

Innovative Approaches to Literacy

Research to Practice

Project Evaluation

Adrian Pasquarella

John Z. Strong

Sharon Walpole

University of Delaware

PR/Award # S215G150117

Research to Practice

This report reviews a two-year Innovative Approaches to Literacy project to refine a K-5 literacy curriculum centered on repeated reading and professional development (PD) in a high-needs, rural school district and to add a wide reading component. Teachers in the Local Education Agency (LEA) had already used a pilot version of the curriculum and professional development modules, but their use had not been studied and the curriculum has not been iteratively designed to maximize feasibility.

In Year 1, the research team added high-quality classroom libraries for all classrooms, and teachers tracked student use of the new texts. In addition, teacher fidelity data and response to online modules were collected and used to refine both the lesson plans and the PD modules. In Year 2, teachers continued to monitor students' free reading and also implemented and evaluated the revised curriculum and PD modules. Across both years, student achievement data were collected to establish the promise of the curriculum and its associated PD intervention to influence teacher practice, student wide reading, and student reading achievement.

The partner schools were involved in a state-level Striving Readers grant. In order to evaluate the project, researchers used comparison data from two comparison districts in that grant. These comparison schools were also using the curriculum, but their PD access was different, and they did not implement the wide reading component. They provided information about their instruction to the state as part of the grant requirements, so self-reported implementation data helped to contextualize their achievement data.

Background

The need for new curriculum and coherent PD services for teachers is especially significant, because full implementation of Common Core State Standards (CCSS) (Council of Chief State School Officers & National Governors Association Center, 2010) and evaluation through high-stakes testing were required in many states in the 2014-15 school year. Shifts in the CCSS for reading included increased difficulty levels, expanded requirements for information text reading in the primary grades, and a move from reader response to text-based response. Standard 10 of the CCSS targets reading complex text. The critical question in this project is not whether the goal is appropriate but rather how schools might feasibly accomplish it.

The CCSS recommended a staircase of text complexity that puts the burden of reading increasingly more complex texts in the elementary grades. Despite teachers' needs for tools to meet the standards, there were no commercial core reading materials that met these text complexity requirements available in 2013 (see, for example, <https://www.edreports.org/ela/reports/compare-k8.html>). State Education Agencies (SEAs) provide a wide variety of curriculum and professional development resources to help schools redesign their curricula, but these combinations have not been associated with gains in literacy achievement (Kane, Owens, Marinell, Thal, & Staiger, 2016).

This project engaged all teachers and their building and district leaders in one rural school district to address the shifts in the CCSS through a curriculum based on daily repeated readings of challenging narrative and informational text supplemented with wide reading from classroom libraries. At the start of the project, the researchers argued that the curriculum was supported by strong theory.

Strong Theory for the Curriculum Design

The curriculum employed in this project, *Bookworms*, is a set of comprehensive daily lesson plans for 135 minutes of English language arts instruction. It is modularly structured into three 45-minute segments: (a) Shared Reading, (b) Interactive Read-Aloud and Writing, and (c) Differentiated Instruction. The curriculum uses a set of researcher-selected trade books chosen to meet the text complexity and genre distributions of the CCSS and employs only evidence-based instructional strategies. Researchers embraced the distinction between constrained foundational skills, which can be addressed rather quickly through targeted instruction, and unconstrained proficiencies, such as vocabulary and comprehension, which continue to develop over time (Paris, 2005). Shared Reading and Interactive Read Aloud address unconstrained skills. Differentiated Instruction addresses constrained skills until they are mastered, and then the focus shifts to unconstrained skills.

Shared Reading includes vocabulary building, instruction in and modeling of comprehension strategies, and text-based discussion, all informed by evidence. The IES publication *Improving Reading Comprehension in Kindergarten through 3rd Grade: A Practice Guide* (Shanahan et al., 2010) rates teaching children to use comprehension strategies as a practice with strong evidence and engaging them in text-based discussion as a practice with minimal evidence. However, the researchers make the claim for *strong theory* based on the provision of daily repeated reading using the partnering system designed for Peer-Assisted Learning Strategies (Fuchs & Fuchs, 1998). Two studies of repeated reading met the What Works Clearinghouse standards (Ellis & Graves, 1990; Wexler, Vaughn, Roberts, & Denton, 2010). Taken together, they yielded a small positive

effect on reading comprehension. Two studies of Peer-Assisted Learning Strategies (PALS) met the What Works Clearinghouse standards (McMaster, Fuchs, Fuchs, & Compton, 2005; Stein et al., 2008), and another met with reservations (Mathes & Babyak, 2001). These revealed small gains in reading comprehension. Shared Reading enables students to engage in paired repeated readings for 20 minutes every day, augmented with vocabulary instruction, comprehension strategy instruction, and text-based discussion.

Interactive Read Aloud and Writing combine several evidence-based practices, including interactive shared book reading, sentence combining, and process writing. The researchers made the claim for strong theory based on reviews of interactive shared book reading, which the What Works Clearinghouse found to have small effects on early reading achievement (U.S. Department of Education, 2015). Most importantly, these effects were found for students at risk, who comprise the majority of students in our LEA. Graham and Perin's (2007) *Writing Next* identifies sentence combining as an effective practice for improving writing quality, while Graham & Hebert's (2010) *Writing to Read* indicates that sentence combining also improves reading comprehension. In addition, the IES publication *Teaching Elementary School Students to be Effective Writers: A Practice Guide* (Graham et al., 2012) rates teaching students to use the writing process for a variety of purposes as a practice with strong evidence.

The final component of the curriculum is Differentiated Instruction. Researchers have designed differentiated instruction that involves more than just a different level of text, especially during the period of literacy acquisition (e.g., Al Otaiba, Connor, Folsom, Greulich, Meadows, & Li, 2011; Connor, Morrison, & Katch, 2004; Juel & Minden-Cupp, 2000). The researchers developed a protocol to group students by achievement and

to provide basic phonological awareness and alphabet knowledge for students with word recognition deficits; additional repeated readings for students with fluency deficits; and additional instruction in comprehension strategies, text structure, and text-based discussion for students with strong fluency. Each of these practices is based on strong theory. The What Works Clearinghouse rates phonological awareness plus letter knowledge training as having positive effects on print knowledge and potentially positive effects on phonological processing and early reading and writing (U.S. Department of Education, 2006). Repeated readings, described above, have small effects on reading comprehension. Finally, described above, instruction in comprehension strategies and text-based discussion have strong evidence and minimal evidence, respectively, for improving reading comprehension, while teaching students to identify and use knowledge of text structure has moderate evidence (Shanahan et al., 2010).

Since the time of this study, researchers have moved beyond *strong theory* to *promise* for the curriculum. Overall, students who received *Bookworms* outperformed a comparison group that received traditional guided reading instruction in both reading fluency and comprehension. On average, *Bookworms* produced medium effects on reading fluency and comprehension in upper elementary grades. For a full description see Walpole, McKenna, Amendum, Pasquarella, & Strong (2017).

Strong Theory for the PD Design

The project also employed PD design with strong theory. The combination of site-based coaching and virtual coaching from the university partners and access to both site-based PD and web-based PD ensured that all teachers had access. The overall design used recommendations from Desimone's (2009) features of PD. The approach was schoolwide

to ensure positive contextual support from the administration. It was content focused to emphasize teacher understanding of core instructional principles in reading: Shared Reading with comprehension strategy instruction and repeated readings, Interactive Read Aloud linked with sentence combining and process writing, and Differentiated Instruction in small groups. Active learning was targeted in video cases in online modules and in iterative collection of feedback through coaching sessions, observation and feedback with fidelity checklists, and surveys. The researchers planned for coherence through extensive connection to standards and through ongoing work in each school's response to intervention system.

Project Services

Project services included design and delivery of classroom libraries. The titles were selected to ensure that all children would have access to a wide range of fiction and nonfiction texts in a wide range of difficulty. The full list of titles included in classroom libraries is available from the authors. The texts were placed in classrooms on carts with checkout folders. These texts were provided in addition to the texts available in the library media center at each school. Teachers provided data on the number of books checked out in each classroom in each year.

Researchers offered optional online office hours three times per month in year 1 and year 2 of the project. Researchers visited each of the three partner schools for 6 days in year 1 and 8 days in Year 2. These visits comprised observations and feedback with building level coaches, grade-level team meetings to resolve feasibility issues, and modeling in classrooms. Based on the feedback from teachers and coaches in Year 1, researchers designed and delivered a large group PD presentation on writing

development, instruction, and assessment in summer of Year 1. Based on feedback from teachers and coaches in Year 2, researchers designed and delivered a large group PD presentation on Differentiated Instruction in summer of Year 2.

Research Questions

This evaluation is organized in two parts. The first part will address feasibility and degree of implementation questions for both the curriculum and the PD design.

1. Based on feedback from teachers and coaches in partner schools, what changes in the curriculum could make it more feasible?
2. To what extent did teachers in partner schools implement the curriculum as intended? Were there differences by grade level or by school?
3. How do teachers in partner schools rate the importance of online modules and face-to-face PD opportunities for building knowledge and influencing instruction? What components of online modules do teachers report as most useful? Were there differences by grade level or by school?
4. To what extent did students in partner schools check out books from the classroom lending libraries? Were there differences by grade level or by school?

The second part will address achievement gains for children during the two-year implementation timeline, first for the partner district, and then compared to two other districts utilizing less intense forms of the curriculum and professional development.

5. To what extent were the combination of curriculum, online and face-to-face PD opportunities, and classroom libraries in partner schools associated with changes in student achievement in reading fluency, comprehension, and a state assessment in English language arts? Were there differences by grade level or by school?

6. How did degree of implementation in partner schools compare to implementation in: (a) a comparison district using the same curriculum and online PD with half the number of face-to-face PD opportunities and (b) a comparison district using the same curriculum and online PD without face-to-face PD opportunities?
7. How did student achievement in partner schools compare to achievement in: (a) a comparison district using the same curriculum and online PD, half the number of face-to-face PD opportunities, and without the addition of classroom libraries, and (b) a comparison district using the same curriculum and online PD without the addition of face-to-face PD opportunities and classroom libraries?

Method

Participants

Two comparison districts were identified based on the Georgia College and Career Ready Performance Index (CCRPI) score assigned to each district in the year before the project started. CCRPI is a school accountability score that takes into account a district's content mastery (English language arts [ELA], math, science, and social studies achievement), progress (ELA and math growth), closing of achievement gaps, and readiness (literacy, student attendance, and fine arts or world language performance). The partner district had a CCRPI score of 65.18. The CCRPI in Comparison District A was 66.29. In Comparison District B, the CCRPI was 67.22.

The partner district included all students ($N = 1,038$) in grades K-5 in three elementary schools. Comparison District A included all students ($N = 1,593$) in one elementary school serving grades K-2 and another school serving grades 3-5. Comparison District B included all students ($N = 1,742$) in grades K-5 in three elementary schools.

Across all three districts and in all schools, 100% of the children were identified as eligible for federal lunch subsidies. In the partner district, 71.9% of students were African American, 19.8% of students were Caucasian, 5.8% of students were Hispanic, 1.3% of students were mixed race, and 1.2% of students were Indian or Pacific Islander. In addition, 14.2% of students received special education services, and 1.9% were English learners. In Comparison District A, 56.9% of students were Caucasian, 33.3% of students were African American, 5.1% of students were mixed race, 4.1% of students were Hispanic, and 0.6% of students were Indian or Pacific Islander. In addition, 15.1% of students received special education services, and 1.4% were English learners. In Comparison District B, 55.6% of students were African American, 31.2% of students were Hispanic, 9.1% of students were Caucasian, 3.4% of students were mixed race, and 0.7% of students were Indian or Pacific Islander. In addition, 14.6% of students received special education services, and 21.6% were English learners.

Measures

Three literacy achievement measures were available for all schools. The Oral Reading Fluency subtest of DIBELS Next (Good et al., 2011) is a measure of fluency. Participants read three passages aloud, scored for words correct per minute (WCPM). The median score is reported. The subtest is administered in fall, winter, and spring.

The HMH Reading Inventory (formerly the Scholastic Reading Inventory; Scholastic, Inc., 2009) is an untimed, computer adaptive measure of comprehension. Participants read passages on a computer screen and answer inferential comprehension questions in the form of a maze. Scores are reported in Lexile scores. Lexile scores are

continuous variables ranging from 200 to 1200. Participants took the HMH Reading Inventory in fall, winter, and spring.

The Georgia Milestones is a summative assessment of the English Language Arts (reading, writing, and language) Georgia Standards of Excellence. The test is administered online in three separate timed segments for 60 to 90 minutes. There are selected response and constructed response items. Scores are reported as scale scores and in four proficiency levels: Beginner, Developing, Proficient, and Distinguished.

