The Effects of Computer-Assisted Culturally Relevant Repeated Readings on English Learners’ Fluency and Comprehension

Mariah Barber
Paradise Valley Unified School District

Gwendolyn Cartledge
The Ohio State University

Morris Council III*
University of West Georgia

Moira Konrad
Ralph Gardner
Alana O. Telesman
The Ohio State University

This study examined the effects of a novel computer software program, Reading RACES (Relevant and Culturally Engaging Stories), on the oral reading fluency and comprehension of three urban first graders who were English language learners (ELLs) and showed reading/special education risk. The individually administered intervention consisted of repeated readings of culturally relevant passages delivered through computer software. A multiple baseline across participants revealed that after intervention, the participants increased their oral reading fluency and comprehension of novel culturally relevant (CR) passages, and to a slightly lesser extent, generalized these skills to AIMSweb (non-culturally relevant, NCR) passages. These findings are discussed relative to study effects for ELLs and the usefulness of this technology in urban classrooms.

Keywords: computer-assisted instruction, culturally relevant, English Language Learner, oral reading fluency, reading comprehension, repeated reading, urban learners.

Many children in the United States are not proficient readers and schools are particularly failing urban minority populations where 50% are likely to be reading below basic (National Assessment of Educational Progress; NAEP, 2013). Basic is defined as, “partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at each grade” (NAEP, 2012, para.4). Even more disturbing is the poorer performance of English language learners (ELLs) whose average reading scores fall below those of native English speaking minorities as well as White children (NAEP, 2013). If children are not competent readers by third grade, the chances of remediation are poor (Moats, 2007; NCES, 2007). Children need to be taught explicitly the five critical building blocks of reading: phonemic awareness, phonics, reading fluency, vocabulary, and reading comprehension (National Reading Panel, NRP, 2000), which are interrelated and build upon each other. For example, fluency
training is most effective if children have the requisite decoding skills (Therrien & Kubina, 2008), and students are less likely to comprehend text, if they are not reading fluently (Yopp & Yopp, 2009). In research, fluency has been positioned as the link between decoding and comprehension, yet is often overlooked in practice (Copple & Bredekamp, 2009).

Speece and Ritchey (2005) described fluency as “the speed and accuracy with which text is read orally” (p. 387). Elsewhere, fluency is defined as reading with speed, accuracy, and prosody (National Institute of Child Health and Human Development, NICHHD, 2000). NICHHD (2000) conducted a meta-analysis and found that fluency depends upon well-developed word recognition skills, which, in turn, depend on decoding skills and practice. The researchers concluded that fluency, which needs to be taught, increases with reading practice. They also contended that guided repeated oral reading procedures were effective in improving reading fluency and overall reading achievement. Considerable research exists on the many variations of evidence-based guided oral repeated reading procedures (e.g., Ardoin, Eckert, & Cole, 2008; Ardoin, McCall, Klubnik, 2007; Baker et al., 2008; Begeny, Daly, & Valleley, 2006; Foorman et. al., 1998; Therrien & Kubina 2007). Overall, this research has pointed towards the efficacy of repeated reading interventions (RRI) (see Strickland, Boon, & Spencer, 2013; Therrien, 2004), involving verbal modeling of a passage, participant practicing the modeled passage, and correcting participant errors during intervention (Therrien, 2004).

**English Language Learners**

Students who have a birth or home language other than English (i.e., ELLs) are steadily increasing in U.S. schools (Castro, 2014; Kena et al., 2014) and, with the exception of children with a Hispanic background, the research on language and literacy with ELLs is particularly scarce (Fitzgerald, 1995). Several studies employed RRI for fluency development showed positive outcomes with native Spanish-speaking learners. Landa and Barbetta (2017), for example, used RRI to assess the fluency and comprehension growth of four third- to fifth-grade Hispanic background English learners with learning disabilities (LD) and obtained reading gains, reading generalization, and maintenance. Some researchers paired RRI with vocabulary instruction resulting in good effects for the RRI and modest or mildly discernable benefits of the language components (Johnston, Mercer, & Geres-Smith, 2018; Leacox & Jackson, 2014). The technology of pre-recorded readings and parent tutoring were emphasized in another RRI study with eight- and nine-year-old Hispanic English learning children, who made steady progress on practiced and untutored passages (Kupzyk, McCurdy, Hofstander, & Berger, 2011).

Much of the research on children who are language different has failed to distinguish English learners according to ethnicities (Hammer et al., 2014), even though these groups with different language backgrounds vary greatly (Winsler et al., 2014). Betts, Bolt, Decker, Muyskens, and Marston (2009), for example, found that the amount of time spent in the U.S. affected reading performance, including reading fluency, for children from Somalia but not for English learning Hispanic children. The researchers attributed this difference to the fact that Somali children, in contrast to Hispanic children, had much less of their culture represented in the dominant
U.S. culture. They also noted that although both groups of third graders made steady progress in reading fluency over the course of the year, they both were well behind their monolingual classmates and warranted direct, explicit interventions in English language and reading strategies. In one of the few published studies including children from a Somali background, the researchers provided fluency training for phoneme segmentation and nonsense word skills to 12 kindergarten ELL students (10 with Somali backgrounds) (Author et al., 2009). The multiple baseline designs showed experimental control and good progress for all of the students; however, more than half (n=7) of the students showed zero response to these fluency (i.e., correct words per minute, CWPM) skills until the intervention was introduced. This finding underscores the importance of direct, systematic instruction for non-responding ELLs. Likewise, Kai, Heward, and Heng (2006) used direct instruction and a similar design to demonstrate increases in Oral Reading Fluency (ORF) for first-grade ELLs. The five students targeted in the Kai et al. study represented Ethiopian, Hispanic, and Cambodian backgrounds. Research findings showed improved fluency, comprehension, and maintenance for all students. Although limited, these studies suggest the potential benefit of direct RRI interventions, the need for more research, and the need to study ELL groups according to their unique characteristics.

