. Lastrapes’ research interests are assessment for students with specific learning disabilities and evidence-based behavioral and academic interventions for teachers working with students with emotional and behavioral disorders.