Data also included two teacher surveys. The first was administered only to the teachers in the three partner schools. The purpose of the survey was to understand teachers' reports of the influence of face-to-face and online PD on knowledge and practice. Participants worked in pairs to complete a series of online PD modules. When they finished each module, they completed confidential online Likert-type ratings of the affordances of each module in general and of specific components of the module.

The second survey was administered confidentially online to teachers in all schools. The purpose of the survey was to describe school-based curriculum and PD choices to the state evaluation team. Teachers worked in grade-level teams to describe their typical whole group, small group, and intervention programs. They also reported the extent to which curricula and practices were used by all members of their grade-level team. In addition, teachers described the PD they had received.

Data Analysis & Results

This evaluation addressed the questions in order. Because of their broad scope, each question includes information about data collection, analysis and results.

Feasibility

Researchers collected ongoing implementation data during partner school visits. Feasibility challenges that could be addressed influenced changes in the curriculum itself and in the provision of face-to-face and on-line PD. Feasibility challenges that could not be addressed will inform future LEAs choosing to use this curriculum.

In the Shared Reading portion of the curriculum, teachers documented challenges for specific lesson plans that were too long for the time allotted. These lesson plans will need to be re-paced to enhance feasibility. In the Interactive Read Aloud portion of the curriculum, nonfiction magazines proved too difficult for students to see and presented teachers with text navigation problems. These texts will be replaced to enhance feasibility. Teachers reported concerns about writing instruction. The researchers responded with a specific training on writing development, instruction, and assessment. Finally, observations revealed inconsistent implementation of the Differentiated Instruction portion of the curriculum. Researchers provided additional face-to-face training on implementing the lessons and created online supports, including a series of YouTube videos (Walpole, 2016).

As part of the partner district's RTI protocols, computer adaptive interventions were added for all students at kindergarten and first grade, and as intensive interventions in upper elementary grades. These choices provided feasibility challenges that could not be resolved fully. In kindergarten and grade 1, time spent in the intensive interventions prevented children from addressing written work essential to the curriculum and from wide self-selected reading. Researchers investigated the potential for these interventions to improve basic skills by comparing diagnostic data generated with the computer

adaptive system with diagnostic data collected by the teachers. In kindergarten, these data yielded largely consistent recommendations. In first grade, they were not. The researchers recommended that the intervention not be used widely in first grade. This recommendation was followed in Year 2.

For upper elementary students, the partner district also adopted two different commercial intensive interventions. One required a 60-minute time slot and the other a 90-minute time slot. Scheduling these intervention services produced wide inconsistencies in access to the tier 1 curriculum in the three schools. The researchers recommended consistent practices, but they were not followed.

Degree of Implementation

Teachers in the partner schools responded to a team implementation survey in grade-level teams by school at the end of the year 2. Teams responded to two open-ended questions in which they described what happens typically during whole class and small group literacy instruction, including computer-based reading interventions that are used.

For whole class literacy instruction, teaching teams described two of the three parts of the Bookworms curriculum in detail. For shared reading (SR) lessons, teaching teams described: word study/vocabulary, focusing on decoding, defining, and spelling words; providing a first focus/purpose for reading; echo or choral reading of a novel; pausing to demonstrate a comprehension strategy using a think aloud; providing a second focus for reading; partner reading; discussing the text using comprehension questions; discussing the text structure using a graphic organizer; and written responses.

For interactive read aloud (IRA) lessons, teaching teams described discussing written responses from the previous day; reading aloud; modeling a comprehension

strategy; asking comprehension questions; participating in discussion about the text; reviewing and discussing tier 2 words; sentence composing activities, including imitating, expanding, and unscrambling; and completing written responses. In addition, teaching teams described providing genre-based writing instruction in narrative, opinion, and informative/explanatory writing using mentor texts, graphic organizers, modeling, and rubrics for evaluating writing samples.

For small group literacy instruction, teaching teams described placing students into three groups to either complete written responses, read self-selected books, or work with the teacher in a needs-based group identified through IDI and DORF data on differentiated instruction (DI) lessons. Teaching teams described the three different groups for DI lessons: (a) phonics lessons in basic alphabet knowledge, letter sounds, letter patterns with high-frequency word instruction, dictated sentences, blends and digraphs, r-controlled vowels, or vowel-consonant-e; (b) fluency lessons that include reading and re-reading a book with support from the teacher through choral or echo reading, partner or whisper reading, and discussion; and (c) vocabulary lessons that include students independently reading fiction or nonfiction silently and discussing the book with the teacher, focusing on text structure, comprehension, and vocabulary.

Teaching teams were also asked to indicate computer-based reading interventions that are used for students with reading difficulties. Kindergarten, first-, and second-grade teams indicated that they use *iRead* as a foundational reading program, third-grade teams indicated that they use *System 44* as a foundational reading program, and fourth- and fifth-grade teams indicated that they use *READ 180* as a reading intervention.

In the team implementation survey, teaching teams also indicated the frequency with which each of the three parts of the curriculum was implemented on an 8-point scale from never (0) to several times a day (7). Table 1 presents means and standard deviations for the degree of implementation of each of the three parts by school and by grade. Overall, teachers reported a high frequency of implementation for each of the three parts. On average, teachers used IRA lessons between several times a week and daily ($M = 5.89$), SR lessons between daily and several times a day ($M = 6.11$), and DI lessons between several times a week and daily ($M = 5.17$).

To determine whether there were differences by school or grade level, data were analyzed using one-way analyses of variance (ANOVAs). Results of the first analysis indicated a statistically significant difference between schools in how often DI lessons were implemented ($F = 4.310$, $df[2, 15]$, $p = .033$). Post hoc comparisons using the Tukey adjustment indicated that teachers at School 1 implemented DI lessons less frequently than other schools. However, this difference was only marginally significant ($p = 0.055$). On average, teachers at School 2 and School 3 reported implementing DI lessons daily ($M = 6.00$), whereas teachers at School 1 reported implementing them between several times a month and weekly ($M = 3.50$). There were no differences between schools in how often IRA and SR lessons were implemented.

Results of the second analysis indicated a statistically significant difference between grades in how often SR lessons were implemented ($F = 4.000$, $df[5, 12]$, $p = .023$). Post hoc comparisons using the Tukey adjustment indicated that fifth-grade teachers implemented SR lessons significantly more frequently than other grade levels ($p = 0.042$). On average, Kindergarten through fourth-grade teachers reported implementing

SR lessons daily ($M = 6.00$), whereas fifth-grade teachers reported implementing them between daily and several times a day ($M = 6.67$). There were no differences between grades in how often IRA and DI lessons were implemented.

Table 1

Degree of Implementation by School and by Grade

	N	Interactive Read Aloud		Shared Reading		Differentiated Instruction	
		M	SD	M	SD	M	SD
<u>School</u>							
School 1	6	6.17	0.41	6.17	0.41	3.50	2.95
School 2	6	5.83	0.41	6.00	0.00	6.00	0.00
School 3	6	5.67	1.37	6.17	0.41	6.00	0.00
<u>Grade</u>							
Kindergarten	3	5.67	0.58	6.00	0.00	6.00	0.00
First	3	5.00	1.73	6.00	0.00	6.00	0.00
Second	3	6.00	0.00	6.00	0.00	6.00	0.00
Third	3	6.00	0.00	6.00	0.00	5.00	1.73
Fourth	3	6.00	0.00	6.00	0.00	4.00	3.46
Fifth	3	6.67	0.58	6.67	0.58	4.00	3.46
Total	18	5.89	0.83	6.11	0.32	5.17	2.01

Note. Reporting is as follows: 0 = never, 1 = several times a year, 2 = monthly, 3 = several times a month, 4 = weekly, 5 = several times a week, 6 = daily, and 7 = several times a day.

Professional Development

Teachers in the partner schools responded to a professional development (PD) survey individually at the end of year 1. Teachers rated how important different types of PD are for improving their knowledge and instruction on a 3-point scale from not as important (0) to most important (2). Table 2 presents means and standard deviations for teachers' ratings of the impact of different types of PD by school.

Overall, teachers rated observation and feedback from their literacy coach as most important for improving both their knowledge ($M = 1.38$) and instruction ($M = 1.54$).

Observation and feedback from an outside consultant was rated the second most important type of PD for improving both knowledge ($M = 1.33$) and instruction ($M = 1.37$). Teachers rated traditional large group presentations as the least important type of PD for improving both knowledge ($M = 0.86$) and instruction ($M = 0.82$).

To determine whether there were differences in teachers' ratings of the importance of different types of PD by school, data were analyzed using a one-way ANOVA. Results indicated a statistically significant difference between schools for modeling in classroom from an outside consultant for improving knowledge ($F = 3.613$, $df[2, 114]$, $p = .03$) and instruction ($F = 3.069$, $df[2, 113]$, $p = .05$). Post hoc comparisons using the Tukey adjustment indicated that teachers in School 1 rated modeling in classroom from an outside consultant significantly higher for improving knowledge than teachers in School 2 ($p = .038$) and significantly higher for improving instruction than teachers in School 3 ($p = .048$). There were no other significant differences between schools for teachers' ratings of the importance of other types of PD.

Teachers also reported the most useful components of online modules on a 5-point scale from not at all useful (0) to very useful (4). Table 3 presents teachers' ratings of online module components by school. Overall, teachers rated video clips ($M = 3.54$) and audio clips ($M = 3.40$) as the most useful components. Handouts ($M = 2.91$) and activities to complete with colleagues after the module ($M = 2.98$) were rated least useful.

To determine whether there were significant differences in teachers' ratings by school, data were analyzed using a one-way ANOVA. Results indicated a statistically significant difference between schools for teachers' ratings of the usefulness of handouts ($F = 4.276$, $df[2, 172]$, $p = .015$), activities to complete in classroom after the module (F

= 3.140, $df[2, 175]$, $p = .046$), and activities to complete with colleagues after the module ($F = 7.283$, $df[2, 172]$, $p = .001$). Post hoc comparisons using the Tukey adjustment indicated that teachers in School 1 rated the usefulness of handouts significantly lower than teachers in School 3 ($p = .011$) but not significantly different from teachers in School 2 ($p = .073$). Upon closer inspection, although teachers in School 3 rated activities to complete in classroom after the module somewhat higher, the differences between teachers in School 1 ($p = 0.13$) and School 2 ($p = 0.07$) were not statistically significant. However, teachers in School 3 did rate activities to complete with colleagues after the module significantly higher than teachers in School 1 ($p = 0.001$) and School 2 ($p = 0.025$). There were no other significant differences between schools for teachers' ratings of the usefulness of any other components of online modules.

Table 2

<i>Ratings of Types of Professional Development on Knowledge and Instruction by School</i>						
Types of Professional Development	Knowledge			Instruction		
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>
<u>Modeling in classroom from outside consultant</u>						
School 1	26	1.50	0.76	26	1.62	0.64
School 2	46	0.98	0.93	45	1.42	0.78
School 3	45	1.33	0.83	45	1.16	0.85
Total	117	1.23	0.87	116	1.36	0.80
<u>Modeling in classroom from literacy coach</u>						
School 1	22	1.18	0.73	27	1.33	0.68
School 2	42	1.29	0.81	46	1.37	0.74
School 3	48	1.23	0.88	48	1.21	0.82
Total	112	1.24	0.82	121	1.30	0.76
<u>Observation and feedback from outside consultant</u>						
School 1	25	1.52	0.71	23	1.52	0.51
School 2	45	1.11	0.83	46	1.28	0.75
School 3	46	1.43	0.83	47	1.38	0.77
Total	116	1.33	0.82	116	1.37	0.72
<u>Observation and feedback from literacy coach</u>						
School 1	25	1.52	0.65	25	1.68	0.48
School 2	46	1.24	0.82	46	1.37	0.77
School 3	46	1.46	0.81	46	1.63	0.64
Total	117	1.38	0.79	117	1.54	0.68
<u>Grade-level team meetings with outside consultant</u>						
School 1	21	1.05	0.67	24	1.00	0.78
School 2	46	1.13	0.62	46	1.00	0.76
School 3	47	1.32	1.07	46	1.33	0.84
Total	114	1.19	0.84	116	1.13	0.81
<u>Grade-level team meetings with literacy coach</u>						
School 1	25	1.24	0.66	23	1.17	0.72
School 2	46	1.13	0.72	46	1.07	0.71
School 3	48	1.42	0.82	47	1.38	0.85
Total	119	1.27	0.76	116	1.22	0.78
<u>Traditional large group presentations</u>						
School 1	21	0.86	0.79	22	0.77	0.81
School 2	46	0.93	0.90	46	0.74	0.74
School 3	45	0.78	1.02	46	0.93	1.10
Total	112	0.86	0.93	114	0.82	0.91
<u>Online modules</u>						
School 1	22	1.14	0.83	23	0.96	0.82
School 2	47	1.34	0.76	48	1.38	0.82
School 3	45	1.27	0.86	45	1.22	0.70
Total	114	1.27	0.81	116	1.23	0.78

Notes. Ratings are follows: 0 = not as important, 1 = moderately important, 2 = most important.