**RRI in the Schools**

There is a growing body of evidence for using RRI to increase ORF skills in young children both with and without disabilities (Author et al., 2014; Therrien, 2004). In one of the more recent reviews for students with LD, Stevens, Walker, and Vaughn (2017) concluded that RRI continues to be the most effective means for improving fluency for students with LD. They also noted that effective outcomes are those that evidence improved comprehension and generalization to novel texts. However, little, if any, of this research has occurred in applied settings (i.e., classrooms), which can often be challenging. Therrien and Kubina (2008), for example, provide an eight-step sequence, which may be difficult for a teacher with limited training to implement with fidelity. Further, RRI requires one-on-one or small group instruction from a competent reader (i.e., an adult or peer), a particularly difficult problem in urban settings with limited resources. In addition, many teachers are unsure when or where to fit RRI into an already rigorous curriculum, and may not be sure what reading material to use when implementing RRI. For example, many state standards mandate that teachers use on-grade level and complex reading materials. However, students who are practicing reading fluency need to read at their instructional level (not at the frustration level), or results may not be favorable (Burns, 2007). Selection of reading material, especially for struggling readers and readers who are culturally and linguistically diverse (CLD) may be another reason implementation of RRI in the classroom is not widespread.

**Culturally Relevant Reading Materials**

Urban schools in the U.S. often have student populations characterized by cultural, linguistic, socioeconomic, racial, and ethnic diversity. Accordingly, school personnel are challenged to utilize meaningful learning materials, particularly for diverse learners who show learning risk. Several authorities argue for the importance
of culturally relevant (CR) materials as a way to affirm, motivate and engage such learners, and to mitigate negative stereotypes (Bishop, 2007; Ebe, 2010; Ma’ayan, 2010). Culturally relevant materials are those with which the learner can identify. Ebe (2012) stated that literature should draw upon the student’s background, experiences, culture, and language. CR literature can be especially valuable for the CLD learner, particularly those with reading risk. Ebe cited studies that provided evidence that CR texts can encourage greater reader engagement and support reading comprehension. Ebe documented her own research findings showing better reading performance for elementary (2010) and adolescent (2012) ELLs with CR texts compared to non-culturally relevant (NCR) materials. In our own studies, we found significantly higher fluency rates with CR passages than with NCR passages for young African American students in urban settings (Author et al., 2015).

To aid in determining cultural relevance, Ebe (2010; 2012) developed a rubric that asks a series of eight questions relative to how closely the reader connects with the story (e.g., Do the characters talk like you and your family?). As discussed elsewhere (Author et al., 2015), we obtained content from the children, their teachers, their parents, contemporaneous culture, and the children’s historical background to create the reading passages used in this study. Like Ebe, we also employed a rubric that required the stories to include at least two elements from the following: (a) affirming events that allow the children to see themselves in realistic, attractive ways, (b) cultural experiences that are familiar to the children (e.g., familiar foods), (c) famous personalities with positive images in the larger culture (e.g., the President looks like me), and (d) positive historical events or figures such as scientists (e.g., George Washington Carver). The use of meaningful text has played an important role in the RRI literature since Samuels’ pioneering research (1979). For students who are CLD, CR material may serve an important role in bridging fluency of practiced CR passages to unpracticed NCR passages.

The main purpose of this study was to examine a multi-component intervention that embeds CR material within RRI delivered through computer software that focuses on errorless learning and learner independence. A particular focus in this study was the beneficial effects of this intervention with ELLs. Therefore, the researchers questioned if this intervention would result in increased ORF and comprehension on novel CR passages for first-grade ELLs. A second related focus was whether the newly acquired fluency and comprehension abilities would generalize to novel NCR passages. A third question was the extent to which increased fluency and comprehension levels would be maintained. Finally, the researchers examined pupil independence by assessing how well the learners could interface with the computer and proceed through the instructional sequence with minimal adult prompting. Several recent studies confirm that technology can be used to deliver effective fluency instruction to urban first- (Author et al., 2011; 2014) and second-grade (Author et al; 2017a; 2017b) students, as well as Hispanic ELLs (Kupzk, et al., 2011; Leacox & Jackson, 2014). Student gains, notwithstanding, there is a need to examine the viability of these interventions with a variety of populations, particularly ELLs (e.g., Author et al., 2015b; Nemeth & Simon, 2013) as well as ascertain the degree to which this technology provides a critical classroom resource to benefit both students and teachers.
Method

Participants and Setting

Participants for this study were recruited from an inner city elementary charter school located in a large Midwestern metropolitan area. The school enrolled 219 low socioeconomic students where 80% were Somali, 18% African American, and 2% Hispanic background. We included three first-grade participants, one male and two females. All three had a Somali background, were designated as ELLs, and showed reading/special education risk. Nico was repeating first grade and evidenced extreme difficulty in reading connected text smoothly and accurately. His English proficiency as determined by his Ohio Test of English Language Acquisition (OTELA), which is a measure of English language acquisition in four domains: Listening, Speaking, Reading, and Writing. A score of two on a five-point scale demonstrated he was slightly above pre-functional or “beginning proficiency.” Arya and Isabelle were in their first year of first grade and although performing below grade level, demonstrated more skill than did Nico. They each demonstrated advanced English proficiency as determined by their scores of four. Each participant had been enrolled in the school one to two full academic years prior to the study. Reading risk and language development were based on assessments as described in the next section. Demographic information can be found in Table 1. The interventions of this study took place in the school’s library/resource room.

Table 1. Participant Information

<table>
<thead>
<tr>
<th>Name</th>
<th>Age</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nico</td>
<td>7-10</td>
<td>M</td>
</tr>
<tr>
<td>Arya</td>
<td>7-3</td>
<td>F</td>
</tr>
<tr>
<td>Isabelle</td>
<td>6-10</td>
<td>F</td>
</tr>
</tbody>
</table>

*a Age: refers to the age of each participant in years and months at the time DIBLES Next Middle of the Year (MOY) sub-tests were administered

Materials

Screening measures and instruments. First graders who showed reading/special education risk and who had parental permissions were screened on the DIBELS NEXT (Good & Kaminsky, 2011) and the Woodcock Reading Mastery Test – Revised (WRMT-R, Woodcock, 1998). Researchers followed the prescribed test procedures and administered the DIBELS NEXT subtests of Nonsense Word Fluency (NWF), Correct Letter Sounds (CLS), and DIBELS Oral Reading Fluency (DORF). They also administered the WRMT-R subtests of Word Attack (WA) and Word Identification (ID) sub-tests. To qualify for the study students needed to demonstrate decoding skills (DIBELS Next NWF-CLS scores at or above 18) and oral reading failure (DORF scores at or below 23). The WRMT-R (Woodcock, 1998) was used to assess reading performance but was not used to qualify students for this study.
For the WRMT-R subtests, reliability at the first-grade level for letter-word identification was $r_{11} = 0.90$ and for word attack was $r_{11} = 0.97$. For first grade, the DIBELS reported alternative-form reliability was as follows: NWF 0.85 for a single test and 0.94 for the three-test form; and 0.91 and 0.96 for the single- and three-test forms, respectively for the DORF (Dewey, Powell-Smith, Good, & Kaminski, 2015). The OTELA rated knowledge of English in five levels, 1 being pre-functional and 5 being full English proficiency. Estimated form reliability coefficients for the OTELA for first and second graders were reported as listening 0.94, speaking 0.95, reading 0.96, and writing 0.94 (OTELA, 2010).

**Culturally relevant passages and mazes.** This study was part of a larger study, which involved the development of the CR passages (Authors et al., 2015a; b). The research team created a novel set of CR passages designed to reflect the interests and background of the target population (i.e., young learners in urban settings who were at risk for reading failure). The passages varied in length from 100 to 120 words with a grade range on the Spache Readability Formula (n. d.) from 2.0 to 2.3. The researchers developed a maze passage for each CR passage to assess reading comprehension. Specific procedures for constructing and equating the passages and mazes are given elsewhere (i.e., Author, 2015a).

**Generalization passages.** The generalization passages were taken from AIMSweb (aimsweb.pearson.com), which is a curriculum-based measurement management system that provides for screening and progress monitoring of basic academic (i.e., reading, language arts, math) and social skills. These passages, viewed as NCR, were chosen based on grade level difficulty (1.7–2.3) as measured by the Spache (n. d.). Eighteen kindergarten stories, six first-grade stories, and three second-grade stories were used in this study. Passage length ranged from 220 to 280 words. The researchers used random.org to randomize the order in which stories were presented to each student.

**Researcher-developed software.** A researcher-developed computer application called Reading RACES (RR) was used to deliver the RRI to the first graders on two Toshiba laptop computers. The program allowed and prompted students to listen to a human voice model, instructed students to read with that model, and had the capability to “listen” to students as they read independently. It also had the capability to calculate total words read (TWR) on one-minute timings based on voice recognition during these independent readings when students verified the last word read by clicking on the word. A headset with a noise-cancelling microphone was used so that participants could listen to the stories with limited distractions and the computer created an audio recording of the students’ oral reading. Using this information and researcher verification, the computer generated a CWPM score and displayed it for the student. During intervention, while the participant was reading practice passages, the computer software provided assistance with unknown words (i.e., reading the word), when a student clicked on the unknown word or when there was a 3-second pause, it prompted the student to click on the word or to continue reading.

**Definition and Measurement of Dependent Variables**

**Fluency for CR passages.** A word was counted correct when a participant
pronounced it within 3 seconds of the previous word being read. A word was counted as incorrect if the participant omitted or mispronounced the word or failed to read it within 3 seconds. Additions (i.e., adding or repeating a word) and self-corrections were not counted as errors. During the timed reading phase, neither the instructor nor the computer gave corrective feedback.

**Correct CR mazes for comprehension.** Maze assessments took place during a 3-min timing. The computer scored a maze comprehension item as being correct if the response given by the participant matched the key and scored incorrect if it did not match the key. Feedback was not given after a maze assessment. The computer then produced a graph of the student's scores after each session.

**Fluency on AIMSweb generalization probes.** AIMSweb passages were inserted into RR. The participant received a generalization probe after three to five successfully completed CR timed readings. Generalization passages were novel AIMSweb passages administered without the voice model on the computer. The same rules used to determine CWPM for CR passages were used for the generalization passages.

**Correct AIMSweb mazes for comprehension.** Correct responses to AIMSweb maze items (also inserted on the computer) were recorded following the generalization probe. Rules for accuracy were the same as those used for CR mazes. Similar to the treatment probes, generalization probes were plotted by the computer following the session they completed.