Table 3

Ratings of the Usefulness of Different Online Module Components by School

Online Module Components	<i>N</i>	<i>M</i>	<i>SD</i>
<u>Written module text</u>			
School 1	36	3.28	0.78
School 2	66	3.20	0.79
School 3	79	3.32	0.79
Total	181	3.27	0.79
<u>Handouts</u>			
School 1	33	2.52	1.03
School 2	63	2.94	0.78
School 3	79	3.05	0.90
Total	175	2.91	0.91
<u>Video clips</u>			
School 1	35	3.46	0.92
School 2	67	3.57	0.84
School 3	81	3.56	0.72
Total	183	3.54	0.80
<u>Audio clips</u>			
School 1	34	3.29	0.97
School 2	65	3.34	0.87
School 3	81	3.49	0.78
Total	180	3.40	0.85
<u>Activities to complete during module</u>			
School 1	31	2.84	0.93
School 2	64	3.08	0.88
School 3	78	3.19	0.81
Total	173	3.09	0.86
<u>Activities to complete in classroom after module</u>			
School 1	36	3.00	0.79
School 2	66	3.02	0.90
School 3	76	3.33	0.81
Total	178	3.15	0.85
<u>Activities to complete with colleagues after module</u>			
School 1	33	2.61	0.86
School 2	65	2.86	0.90
School 3	77	3.25	0.85
Total	175	2.98	0.90

Note. Ratings are follows: 0 = not at all useful, 1 = somewhat useless, 2 = neutral, 3 = somewhat useful, 4 = very useful.

Classroom Libraries

Students in grades 1-5 checked out books from classroom lending libraries during years 1 and 2. The number of books checked out in each classroom was counted for both years. Table 4 presents means and standard deviations for the number of books checked out by school and grade for each year. Overall, students checked out 5,877 books in year 1 and 16,016 books in year 2. A paired samples *t*-test indicates that students checked out significantly more books from classroom libraries in year 2 than in year 1 ($t = 6.497$, $df = 39$, $p < .001$).

To determine whether there were differences in the number of books checked out by school, data were analyzed using one-way multivariate analyses of variance (MANOVA) with number of books checked out in year 1 and year 2 as dependent variables. Results of the first analysis indicated a statistically significant difference between schools for number of books checked out in year 1 ($F = 3.223$, $df [2, 47]$, $p = .049$). There was no significant difference between schools in year 2 ($F = 1.504$, $df [2, 48]$, $p = .232$). Post hoc comparisons using the Tukey adjustment indicated that students in School 3 checked out significantly more books in year 1 than students in School 1 ($p = .039$). Results of the second analysis indicated there were no significant differences between grade levels in the number of books checked out in year 1 or year 2.

Table 4

Number of Books Checked Out by School and Grade in Year 1 and Year 2

School	Grade	Year 1			Year 2				
		<i>N</i>	Sum	<i>M</i>	<i>SD</i>	<i>N</i>	Sum	<i>M</i>	<i>SD</i>
School 1	1	2	125	62.50	4.95	2	513	256.50	167.58
	2	2	151	75.50	10.61	2	519	259.50	71.42
	3	2	153	76.50	60.10	2	199	99.50	54.45
	4	2	143	71.50	50.21	2	567	283.50	169.00
	5	2	65	32.50	9.19	2	446	223.00	59.40
	Total	10	637	63.70	31.68	10	2244	224.40	111.03
School 2	1	4	407	101.75	12.34	4	1202	300.50	164.46
	2	4	605	151.25	42.95	4	1156	289.00	157.23
	3	4	496	124.00	43.60	4	1432	358.00	45.56
	4	4	407	101.75	14.45	4	646	161.50	21.39
	5	3	384	128.00	50.48	4	1647	411.75	234.10
	Total	19	2299	121.00	36.69	20	6083	304.15	156.88
School 3	1	4	427	106.75	46.41	4	2078	519.50	421.83
	2	5	615	123.00	95.45	4	2060	515.00	470.09
	3	4	747	186.75	79.94	4	1150	287.50	148.79
	4	4	814	203.50	219.92	5	1376	275.20	84.53
	5	4	338	84.50	48.52	4	1025	256.25	135.65
	Total	21	2941	140.05	113.42	21	7689	366.14	286.71

Table 5

Student counts and percentages of performance ranks on DORF and RI assessments from Fall 2015 – Spring 2017.

School	Performance Rank	Fall 2015		Spring 2016		Fall 2016		Spring 2017	
		Count	Percent	Count	Percent	Count	Percent	Count	Percent
<i>DORF: Reading Fluency</i>									
School 1	Well Below GL	27	23.5	22	18.5	14	15.9	8	9
	Below GL	15	13	24	20.2	6	6.8	5	5.6
	At or Above GL	73	63.5	73	61.3	68	77.3	76	85.4
	Total	115	100	119	100	88	100	89	100
School 2	Well Below GL	78	29	57	20.7	44	22.1	28	14
	Below GL	50	18.6	64	23.3	30	15.1	33	16.5
	At or Above GL	141	52.4	154	56	125	62.8	139	69.5
	Total	269	100	275	100	199	100	200	100
School 3	Well Below GL	122	37.7	102	30.1	91	40.1	48	21.1
	Below GL	58	17.9	98	28.9	43	18.9	43	18.9
	At or Above GL	144	44.4	139	41	93	41	136	59.9
	Total	324	100	339	100	227	100	227	100
<i>RI: Reading Comprehension</i>									
School 1	Below GL	62	71.3	41	47.1	56	64.4	37	41.6
	At or Above GL	25	28.7	46	52.9	31	35.6	52	58.4
	Total	87	100	87	100	87	100	89	100
School 2	Below GL	136	67.3	102	50.5	136	70.1	90	45.7
	At or Above GL	66	32.7	100	49.5	58	29.9	107	54.3
	Total	202	100	202	100	194	100	197	100
School 3	Below GL	176	71.3	131	53	150	69.8	106	48.6
	At or Above GL	71	28.7	116	47	65	30.2	112	51.4
	Total	247	100	247	100	215	100	218	100

Notes. GL = Grade Level.

Student Achievement

Table 5 displays performance rank counts and percentages by school from Fall 2015 to Spring 2017 for the DIBELS Oral Reading Fluency (DORF) and the Reading Inventory (RI) assessments. The top portion of the table displays the descriptive statistics for reading fluency in words correct per minute (WCPM), and the bottom displays descriptive statistics for reading comprehension in Lexiles.

For reading fluency, there was a large reduction in the number of students performing *well-below* and *below* grade level on the DORF assessment and a large increase in the number of students performing *at or above* grade level. School 1 saw the largest increase, with 22% more students performing at or above grade level, two years later. In Spring, only 9% and 5.6% of students were performing well-below or below grade level, respectively. Eighty-five percent of students were performing at or above grade level. School 2 and School 3 saw remarkable gains as well, with 17.1% and 15.5% more students, respectively, performing at or above grade level.

For reading comprehension, large gains are clearly evident. All three partner schools started out with over two-thirds of students performing below grade level on the RI assessment. Two years later, all partner schools had over 50% of students performing at or above grade level. School 1, again, stood out with the highest increase of students performing at or above grade level (30.3%), followed by School 3 (22.7%) and School 2 (21.6%).

Table 6.

Student counts and percentages of performance ranks on Milestones State Assessment in Spring 2016 and Spring 2017.

School	Performance Rank	Spring 2016		Spring 2017	
		Count	Percent	Count	Percent
School 1	Beginner	40	0.40	25	0.27
	Developing	39	0.39	52	0.55
	Proficient	17	0.17	15	0.16
	Distinguished	5	0.05	2	0.02
	Total	101	1.00	94	1.00
School 2	Beginner	114	0.53	83	0.39
	Developing	68	0.31	86	0.40
	Proficient	29	0.13	39	0.18
	Distinguished	5	0.02	6	0.03
	Total	216	1.00	214	1.00
School 3	Beginner	132	0.52	114	0.44
	Developing	80	0.31	101	0.39
	Proficient	35	0.14	34	0.13
	Distinguished	7	0.03	9	0.03
	Total	254	1.00	258	1.00

Table 6 displays descriptive statistics for the Milestones State Assessment in English language arts. This assessment was first introduced in Spring 2015. Georgia Department of Education considers students scoring within the *developing* range to be performing at grade level. Overall, all three partner schools saw substantial increases in the number students performing within with *developing* range and large reductions in the number of students performing within the *beginner* range. *Proficient* and *distinguished* rankings remained relatively stable across the years, except for at School 2, which saw the largest increase of the number of students in the *proficient* range.

Table 7.

Mean scores and standard deviations for oral reading fluency in WCPM by school and grade from Fall 2015 to Spring 2017

School	Initial Grade	N	Fall 2015		Winter 2016		Spring 2016		Fall 2016		Winter 2017		Spring 2017	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
School 1	2	27	65.85	15.80	90.07	16.55	104.78	15.63	87.15	14.41	110.85	17.17	127.44	20.11
	3	25	99.00	22.32	119.00	19.93	135.72	25.00	121.68	29.11	146.80	23.70	161.96	26.68
	4	31	97.74	37.44	121.87	38.24	134.52	38.88	119.71	41.65	139.90	36.55	149.55	43.45
	Total	83	87.75	31.18	110.66	30.76	125.20	31.89	109.71	34.60	132.53	31.37	146.10	34.97
School 2	2	60	76.22	25.52	96.32	27.68	114.23	31.73	93.72	30.89	117.73	34.18	130.83	41.87
	3	74	77.84	33.47	102.97	36.01	119.27	38.91	97.61	37.27	116.23	38.41	135.64	37.38
	4	59	80.73	37.14	109.34	37.10	127.14	38.99	117.44	40.02	136.85	39.84	142.71	43.85
	Total	193	78.22	32.34	102.85	34.21	120.11	37.02	102.46	37.51	123.00	38.53	136.31	40.90
School 3	2	62	66.08	28.19	88.79	26.92	101.73	28.96	79.74	30.09	99.05	31.87	117.50	36.42
	3	66	70.67	32.73	90.09	29.85	106.02	30.20	88.26	31.66	102.95	31.53	128.15	28.34
	4	68	84.29	37.99	104.79	36.52	120.78	36.35	103.93	39.57	121.93	40.26	132.22	35.28
	Total	196	73.94	34.09	94.78	32.18	109.78	32.98	91.00	35.44	108.30	36.16	126.19	33.89

Figure 1.