**Fluency maintenance on generalization passages.** One week after intervention was completed, participants read an AIMSweb passage and total CWPM was recorded. These data were collected using the same procedures as those used during the generalization assessments.

**Comprehension maintenance of generalization passages.** One week after intervention was completed, correct responses to AIMSweb mazes were measured using the same procedures used during the generalization assessments.

**Fluency maintenance on CR cold passages.** The computer plotted on a graph the student’s CWPM maintenance for CR passages. These data were collected using the same procedures as those used during the CR CWPM intervention assessments.

**Pupil independence.** Independent pupil use of the program and related behavior were recorded, when interacting with the computer software. After initial training, the researcher recorded the number of additional human verbal prompts needed for the student to complete the instructional sequence according to the researcher-created procedural integrity checklist.

**Experimental Design and Conditions**

A concurrent multiple probe experimental design across participants during intervention was used for this study. The median CWPM score on DIBELS NEXT ORF test was used to decide whether the student would enter the intervention in either the first, second, or third tier. Student middle- and end-of-year DORF scores can be found in Table 2. All students began baseline at the same time. Once baseline data were stable, the first student entered into intervention and following the first generalization probe, the data were assessed for trends and
continued stability. Once there was an increasing trend on the intervention passages, the next participant entered the intervention, and so on. All participants entered into intervention condition 1 to 5 weeks after the baseline condition was complete. Each session included 1 to 3 participants and 1 to 3 researchers. There were 1 to 4 sessions a week with each session lasting 20 to 30 minutes. Participants were in intervention from 7 to 11 weeks.

Table 2. DIBELS Next Scores on Oral Reading Fluency for Middle of the Year, and End of the Year

<table>
<thead>
<tr>
<th>Name</th>
<th>MOY (Percentage Growth)</th>
<th>EOY (Percentage Growth)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Raw Score</td>
<td>Risk Level</td>
</tr>
<tr>
<td>Nico</td>
<td>6</td>
<td>Well Below Benchmark (100%)</td>
</tr>
<tr>
<td>Arya</td>
<td>10</td>
<td>Well Below Benchmark (250%)</td>
</tr>
<tr>
<td>Isabelle</td>
<td>18</td>
<td>Below Benchmark (172%)</td>
</tr>
</tbody>
</table>

Baseline. Classroom instruction included whole group instruction using a blend of phonics and whole word instruction. There was no available pull-out reading intervention offered at the school for struggling students. Further, there was no formal measurement used to track student progress in reading. The participants’ baselines consisted of the students reading AIMSweb passages (i.e., NCR passages used for generalization) and CR passages delivered on the computer. The number of passages read during baseline varied based upon the variability of student data. The students listened to the instructions of the computer and read the passage using the timed reading button as the researcher followed along recording the errors on a separate sheet. The researcher collected data on the students’ errors, CWPM, and number of items scored correctly on the maze comprehension passage. These data were used to verify the computer’s calculations and fidelity.

Training. The training phase included training of experimenters as well as training of student participants.

Experimenters. Two graduate students (one doctoral student and one master’s student) and one research assistant (early childhood undergraduate student) were trained to serve as both a primary experimenter and second observer. This training required the trainees to achieve 90% procedural integrity and 90% IOA for each timing before he/she was able to conduct or observe participants for this study. IOA was calculated using exact agreement. Specifically, agreements were divided by the number of agreements plus disagreements, multiplied by 100, and rounded to the nearest tenth.

Participants. During the session immediately preceding intervention, participants were trained on how to use the computer and RR software. The same CR story was used for all participants (i.e., Football in the Park). Based on a training checklist, all participants met the predetermined training criterion after one training session (i.e., they were able to demonstrate all the skills in the correct sequence).
Intervention and procedures. The intervention/independent variable was the researcher-designed repeated reading application with CR passages delivered through voice-recognition software with a human voice model titled Reading RACES. The following describes the intervention.

Goal setting and beginning. Each participant was shown his or her baseline data (presented as a graph on the computer) and a goal (i.e., 20 CWPM, 25 CWPM, or 40 CWPM) before the computer session began. The participant was prompted to put the headphones on and log into the computer by finding his/her initials. The computer then reviewed the previous day’s data and had the student adjust to the appropriate goal. According to DIBELS the oral reading fluency at which participants (i.e., first graders) are expected to read is 23+ CWPM by the middle of the year and 47+ by the end of the year. The researcher set the participant’s goal in the computer to an interval of 5, beginning from 20. The goals were slightly higher than the benchmark goals set by DIBELS NEXT to encourage transfer of ORF skills to generalization passages (i.e., AIMSweb). To choose the goal, the student’s goal was set at 20 if the median baseline was below 20, or the nearest interval if baseline was above 20. If the participant’s goal was between 20 and 40 CWPM the computer played the voice model for 40, if the participant’s goal was between 45 and 60 CWPM the voice model played at 60 WPM, if the participant read above 60 the computer provided the model at 90 WPM. Through a series of steps the student was directed through the computer to listen to introductions, find the correct passage and click on the go button. The 25 CR passages were divided into five sets of five stories. Each story set was randomly assigned to students and they progressed through the stories in the set. There was one expository text in each set, except one; the remaining passages were fictional.

Cold read. After selecting the first CR story, the student was directed to read the story as a cold read. If the student met his/her goal without practicing, he/she was congratulated and went to the next story at a slightly higher goal. If not, the student was told how many CWPM he/she read and told that he/she would reach that goal with some practice.

Practice words. After the participant read the story for one minute during the cold read, the participant went through prescribed practice words with the computer. The computer showed the student a word (e.g., football), and said, “This word is ______________. Say ___________.” Utilizing the voice-recognition component of the software, the computer then waited for the student to successfully say the word, then the word was read in a sentence, which the student also read. Participants engaged in error correction after each ORF timed reading (i.e., cold read, listen to me, and timed reading). During the Listen to Me and Timed Reading, participants reviewed specific errors made during their timing. The researcher manually selected the errors to be practiced by the participant.

Read to me. During the Read to Me session the computer played the pre-recorded human voice model for the assigned story at the participant’s goal speed (i.e., 40 or 60 CWPM). The blue-highlighting feature kept pace with the voice recording and the student was prompted to follow along.

Read along. In this condition, the student read the story along with the computer human voice recording. The student was directed to click on unknown
words(s) if the student needed help. In which case, the computer stopped, said the word, and started the sentence from the beginning.

**Listen to me.** During the next condition, the student was directed to click on the *Listen to Me* button and do his/her best and smoothest reading. Students again could click on unknown words for help.

The participants had a chance to practice three one-minute timings, trying to reach their set criterion with 90% accuracy. At the end of the 1-minute timing, the computer displayed the CWPM. If the participant did not reach his/her goal, the researcher highlighted the word(s) that the participant needed to read to reach his/her goal. The participant then had another one-minute timing.

If a participant did not qualify for the timed reading after three *Listen to Me* trials, the session ended and the participant started the *Listen to Me* phase over with the same story the next session. If a participant did not reach his/her goal after two days, the goal was lowered systematically for the third day. In this case, the goal was re-set to a criterion that was more attainable for the participant. To systematically lower a participant’s goal, researchers *sliced back* the goal by 5 CWPM. For example, if a student’s goal was 30, his/her goal was sliced back to 25. Once the student reached the new goal, the goal was raised to the previous level (e.g., 30) the next session.

If there was an announcement or other noise disturbance during the reading, the student hit the stop reading button, waited for the noise to pass, and re-started the timing. This was noted on the researcher data sheet.

**Timed reading.** The *Timed Reading* was also known as the treatment probe. During this phase, the computer gave the participant one minute to read the practiced passage and did not provide corrective feedback (i.e., participants were not able to click on unknown words or told a word after a 3-second pause).

If the participant reached his/her goal, the student was congratulated and then went on to the comprehension maze. If the participant did not reach his/her goal the first time, the student was directed to try a second time. If the participant did not reach his/her goal after the second Timed Reading, the session was stopped and the passages practiced the next session.

In total (including the read along, listen to me, and timed reading phases of the intervention) the participants had the opportunity to read the same story 5-8 times per session. The maximum number of times that one participant stayed on one story was three sessions (i.e., a participant practiced the same story no more than three days).

**Maze.** After the participant met his/her goal during the *Timed Reading*, the computer prompted the participant to select the *Maze Comprehension* button, provided directions, and timed the student for 3-minutes. There were approximately 11 to 15 response opportunities for each maze passage. Participants had three fill-in-the-bubble choices on the computer for each missing word. Stored computer data served as a permanent product and the computer calculated the correct score. Mazes were only taken one time and the computer then plotted the correct raw scores on a graph.

**Generalization.** After a participant met his/her goals on three consecutive CR stories or every 5 sessions, if their goal was not met, the participant read a *Generalization Probe*. The number of errors and correct words read per minute were
computed as generalization probes. After the participant completed the 1-minute timing on the AIMSweb (i.e., NCR) passage, regardless of the CWPM read, he or she was given a corresponding AIMSweb comprehension maze passage. These probes were selected randomly.

**Maintenance.** One week after intervention ended, the researcher asked each participant to complete one novel CR passage and one novel AIMSweb passage on the computer. This phase was identical to baseline with the exception that only one story each was read. As in baseline, errors made, CWPM, and items scored correct on the maze passage were graphed on the computer.

**Pupil independence.** The computer software was designed to systematically guide students through the instructional sequence with little or no adult continuous monitoring. The researcher used a check sheet to record the number of prompts students needed over the course of the intervention. To aid independence, the researcher periodically would say to the student, “What should you do next?” Fewer number of prompts indicated greater pupil independence.

**Behavior management.** A participant’s behavior was reinforced in many different ways. First, the computer displayed the results of the TWR and CWPM, which were used as a visual reinforcer. Second, the participants saw their charts where they plotted their progress on both treatment and generalization probes across sessions, allowing them to see their progress. Each student had a research plan that showed the order in which the stories would be read. After successful completion of each story, students selected a sticker (e.g., Disney Frozen characters) to put over the story title to indicate completion.

**Social Validity**

Participants were interviewed using a questionnaire after the study ended to assess their reactions to RR and their perceived progress.

**Treatment Integrity**

The purpose of collecting treatment integrity during these studies was three-fold. The performances observed were of the primary experimenter (51.1% by second observer), the participants (100% by researcher, 51.1% by second observer), and the computer (100% researcher, 51.1% second observer). In all instances, checklists were used to calculate the percentage of steps completed correctly.