Fluency growth in WCPM by school for students in grades 2-3 from 2015 to 2017

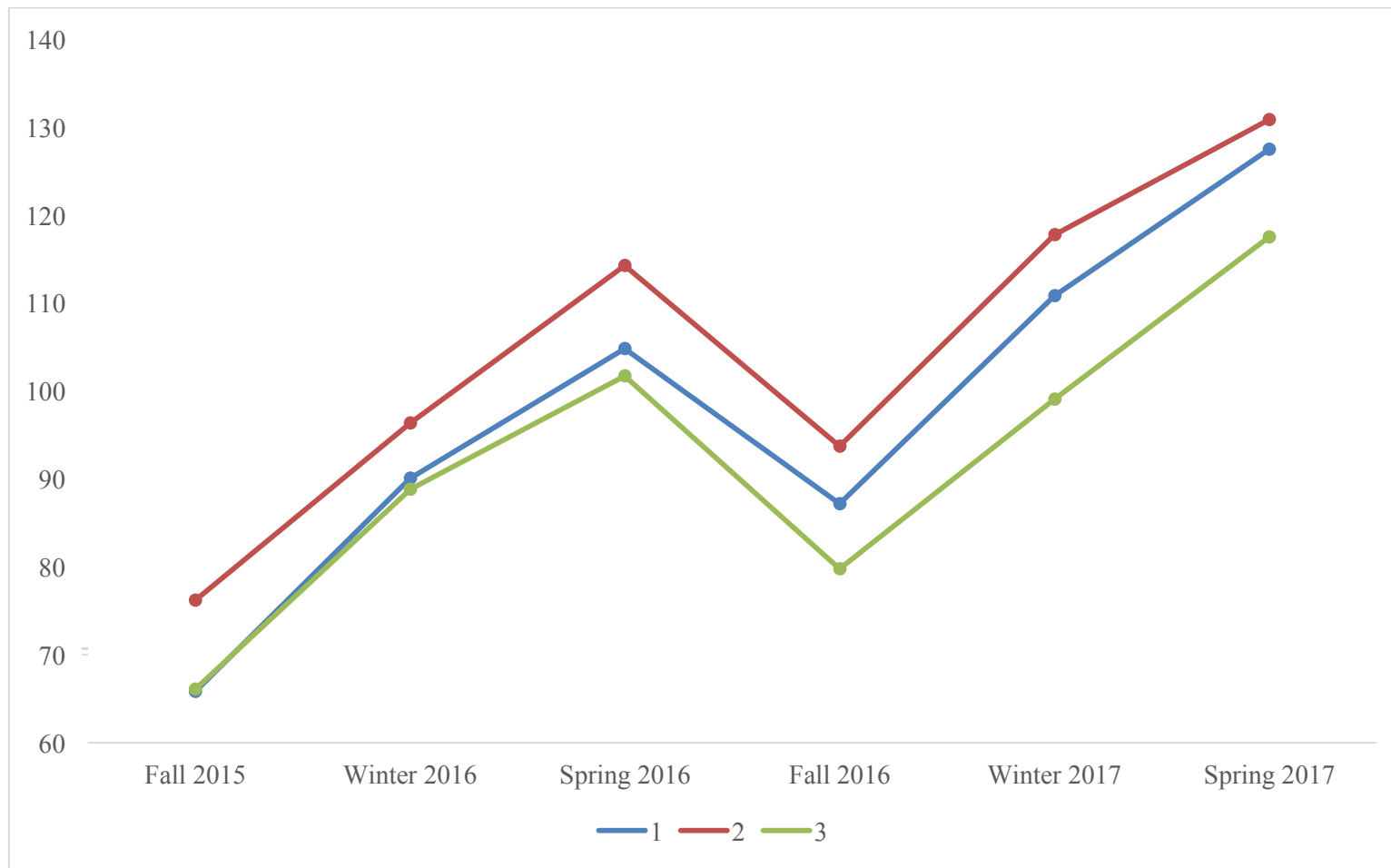


Figure 2.

Fluency growth in WCPM by school for students in grades 3-4 from 2015 to 2017

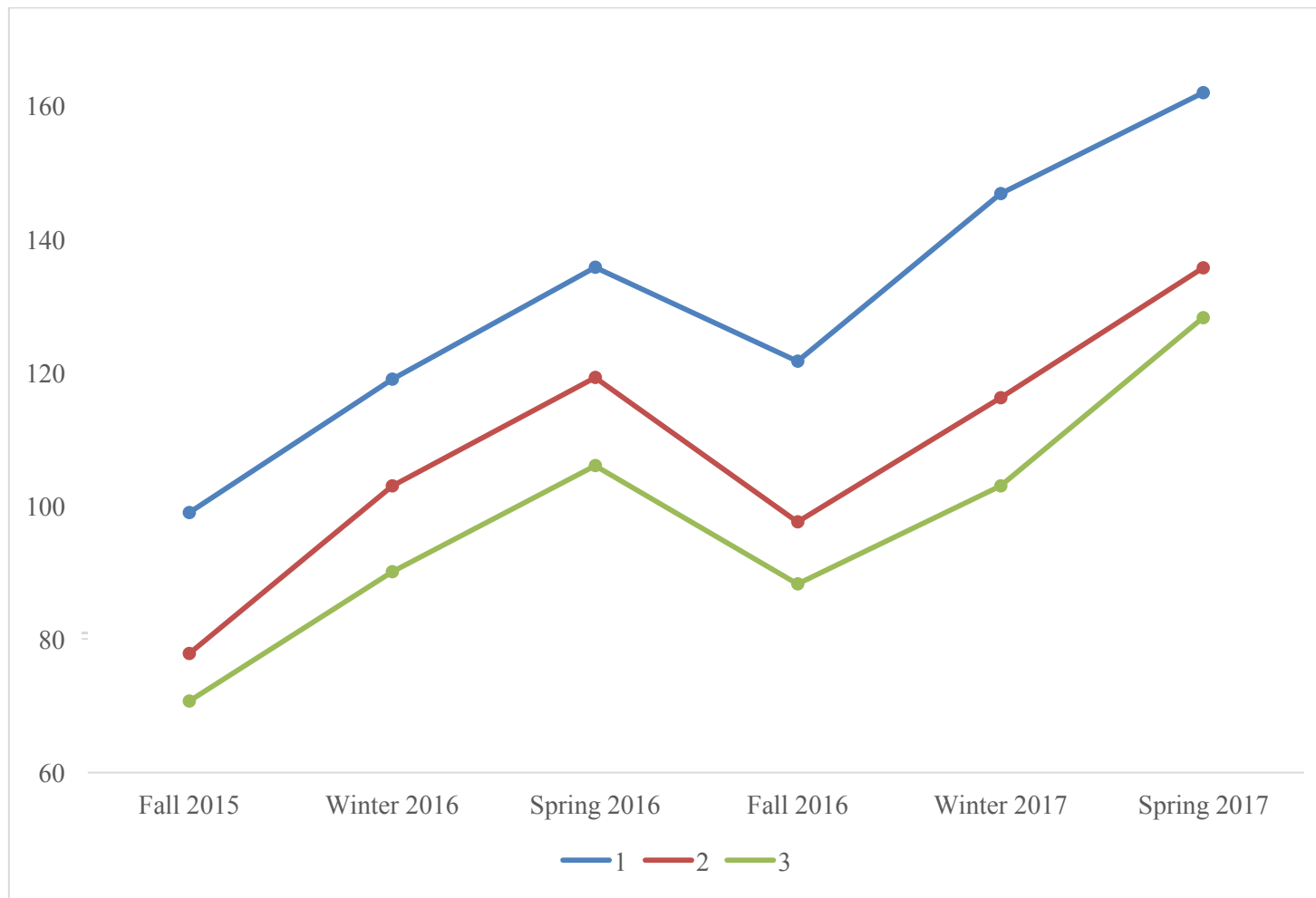
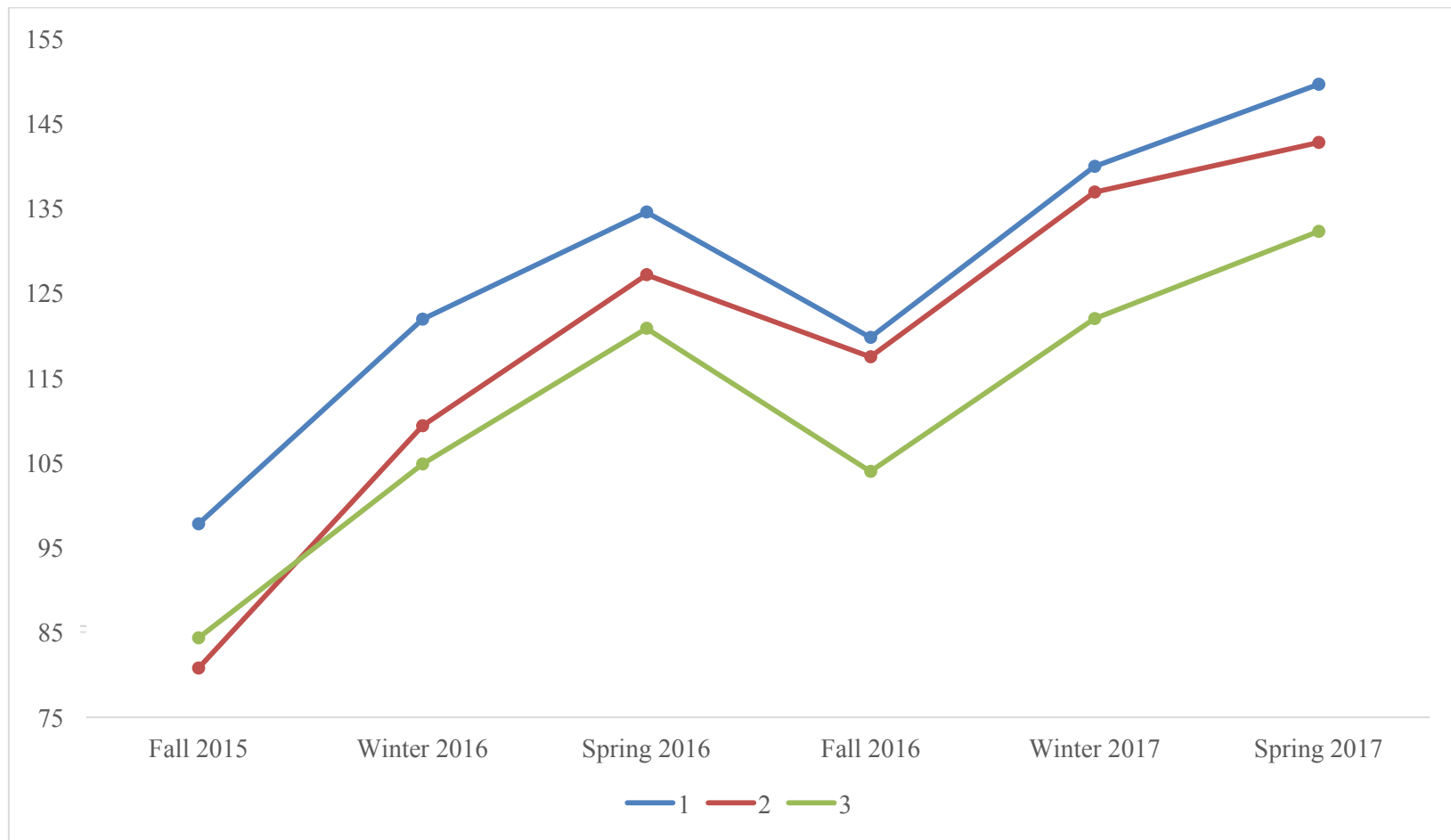


Figure 3.

Fluency growth in WCPM by school for students in grades 4-5 from 2015 to 2017



We examined school and grade level differences in the growth of oral reading fluency by conducting a 6 (time) x 3 (school) x 3 (grade) repeated measures ANOVA. There was a significant three-way interaction between time, school and grade ($F = 2.863$, $df[20, 1952]$, $p < .001$). Table 7 displays the descriptive statistics for the DORF assessment in fall, winter, and spring from 2015 to 2017, separated by school and grade. Figures 1-3 plot means for each school separately for students who progressed from grades 2-3, 3-4 and 4-5, respectively, between years 1 and 2 in the study. Post hoc analyses were conducted to better understand the three-way interaction. Students in grades 2-3 in School 1 and School 2 experienced equivalent growth ($p = .697$), which was significantly higher than students in School 3 ($p = .038$). In grades 3-4, students in School 1 experienced significantly more growth than students in School 2 ($p = .029$), and students in School 2 experienced significantly more growth than students in School 3 ($p = .011$). In grades 4-5, growth was equivalent between all schools (p range .139 to .767).

Table 8.

Mean scores and standard deviations for reading comprehension in Lexiles by school and grade from Fall 2015 to Spring 2017

	Initial Grade	N	Fall 2015		Winter 2016		Spring 2016		Fall 2016		Winter 2017		Spring 2017	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
School 1	2	27.00	304.11	136.25	372.04	133.57	402.93	130.78	460.04	159.71	521.85	158.52	584.00	168.97
	3	25.00	451.28	153.63	497.60	154.90	590.84	147.51	624.24	154.16	664.24	157.87	722.68	156.60
	4	31.00	562.55	222.16	611.71	242.19	696.81	235.55	732.87	223.58	771.29	224.28	808.42	230.48
	Total	83.00	444.96	206.45	499.37	210.36	569.29	217.93	611.40	215.65	657.90	211.26	709.59	211.20
School 2	2	60.00	356.75	210.87	414.02	210.91	476.32	192.47	488.10	187.59	546.48	177.15	619.13	180.01
	3	74.00	408.27	229.47	480.82	225.25	538.82	220.76	554.14	231.71	616.34	224.10	676.50	211.78
	4	59.00	498.42	241.01	571.98	231.10	637.34	223.87	678.05	223.29	730.24	223.12	795.20	228.49
	Total	193.00	419.81	233.29	487.92	230.20	549.51	221.68	571.49	228.18	629.44	221.62	694.95	218.63
School 3	2	62.00	322.76	204.59	417.79	178.18	445.58	189.16	462.15	200.32	507.87	191.98	545.48	206.45
	3	66.00	398.27	245.75	477.89	223.29	537.45	210.72	564.21	221.77	637.15	217.04	715.79	199.43
	4	68.00	536.38	186.31	600.69	181.19	665.78	169.17	702.18	180.88	764.54	182.29	813.79	179.18
	Total	196.00	422.30	230.29	501.48	209.00	552.91	209.79	579.79	223.37	640.45	222.74	695.92	223.15

Figure 4.