**Results**

This study examined the effects of a multicomponent, supplemental intervention (i.e., RRI with CR passages) engineered into a novel computer software program, RR. A multiple baseline probe across participants demonstrated experimental control and instructional efficiency, essential in helping students with risk make gains needed for academic competence. The researchers questioned the beneficial effects of this intervention on the reading fluency of cold-reads for CR and AIMSweb passages for first-grade ELLs. Figure 1 shows that all three students made consistent progress in their fluency of CR and AIMSweb passages. A functional relation, indicating the reading behavior was under the control of the intervention, is evident in that increases are not noted until the treatment is introduced. Also noteworthy
is the dotted horizontal line depicting the mean fluency for each student with the CR practice passages. Tau-U was used to calculate effect size for each participant. Tau-U is an estimate of the percentage of data that shows improvement between phases (Rakap, 2015). Table 3 shows that each participant demonstrated medium to large effect sizes for reading fluency and comprehension with the exception of Arya (small) on AIMSweb passages. Individual participant results are discussed below.

### Table 3. Tau-U Effect Size for Reading Fluency and Comprehension

<table>
<thead>
<tr>
<th>Student</th>
<th>Tau-U CR Passages</th>
<th>Tau-U CR Mazes</th>
<th>Tau-U AIMSweb Passages</th>
<th>Tau-U AIMSweb Mazes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isabell</td>
<td>.94</td>
<td>.99</td>
<td>.98</td>
<td>.75</td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>Large</td>
<td>Large</td>
<td>Medium</td>
</tr>
<tr>
<td>Arya</td>
<td>.75</td>
<td>.95</td>
<td>.57</td>
<td>.67</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>Large</td>
<td>Small</td>
<td>Medium</td>
</tr>
<tr>
<td>Nico</td>
<td>.77</td>
<td>.96</td>
<td>.68</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>Large</td>
<td>Medium</td>
<td>Large</td>
</tr>
<tr>
<td>Weighted Average</td>
<td>.82</td>
<td>.97</td>
<td>.74</td>
<td>.81</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>Large</td>
<td>Medium</td>
<td>Medium</td>
</tr>
</tbody>
</table>

**Note.** Small Effect = .65 and lower; Medium Effect = .66–.99; Large Effect = .93 and higher.
Figure 1. CWPM across participants.

Note. The horizontal dashed line represents an average of the participant’s scores on practiced passages during the Timed Reading step of intervention.

Nico

During baseline Nico obtained a mean fluency score of 9 (range: 6-10) for CR passages and increased his fluency mean to 18.1 CWPM (range: 10–30) in intervention, which was 101% over baseline. Similarly, on the AIMSweb generalization passages, Nico averaged 7 (range: 4 – 9) CWPM during baseline but increased by 148% to a mean of 17.33 (range: 6–25) during intervention (see Figure 1). Nico
showed comparable progress on his maze comprehension scores, moving from a baseline mean of 1.0 (range 0-2) to an intervention mean of 7.6 (range: 2-11) for CR passages. Likewise, on the generalization AIMSweb passages Nico improved 333% from a baseline mean of 1 (range: 0-2) to an intervention mean of 4.3 (range: 3-6).

**Arya**

During baseline Arya obtained a mean fluency score of 15 (range: 7-25) for CR passages and increased her fluency mean to 24.8 CWPM (range: 10–50) in intervention, which was 65% over baseline. Similarly, on the AIMSweb generalization passages, Arya averaged 18.7 (range: 14 – 27) CWPM during baseline but increased by 35% to a mean of 25.2 (range: 16-38) during intervention (see Figure 1). Arya showed comparable progress on her maze comprehension scores, moving from a baseline mean of 3.75 (range 1-6) to an intervention mean of 8.8 (range: 6-12) for CR passages. Likewise, on the generalization AIMSweb passages Arya improved 127% from a baseline mean of 2.7 (range: 2-4) to an intervention mean of 6.17 (range: 5-9).

**Isabelle**

During baseline Isabelle obtained a mean fluency score of 21.7 (range: 9-26) for CR passages and increased her fluency mean to 41.6 CWPM (range: 28–58) in intervention, which was 92% over baseline. Similarly, on the AIMSweb generalization passages, Isabelle averaged 25.2 (range: 13– 43) CWPM during baseline but increased by 71% to a mean of 43.4 (range: 35-51) during intervention (see Figure 1). Isabelle showed comparable progress on her maze comprehension scores, moving from a baseline mean of 2.8 (range 2-4) to an intervention mean of 12.78 (range: 10-15) for CR passages. Likewise, on the generalization AIMSweb passages Isabelle improved 157% from a baseline mean of 4.8 (range: 1-8) to an intervention mean of 12.4 (range: 9-16).

For maintenance the researchers questioned if levels of reading fluency and comprehension achieved during intervention would persist. Figure 1 shows that one week after intervention all three students obtained CWPM scores on both CR and AIMSweb passages comparable to those obtained by the end of intervention. Figure 2 shows CR and AIMSweb comprehension maintenance scores slightly below intervention means but well above baseline levels for all participants with the exception of Arya (slightly above baseline). Table 4 present specific maintenance means for CR and AIMSweb mazes.
### Table 4. Baseline and Post Intervention Scores

<table>
<thead>
<tr>
<th>Participant</th>
<th>Mean AIMSweb CWPM</th>
<th>Mean CR CWPM</th>
<th>Mean AIMSweb Maze</th>
<th>Mean CR Maze</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
<td>Intervention (% growth)</td>
<td>Baseline</td>
<td>Intervention (% growth)</td>
</tr>
<tr>
<td>Nico</td>
<td>7.0</td>
<td>17.3 (148%)</td>
<td>9.0</td>
<td>18.1 (101%)</td>
</tr>
<tr>
<td>Arya</td>
<td>18.7</td>
<td>25.2 (35%)</td>
<td>15.0</td>
<td>24.8 (65%)</td>
</tr>
<tr>
<td>Isabelle</td>
<td>25.4</td>
<td>43.4 (71%)</td>
<td>21.7</td>
<td>41.6 (92%)</td>
</tr>
</tbody>
</table>