Comprehension growth in Lexiles by school for students in grades 2-3 from 2015 to 2017

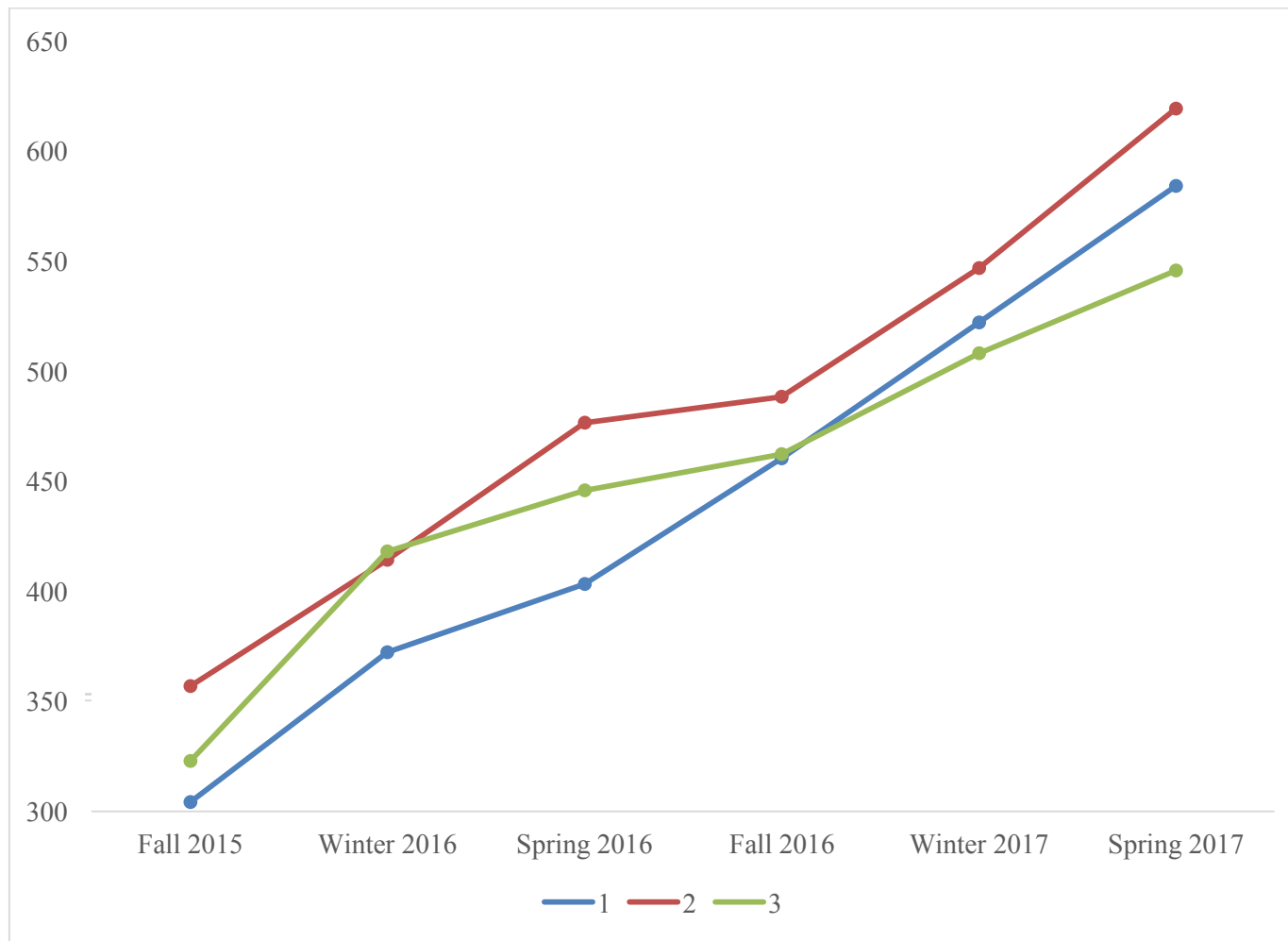


Figure 5.

Comprehension growth in Lexiles by school for students in grades 3-4 from 2015 to 2017

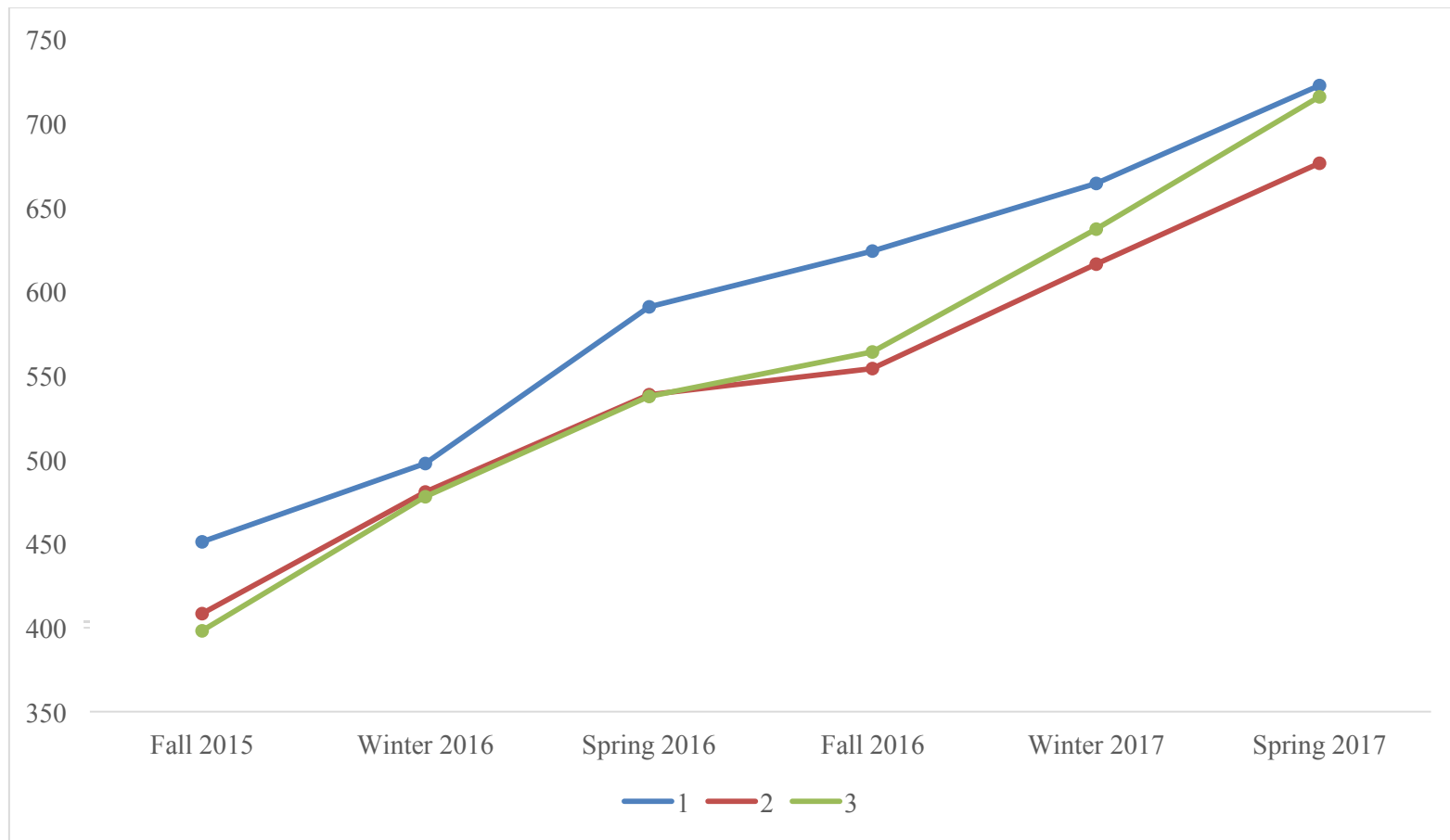
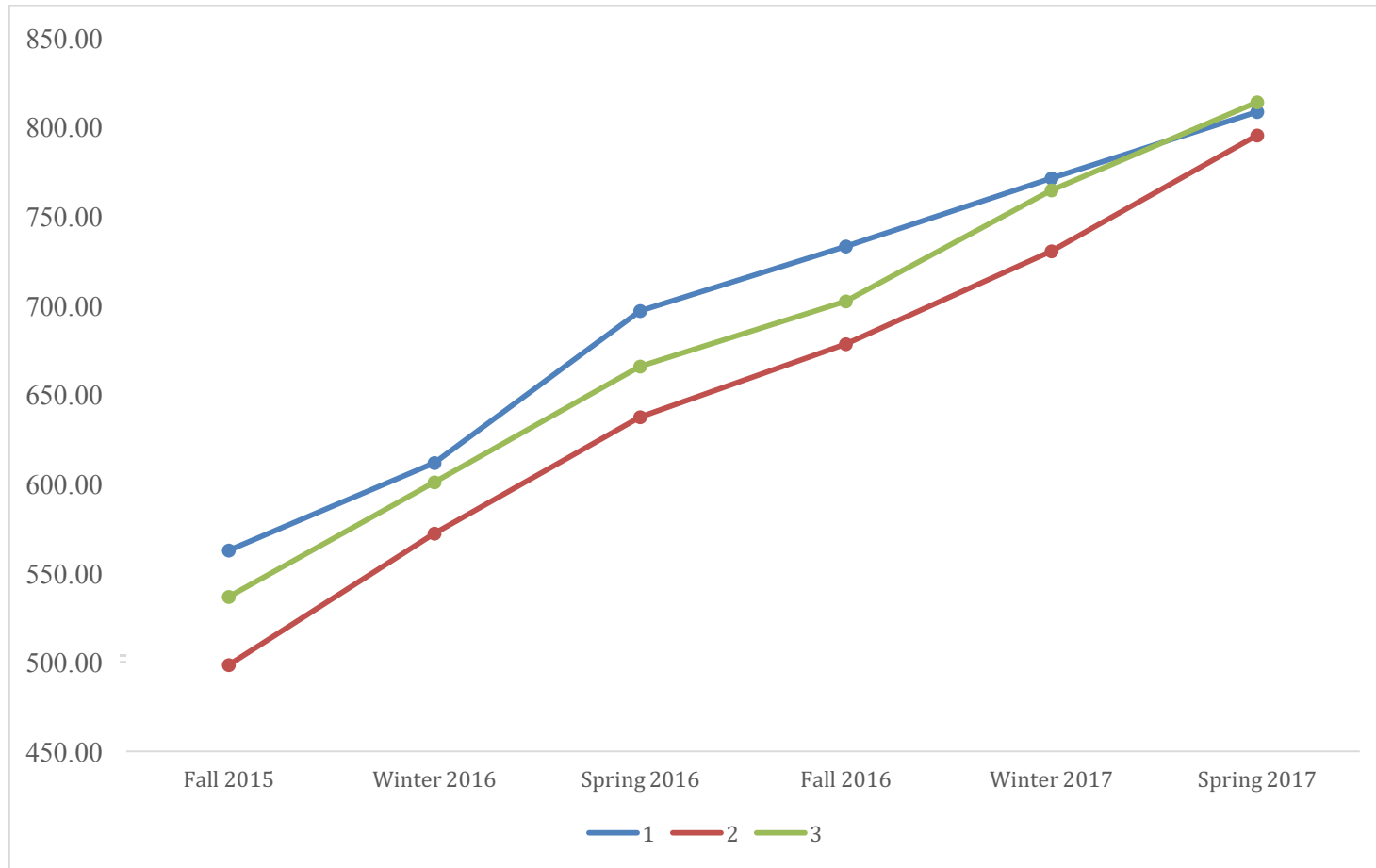


Figure 6.

Comprehension growth in Lexiles by school for students in grades 4-5 from 2015 to 2017



School and grade-level differences of growth in reading comprehension were examined with a 6 (time) x 3 (school) x 3 (grade) repeated measures ANOVA. There was a significant three-way interaction between time, school, and grade, ($F = 2.93$, $df [20, 1868]$ $p < .001$). Table 9 displays the descriptive statistics for the RI assessment in fall, winter, and spring from 2015 to 2017, separated by school and grade. Figures 4-6 plot means for each school separately for students who progressed from grades 2-3, 3-4 and 4-5, respectively, between years 1 and 2 in the study. Post hoc analyses were conducted to further understand the difference between schools and across grades. There was no significant difference in reading comprehension growth experienced by schools over the duration of the study. The significant three-way interaction was determined by slightly different patterns of non-linear change. In grades 2-3, School 1 and School 2 experienced similar growth, while School 3 displayed a slightly different trajectory. In grades 3-4 and 4-5, comprehension growth was not significantly different between the three schools.