**Note.**
Baseline and post intervention scores by mean on CR and AIMSweb passages for CWPM and correct responses on maze passages, and percentage growth for both measures.
A final question on pupil independence was addressed by assessing the number of adult prompts the students needed to complete the instructional sequences. That is, to what extent were the students able to complete the sessions with minimal or no supervision? Figure 3 shows the researcher prompting data. Although all students required fewer prompts by the end of the intervention, these data show that the students performed at different levels of independence. Nico required many more prompts throughout the intervention, while Arya and Isabelle were much less dependent on adult supervision.
Reading Growth Measures

During the 1-week maintenance probes, the researcher administered the DIBELS end of the year (EOY) assessments. As can be seen in Table 2, two of the three students improved their DIBELS status. Arya went from Well Below Benchmark to Below Benchmark and Isabelle went from Below Benchmark to At or Above Benchmark. Although Nico doubled his DORF score, his status remained Well Below Benchmark. See Table 2.
Interobserver Agreement

Each participant was observed by a second observer for 50% of each subtest during pre- and post- tests, 58.5% of baseline, 51% of intervention sessions, 46.7% of treatment probes, and 100% of generalization probes for agreements and ranges of 95.3% (86.6% –100%), 97.6% (86.8%–100%), 98.3% (81%–100%), and 99.0% (97.1%–100), respectively.

Treatment Integrity

The primary experimenter used a checklist to assess treatment integrity of participants with computer software across 100% of the sessions. A second observer used the same checklist over 51.1% of the sessions. Agreements revealed that participants followed 94.8% of the directions given by the computer (individual mean scores ranged from 92%–98%). The mean agreement between primary and secondary experimenters was 99% (range of 98%–100%).

Social Validity

Social validity data was formally collected via survey at the conclusion of the study. In general, all the participants said they enjoyed reading on the computer, that they thought their reading had improved, they liked seeing their progress on the computer-generated graph, and two indicated wanting to continue the following year. When prompted why, one student answered the “stories are interesting,” and another stated it “is fun.”

Discussion

The findings of this study support other research in this area showing RRI delivered through computer software effective in increasing the reading fluency and comprehension of young learners in urban settings (e.g., Author, 2014; 2017 a). The data show that all participants quickly responded to the intervention and made fluency gains on both the CR passages and the AIMSweb generalization passages by the end of the intervention. Although the fluency rates across participants varied, mean fluency levels depicted on the broken horizontal lines in Figure 1 underscore the considerable amount of pupil practice needed for improvements on unpracticed passages to occur. This is consistent with the research literature on RRI in general (e.g., Therrien, 2004) and supports the position that RRI interventions can be effective with ELLs.

Individual differences are instructive. Nico, for example, averaged approximately 50 CWPM during his practice sessions to achieve 20 CWPM for his cold reads at the end of the study. One of the lowest performing students in his grade, and although not formally diagnosed, considering his grade retention and performance pattern, some cognitive impairment was suspected. Nevertheless, he more than doubled his fluency scores during his 11 weeks of intervention and, given this was his second year of first grade and considering the pattern of his baseline scores, it is seriously questioned if he would have made this level of progress without the intensive, systematic instruction provided by this intervention.

Arya and Isabelle, who were younger than Nico but were age appropriate for their grade and who scored better on the OTELA English assessment, made considerably more fluency gains with fewer sessions (10 and 7 weeks, respectively).
Both girls improved well over baseline levels and it is particularly noteworthy that Isabelle went from 21.7 CR CWPM in baseline to 50 CR CWPM within 7 weeks. Furthermore, she met benchmark for the end of year assessment.

**CR Stories and Generalization**

A special feature of this intervention was the use of CR stories for the repeated reading practice exercises. The rationale for developing a set of stories that reflected the students’ background and interest was based on some professional literature that indicated students responded more favorably to literature that reflected their backgrounds (e.g., Ebe, 2010) and their particular interests (Author et al, 2015 b). The assumption was that students would be more willing to read and re-read stories they found attractive and had common interest, not that they would necessarily perform better on the CR stories compared to the NCR stories. The researchers questioned if practice with CR stories would aid students in becoming more fluent in reading NCR passages (i.e., AIMSweb passages that were less directly linked to their backgrounds and interests).

For all three students, progress made on the CR practice passages carried over not only to the cold reads of the CR passages, but also to the cold reads of the NCR passages. The graphs show no discernible differences in these two types of cold reads for any of the students, leading the researchers to conclude that RRI with CR material can be effective in aiding reading fluency with NCR or typical classroom literature. Although these findings do not tell us if RRI with NCR practice materials would produce the same effects, a previous investigation does give a slight edge for fluency with CR compared to NCR materials (Author et al., 2015 a).

In the current study, maze passages, which are considered to be especially effective with primary-aged students (Parker, Hasbrouck, & Tindal, 1992) were used to assess comprehension. These data as displayed in the multiple-baseline graphs are quite convincing, showing pronounced increases in maze completions immediately upon introducing the intervention. For the most part, the upward trajectory for all participants is similar, but the progress of Isabelle is particularly dramatic and noteworthy. During intervention Isabelle typically finished each of the mazes before the timer and probably would have achieved higher scores if more opportunities were available.