Comparison Districts

The second part of the evaluation compared implementation of *Bookworms* in the three partner schools to the five schools in the two comparison districts. Both districts implemented the same curriculum but received different numbers of PD opportunities. The researchers visited each of the schools in Comparison District A to provide face-to-face PD opportunities for a total of 7 days over years 1 and 2. In comparison, the researchers visited the partner district to provide face-to-face PD opportunities for a total of 14 days in years 1 and 2. The researchers did not visit the schools in Comparison District B to provide face-to-face PD opportunities in year 1 or 2 of the partner. Teachers

in both comparison districts had access to the online PD modules, but neither comparison district had access to the additional classroom libraries used in the partner schools.

Teachers in the two comparison districts responded to the same team implementation survey in grade-level teams at the end of year 2. In describing their whole class and small group literacy instruction, teachers in Comparison District A described the three parts of the *Bookworms* curriculum: IRA, SR, and DI. While teachers in Comparison District B also described the three parts of the *Bookworms* curriculum in their whole class and small group literacy instruction, they also provided a more extensive list of activities not associated with the *Bookworms* curriculum.

For IRA lessons, teachers in Comparison District A mentioned reading aloud, comprehension questions and discussion, tier 2 vocabulary, sentence composing, and written responses. For SR lessons, teachers described word study or vocabulary; fluency practice through echo, choral reading, and partner reading; comprehension questions and discussion; journal or written responses; and super sentences. One team also mentioned providing writing instruction during whole class literacy instruction.

For small group literacy instruction, teachers in Comparison District A described placing students into groups for DI lessons based on assessment data and their instructional needs. For DI lessons, teachers mentioned direct instruction in phonics skills, leveled novels for additional comprehension and vocabulary practice, and fluency and comprehension lessons with sight word practice. They also mentioned that other students completed classwork, including word study practice, super sentences, and written responses.

Overall, teachers in Comparison District B described their whole class literacy instruction in the following way: activating strategy to discuss a focus or purpose for reading, word work or vocabulary instruction, interactive read alouds, think-alouds, comprehension strategies (e.g., text structure, main idea, and context clues), shared reading (with choral or echo reading then partner reading), story maps, sentence construction, discussion, constructed responses, writing (informational, opinion, and narrative), summarizing, grammar mini-lessons, independent reading, skills groups, and working on computers.

For small group literacy instruction, teachers in Comparison District B described providing instruction to students in different ability groups by reading level. Many small group rotations were provided, including phonics instruction, fluency practice with repeated reading, vocabulary and comprehension strategies with graphic organizers, guided reading with leveled books, handwriting practice, independent reading, grammar worksheets, computers, and writing.

Teaching teams were also asked to indicate computer-based reading interventions that are used for students with reading difficulties. Fourth- and fifth-grade teams in Comparison District A indicated that they use *Classworks* for reading and language arts practice. Similar to the partner schools, teaching teams in two schools in Comparison District B reported using *iRead* in kindergarten through third grade and both *System 44* and *READ 180* in grades four and five. In one school, teaching teams reported using *Fast ForWord* as a reading and language intervention program.

Comparison in Degree of Implementation

Grade-level teams in both comparison districts indicated the frequency with which each of the three parts of the curriculum was implemented on an 8-point scale from never (0) to several times a day (7). Table 9 presents means and standard deviations for the degree of implementation of each of the three parts by grade.

In Comparison District A, teachers reported a high frequency of implementation for each of the three parts. On average, teachers reported using IRA lessons between weekly and several times a week ($M = 4.67$), SR lessons between daily and several times a day ($M = 6.11$), and DI lessons between several times a week and daily ($M = 5.17$). In Comparison District B, teachers reported a lower frequency of implementation. On average, teachers reported using IRA lessons monthly ($M = 2.00$), SR lessons between several times a year and monthly ($M = 1.62$), and DI lessons between weekly and several times a week ($M = 4.41$).

To determine whether there were differences in degree of implementation between the partner and comparison districts, data were analyzed using a one-way ANOVA. Results indicated statistically significant differences between districts in frequency of implementation of IRA lessons ($F = 20.106$, $df[2, 38]$, $p < .001$) and SR lessons ($F = 37.394$, $df[2, 37]$, $p < .001$). There was no significant difference between districts in how often DI lessons were implemented ($F = 0.458$, $df[2, 41]$, $p = .636$). Post hoc comparisons using the Tukey adjustment indicated that teachers in Comparison District B implemented IRA lessons less frequently than teachers in the partner district ($p < .001$) and Comparison District A ($p < .001$). Teachers in Comparison District B also implemented SR lessons less frequently than teachers in the partner district ($p < .001$) and

Comparison District A ($p < .001$). There was no difference in how frequently teachers in Comparison District A and the partner district implemented both IRA and SR lessons.

Table 9

Degree of Implementation by Grade for Comparison Districts by Grade

District	Grade	N	Interactive Read Aloud		Shared Reading		Differentiated Instruction	
			M	SD	M	SD	M	SD
Comparison A	K	2	6.00	0.00	6.00	0.00	6.00	0.00
	1	2	4.50	0.71	6.00	0.00	6.00	0.00
	2	1	6.00	–	6.00	–	6.00	–
	3	2	6.00	0.00	6.00	0.00	5.00	1.73
	4	1	6.00	–	6.00	–	0.00	–
	5	1	7.00	–	7.00	–	0.00	–
	Total	9	4.67	0.83	6.11	0.33	4.67	2.65
Comparison B	K	2	0.00	0.00	0.00	0.00	3.00	4.24
	1	3	0.00	0.00	0.00	0.00	5.67	0.58
	2	3	2.33	4.04	2.33	4.04	4.00	3.46
	3	3	4.33	3.79	3.00	4.24	4.00	3.46
	4	3	1.00	1.73	1.00	1.73	3.67	3.22
	5	3	5.00	0.00	5.00	0.00	5.67	0.58
	Total	17	2.00	2.94	1.62	2.66	4.41	2.55

Note. Reporting is as follows: 0 = never, 1 = several times a year, 2 = monthly, 3 = several times a month, 4 = weekly, 5 = several times a week, 6 = daily, and 7 = several times a day.

Table 10.

Mean scores and standard deviations for oral reading fluency in WCPM by district, school, and grade from Fall 2015 to Spring 2017.

District	School	Grade	N	Fall 2015		Winter 2016		Spring 2016		Fall 2016		Winter 2017		Spring 2017	
				Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Partner District	School 1	2	27	65.85	15.80	90.07	16.55	104.78	15.63	87.15	14.41	110.85	17.17	127.44	20.11
		3	26	95.54	28.10	115.46	26.58	131.73	31.84	118.00	34.14	142.81	30.88	157.62	34.26
		4	33	96.00	38.81	118.64	40.70	131.58	41.68	117.27	43.08	137.06	38.83	147.64	44.48
	School 2	2	63	74.87	25.94	94.54	28.60	112.49	32.53	91.94	31.69	115.48	35.57	128.32	42.81
		3	76	76.05	34.79	100.66	38.26	116.47	42.04	95.34	39.30	113.55	41.30	132.46	41.70
		4	59	80.73	37.14	109.34	37.10	127.14	38.99	117.44	40.02	136.85	39.84	142.71	43.85
	School 3	2	64	64.55	29.10	86.91	28.52	99.95	30.20	78.33	30.66	97.59	32.41	116.06	36.75
		3	76	63.43	35.97	81.39	35.97	96.79	37.29	79.84	36.93	93.68	38.20	117.51	39.36
		4	74	83.38	37.46	103.39	36.17	118.61	36.93	101.97	39.29	120.28	39.73	130.41	35.99
Comparison District B	School 6	2	43	54.84	24.44	69.67	30.23	79.40	30.90	71.02	27.32	81.33	30.78	96.70	32.38
		3	39	68.23	30.96	78.00	33.73	91.44	33.63	82.08	35.91	90.54	34.43	108.23	35.65
		4	37	82.49	35.18	96.81	35.21	101.78	36.75	93.51	32.90	110.16	31.61	115.46	35.75
	School 7	2	58	52.76	31.98	63.64	36.86	73.69	37.23	60.62	37.10	78.09	36.80	84.83	36.65
		3	60	70.48	31.76	82.15	31.83	94.57	37.13	81.73	36.15	94.97	36.74	105.93	36.53
		4	57	75.39	35.77	85.70	34.55	100.23	35.02	90.44	31.42	104.82	34.98	105.65	33.34
	School 8	2	62	50.74	34.35	69.98	39.72	83.31	41.81	68.98	38.40	82.71	41.81	94.52	39.63
		3	74	65.24	37.32	78.18	38.64	86.51	41.20	81.24	41.23	101.66	44.68	117.31	48.67
		4	70	78.09	40.85	93.77	40.95	105.94	42.33	98.36	39.52	112.10	36.14	118.04	41.24
Comparison District A	School 4	2	253	50.22	30.28	68.68	36.07	82.52	39.81	69.58	36.45	83.25	40.09	98.38	42.85
	School 5	3	266	65.25	33.80	78.45	34.67	95.12	38.11	87.88	41.09	105.42	41.05	121.42	42.02
		4	282	98.24	31.79	109.12	35.85	126.67	38.46	116.13	37.96	130.63	37.08	154.38	42.39

Figure 7.

Fluency growth in WCPM by district and school for students in grades 2-3 from 2015 to 2017

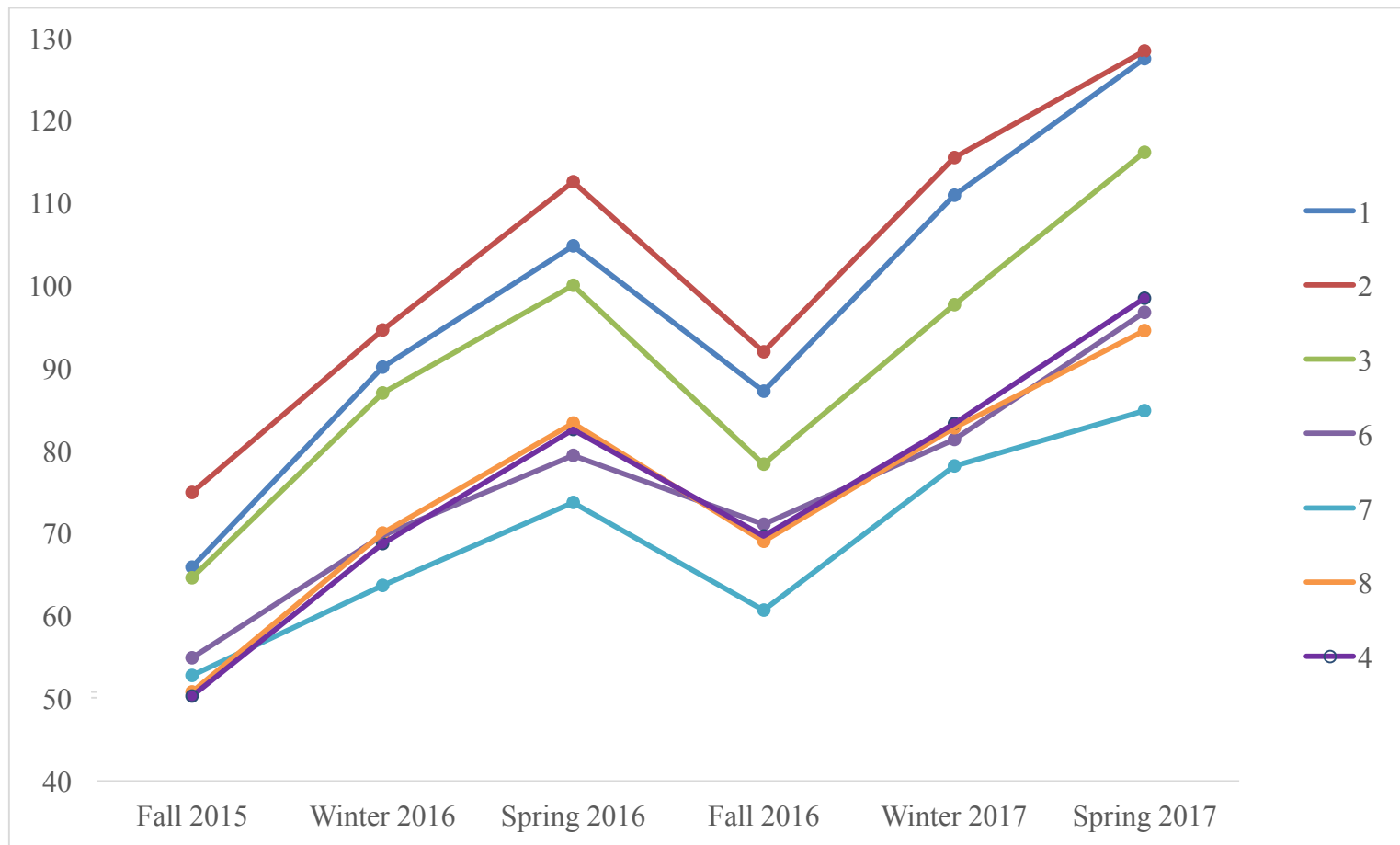


Figure 8.

Fluency growth in WCPM by district and school for students in grades 3-4 from 2015 to 2017

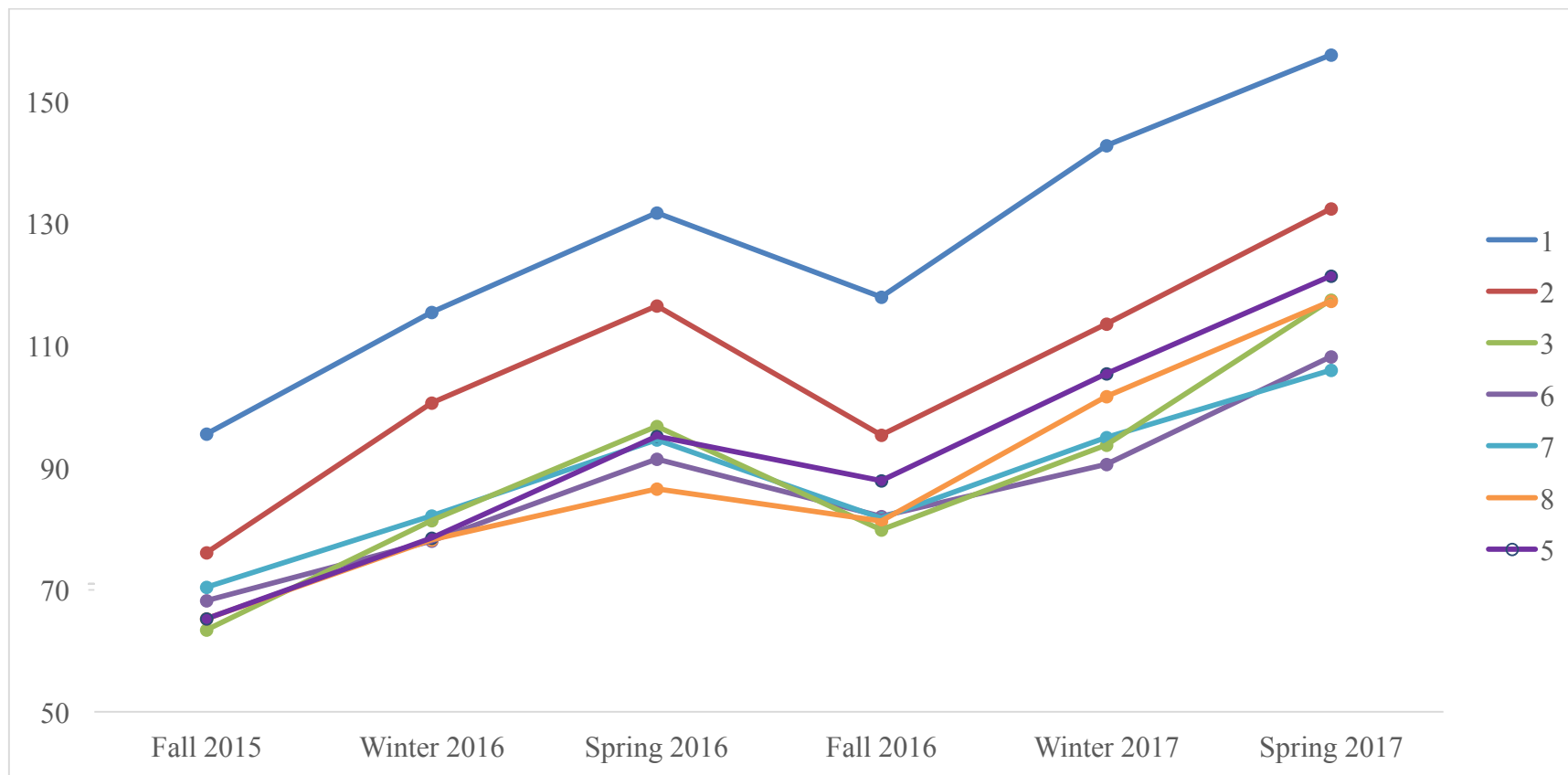
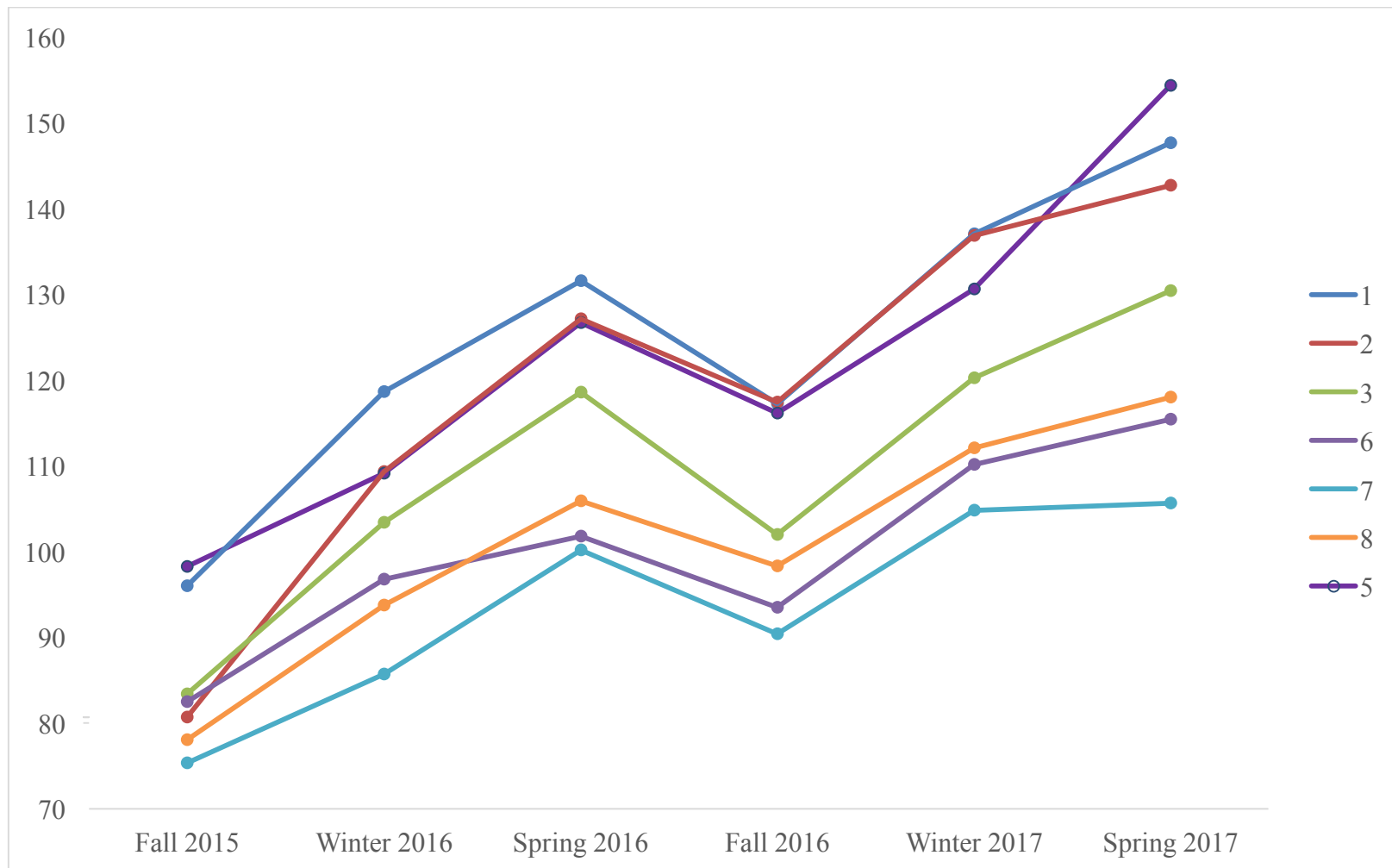


Figure 9.

Fluency growth in WCPM by district and school for students in grades 4-5 from 2015 to 2017



Comparison in Student Achievement

Growth in student achievement in reading fluency and comprehension were compared between the partner schools and schools in both comparison districts.

First, school-level and grade-level differences of growth in oral reading fluency were examined with a 6 (time) x 7 (school) x 3 (grade) repeated measures ANOVA. There was a significant three-way interaction between time, school and grade, ($F = 3.165$, $df[30, 8030]$, $p < .001$). Table 10 displays the descriptive statistics for oral reading fluency measured in WCPM in fall, winter and spring from 2015 to 2017, separated by school and grade. Figures 7-9 plot means for each school separately for students who progressed from grades 2-3, 3-4, and 4-5, respectively, in years 1 and 2 of the study.

Post hoc analyses were conducted to further understand the difference between schools and across grades. Trends in grades 2-3 demonstrate that School 1, School 2, and School 3 in the partner district experienced the most growth, while School 4 in Comparison District A and all three schools in Comparison District B experienced significantly less growth. In grades 3-4, School 1 and School 2 in the partner district experienced significantly more growth than all other schools ($p < .001$), and there was no difference between School 1 and School 2 ($p = .151$). All other differences between schools in grades 3-4 were non-significant. In grades 4-5, School 1, School 2, and School 5 were among the schools who experienced the most growth, while School 3, School 6, School 7, and School 8 experienced significantly less growth.

Table 11.

Mean scores and standard deviations for comprehension in Lexiles by district, school, and grade from Fall 2015 to Spring 2017.

District Grade	School	N	Fall 2015		Winter 2016		Spring 2016		Fall 2016		Winter 2017		Spring 2017		
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
Partner District	School 1	3	25	451.28	153.63	497.60	154.90	590.84	147.51	624.24	154.16	664.24	157.87	722.68	156.60
		4	31	562.55	222.16	611.71	242.19	696.81	235.55	732.87	223.58	771.29	224.28	808.42	230.48
	School 2	3	74	408.27	229.47	480.82	225.25	538.82	220.76	554.14	231.71	616.34	224.10	676.50	211.78
		4	59	498.42	241.01	571.98	231.10	637.34	223.87	678.05	223.29	730.24	223.12	795.20	228.49
	School 3	3	67	393.06	247.58	475.36	222.56	534.01	211.01	563.30	220.21	635.10	216.04	716.82	198.09
		4	68	536.38	186.31	600.69	181.19	665.78	169.17	702.18	180.88	764.54	182.29	813.79	179.18
Comparison District B	School 6	3	32	394.88	230.94	421.81	223.87	495.28	204.55	494.81	196.42	543.06	196.25	608.81	166.47
		4	32	494.34	187.36	531.25	214.12	571.69	232.90	578.13	241.79	656.56	226.20	718.31	202.26
	School 7	3	56	360.59	262.51	429.07	219.88	480.36	201.64	543.59	203.93	565.71	224.53	657.48	211.75
		4	56	407.98	241.31	451.18	235.17	501.64	225.48	570.13	205.87	635.66	197.32	714.98	177.79
	School 8	3	65	344.08	248.21	403.22	262.91	478.63	249.49	532.12	239.76	576.25	244.26	645.02	255.31
		4	56	495.23	289.80	517.21	299.44	552.14	300.26	649.84	288.24	713.05	265.69	808.38	242.16
Comparison District A	School 5	3	185	364.25	241.01	438.31	233.75	487.85	236.71	599.72	210.01	678.84	200.19	733.93	204.63
		4	150	560.69	237.40	637.74	228.72	718.01	223.27	729.20	222.97	767.77	223.06	793.85	239.87

Figure 10.