All students made similar gains in comprehension on the generalization (AIMSweb) passages as they did on the CR passages. Again, this was especially true of Isabelle. Nico and Arya, who did show growth over baseline with the generalization passages, performed at a slightly lower level compared to the CR mazes. It may be that the assumed greater relevance of the CR passages facilitated comprehension, strengthening the rationale for using CR passages in this intervention. The high fluency rates of the practiced passages during intervention, averaged 50, 60, and 70 CWPM for Nico, Arya, and Isabelle, respectively, apparently aided both fluency in connected text and comprehension. Given that comprehension is the essence of reading these findings are encouraging and extend the existing research (e.g., Fuchs et al., 2006), showing the positive direct relationship between fluency and comprehension.

The maintenance data for these students must be interpreted with some caution. First, because of time limits due to the end of school, the researchers only
had one week between intervention and maintenance assessments. Further, although there was an ascending trend for all three students during intervention, the variability of these data suggest a need for several data points following intervention to get a clearer picture of the students’ performance and stability. Nevertheless, the stability of the fluency scores for all three participants was maintained at intervention levels. This is not the case for the maze maintenance, where although above baseline levels, there were noticeable drops for all three students. Again, it is difficult to draw conclusions from one data point. These assessments occurred at the end of the school year on a day when several other tests were given. It is possible that test fatigue was a factor in the maintenance maze assessments, which occurred after the fluency measures.

One important question guiding this research is whether first-grade students could be relatively independent or self-directed in using this computer-based intervention with the desired effects. To assess this outcome the researchers counted the number of task-prompts (e.g., What should you do now?) to complete each instructional sequence. The number of prompts were systematically faded for all three students but Nico consistently showed the most difficulty with independence. Although the oldest of the three, Nico was the lowest performing with suggestions of intellectual impairment. Arya and Isabelle displayed much greater independence but they too evidenced physical/developmental problems that hampered their performance. Despite using the smallest sized mouse, Arya, for example, had difficulty holding and moving the mouse to make the computer responses. On occasion, the researcher had to hold her hand, which often shook, to help her make the response. Isabelle was very small for her age and had a very soft voice. She was constantly prompted to speak louder so that she could be recorded in the voice-activated computer. This was difficult for Isabelle, who said she enjoyed reading the stories, but did not like reading on the computer because she could not read loud enough. These difficulties point to the problems that may accompany children who have both immigrant and refugee status. Although specific records were unavailable or the researchers could not access, these children were born in Somalia where difficult birth conditions were highly likely and would subsequently impact formal schooling. The speaking/voice problems of Isabelle, for example, may be indicative of low birth weight and underdeveloped lung capacity, limiting her ability to project and be heard. All of the students continued to be supervised individually by the end of the study. Nevertheless, there was sufficient evidence (both prompting data and informal observations) that after an initial training period, first graders showed enough independence that they possibly could be supervised in small groups (e.g., four to five students per group) to achieve desired outcomes. The highly prescribed nature of this software suggests that the supervising personnel could be trained paraprofessionals as well as certified professionals. Further investigations are needed for this purpose.

**Limitations**

Although the data from this study support previous research (Author, 2017b), there is still need for further study to expand both the number of participants and diversity of backgrounds. Additionally, although each of the participants made significant improvement in fluency, the maintenance declines in the comprehension...
data suggest there may be a need to strengthen both fluency and decoding skills so that comprehension is more automatic. Eleven weeks of explicit fluency instruction is negligible compared to over 30 weeks for the typical school year. It is noteworthy, however, that comprehension maintenance is still above baseline levels. Even though two of the three participants became considerably more independent in using the software, individualized guidance continued until the end of the study. Much more research is needed to conclude pupil independence with this intervention and under which conditions.

**Conclusions and Implications for Classroom Applications**

The findings of this study strongly support the beneficial effects of RRI to increase the reading fluency of first-grade, ELLs who evidence reading/special education risk. These data also underscore the facilitating relation between fluency and comprehension. Along with validating existing RRI investigations, these data support the value of using CR passages that represent the children’s experiences/background for fluency practice.

Children who are English learners and beginning readers need to make meaningful connections with the text. This is particularly true for students with reading risk. School personnel need to seek out and develop materials that serve this purpose. For children from less common environments such as Somali where written materials for children are extremely sparse, it will be necessary to access a variety of resources such as the immediate community, the child, the child’s family as well as the Somalia history to construct engaging and affirming learning materials. Feedback and input from the children can be highly valuable, relating which stories they like and why they like them. As students progress in their literacy, they can be engaged in creating stories that reflect their interests and backgrounds. These and related activities can lead to English learning, literary enjoyment, and a wide range of reading beyond CR texts. Although all three children in this study made substantial progress, it is important to note that the intervention was supplementary, tailored to the children’s backgrounds, and individualized. The importance of school-based supplementary interventions cannot be overstated for young school children (especially low-income children in urban settings) who show risk.

**Implications for Future Research**

More research with other language different populations would greatly enhance this line of work. Researchers could systematically compare RRI intervention with CR materials versus NCR materials to compare fluency and comprehension gains as well as generalization. Also included in these studies would be students with definite diagnosis of disabilities compared to those who simply show risk without diagnoses. The limited resources typically encountered in urban and low-income settings might be addressed through the use of technology such as that used in this study. One question not fully answered is whether first graders, who are ELLs and have special needs, could use such technology and make good progress with minimal adult supervision. Additional research is needed to more fully address this issue.
REFERENCES


Read Naturally. www.readnaturally.com


Spache Readability Formula (nd). (http://www.readabilityformulas.com/spache-readability-formula.php


Yopp, H. K., & Yopp, R. H. (2009). Phonological awareness is child’s play!. *Young Children*, 64(1) 12–21.