Comprehension growth in Lexiles by district and school for students in grade 3-4 from 2015 to 2017

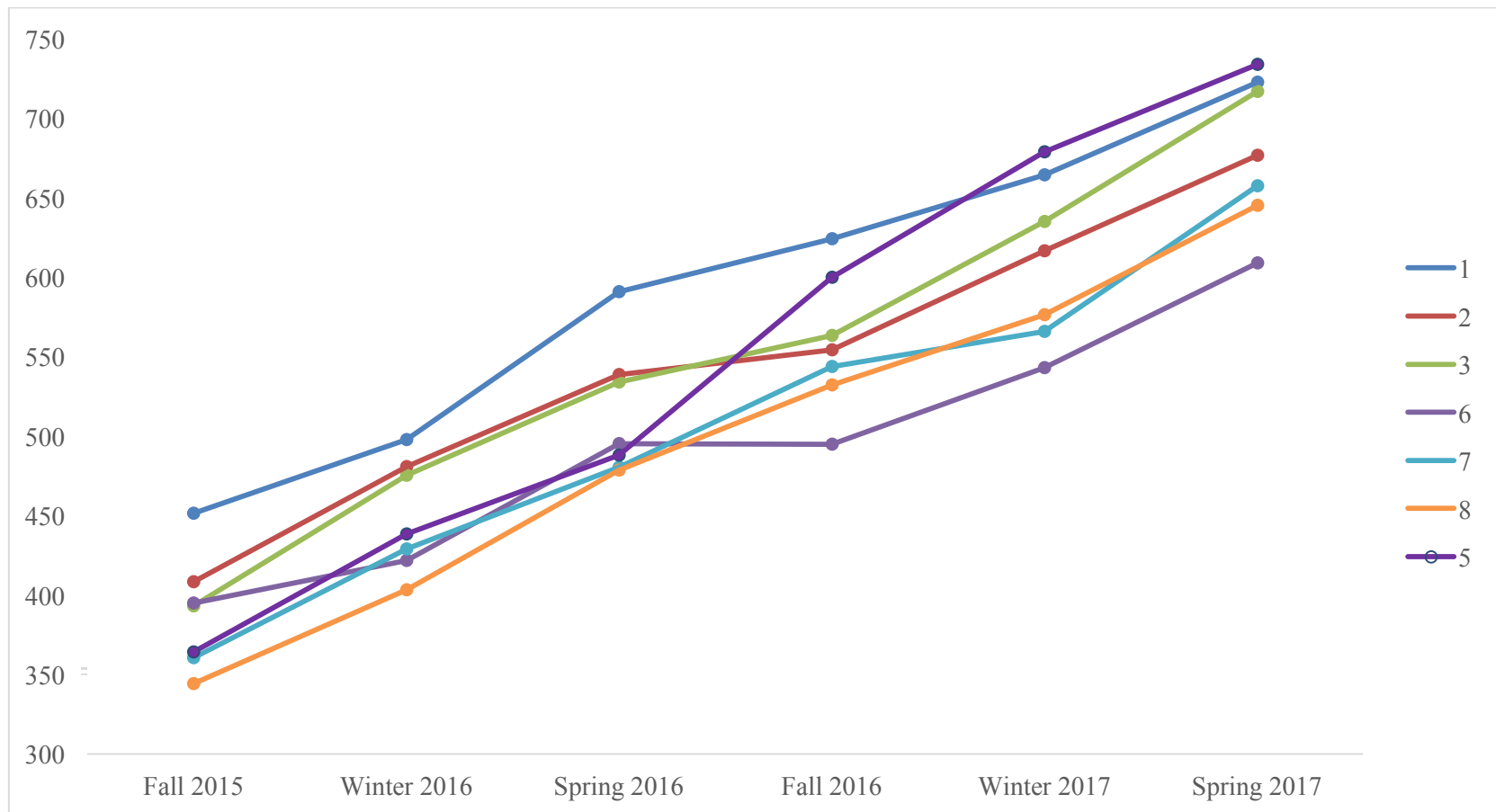
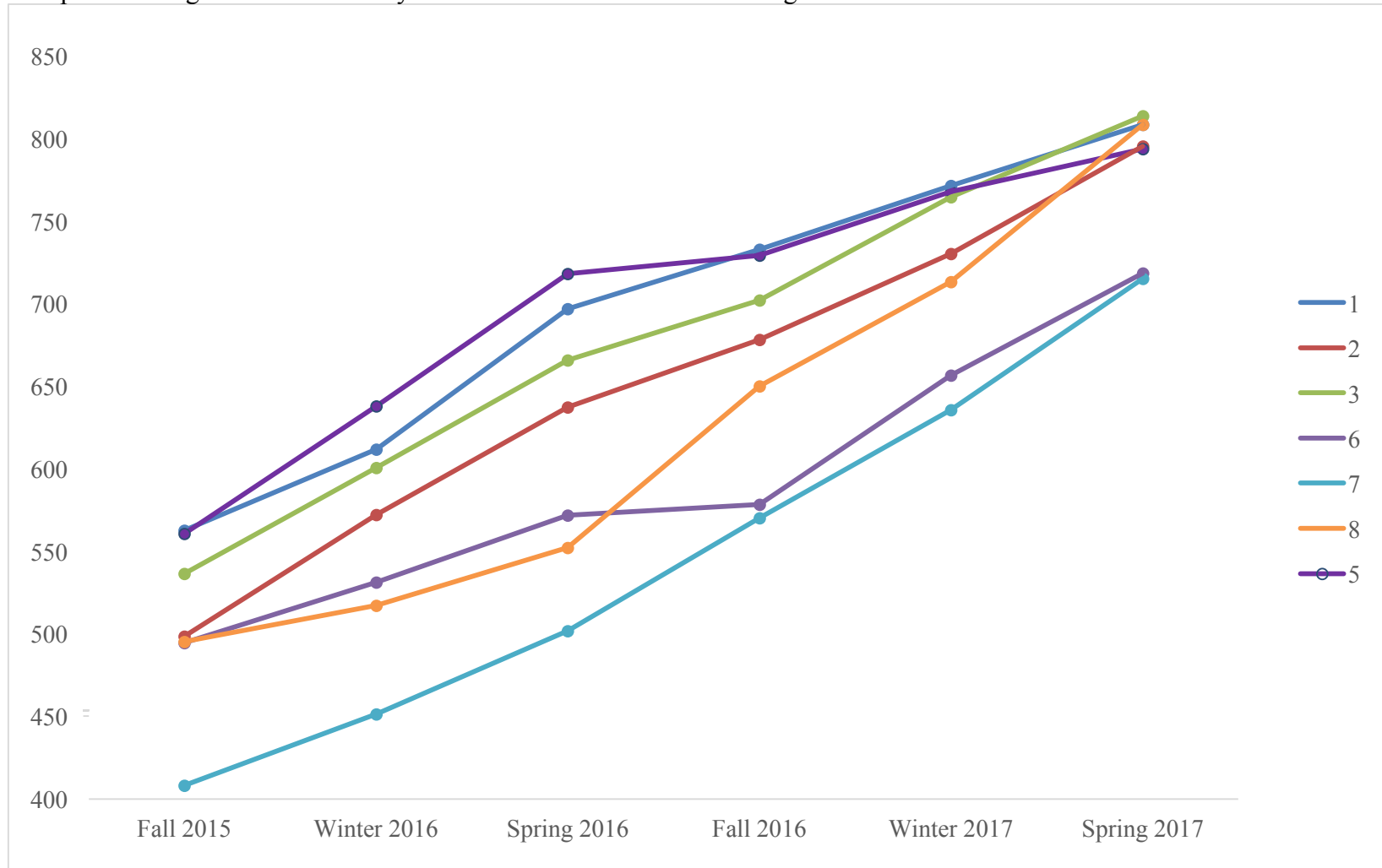


Figure 11.

Comprehension growth in Lexiles by district and school for students in grades 4-5 from 2015 to 2017



Next, school and grade-level differences of growth in reading comprehension were examined with a 6 (time) x 7 (school) x 2 (grade) repeated measures ANOVA. There was a significant three-way interaction between time, school and grade ($F = 3.165$, $df[30, 4710]$, $p < .001$). Table 11 displays descriptive statistics for reading comprehension in Lexiles in fall, winter, and spring from 2015 to 2017, separated by school and grade. Figures 10-11 plot means for each school separately for students in grades 3-4 and 4-5, respectively, in years 1 and 2 of the study. Students in grades 2-3 were not included in the analysis, because schools in the comparison districts did not administer the RI assessment in grade 2. Post hoc analyses were conducted to further understand the difference between schools and across grades. Trends in grades 3-4 demonstrate that, on average, all schools had similar rates of growth in comprehension. In grades 4-5, students in School 1, School 3, and School 5 experienced the most growth.

Overall, trends in fluency and comprehension growth demonstrated that schools in the partner district, as well as Comparison District A, experienced significant growth over the course of the project. Schools in Comparison District B did not make comparable growth in reading achievement.

Table 12.

Student counts and percentages of performance ranks on Milestones State Assessment by district and school in Spring 2016 and Spring 2017.

District	School	Performance Rank	Spring 2016		Spring 2017	
			Count	Percent	Count	Percent
Partner	School 1	Beginner	40	0.4	25	0.27
		Developing	39	0.39	52	0.55
		Proficient	17	0.17	15	0.16
		Distinguished	5	0.05	2	0.02
		Total	101	1.00	94	1.00
	School 2	Beginner	114	0.53	83	0.39
		Developing	68	0.31	86	0.40
		Proficient	29	0.13	39	0.18
		Distinguished	5	0.02	6	0.03
		Total	216	1.00	214	1.00
	School 3	Beginner	132	0.52	114	0.44
		Developing	80	0.31	101	0.39
		Proficient	35	0.14	34	0.13
		Distinguished	7	0.03	9	0.03
		Total	254	1.00	258	1.00
District B	School 6	Beginner	18	0.48	84	0.55
		Developing	15	0.41	49	0.32
		Proficient	4	0.11	18	0.12
		Distinguished	0	0	1	0.01
		Total	37	1.00	152	1.00
	School 7	Beginner	23	0.41	82	0.38
		Developing	21	0.38	85	0.40
		Proficient	12	0.21	39	0.18
		Distinguished	0	0	7	0.04
		Total	56	1.00	213	1.00
	School 8	Beginner	46	0.61	106	0.45
		Developing	17	0.22	68	0.29
		Proficient	12	0.16	50	0.19
		Distinguished	1	0.01	13	0.05
		Total	76	1.00	237	1.00
District A	School 5	Beginner	89	0.32	179	0.29
		Developing	117	0.42	270	0.43
		Proficient	66	0.23	153	0.25
		Distinguished	9	0.03	21	0.03
		Total	281	1.00	623	1.00

Table 12 presents counts and percentages for the performance ranks of the Milestones State assessment. School 1 and School 3 saw substantial increases in the number of students in the developing range, decreases in the number of students at the beginner level, and the proficient level remained stable. School 2 saw the same trends, as well as an increase in the number of students at the proficient level. Comparisons over time for the two comparison districts are difficult to make, because the comparison schools did not administer the Milestones assessment to as many students as the partner district did in Spring 2016. However, that is evident in Spring 2017 is that both districts have a high proportion of children scoring in the beginner and developing range.

Discussion

Intense work in the partner district revealed that the *Bookworms* curriculum was generally feasible and identified specific areas where the curriculum could be adjusted to improve feasibility. An unresolvable feasibility issue was the scheduling of computer-based reading interventions. In the future, schools using the *Bookworms* curriculum should only adopt interventions that can be implemented in 45-minute blocks (to replace Differentiated Instruction) or in 90-minute blocks (as a replacement core program during Shared Reading and Differentiated Instruction).

Observation data collected by the researchers and implementation data reported by teachers in the partner schools revealed that teachers had accurate understanding of the curriculum design and implemented it with the intended frequency. Small differences in frequency are likely explained by different departmentalization strategies, with one teacher teaching a segment more than once a day to different groups of students.

In terms of PD preferences, teachers in the project schools rated direct services in their classrooms, provided either by their literacy coach or by one of the researchers, as most effective in improving knowledge and instruction. This finding is similar to their ratings of video and audio clips as the most useful components of online PD modules; such offerings allow for proxy classroom modeling.

The addition of classroom libraries within the constraints of the curriculum was successful. Students chose to read the books provided. In future studies including provision of specialized classroom libraries, pre- and post-intervention motivation to read would add value to our understanding of the potential of wide reading to influence student achievement.

Student achievement in the partner schools was monitored with sensitive and specific assessments of fluency and comprehension. While the differences were not the same in the three schools, they were meaningful. For fluency, 63.5% of students in School 1 were at or above grade level at the beginning of the project, and 85.4% were fluent at the end. In School 2, 52.4% of students were fluent at the beginning of the project compared with 69.5% at the end. In School 3, 44.4% were fluent at the beginning, and 59.9% were fluent at the end of the project. These fluency status increases were realized even as the texts read were designed to be one year above grade level in difficulty. Differences in grade-level growth between schools revealed that one of the partner schools was significantly less successful than the other two in increasing fluency. Further observational data would be required to understand this difference.

Similar increases were realized in reading comprehension. It is typical that fluency improves more quickly than comprehension, and that was the case here. As with

the fluency data, the year 2 comprehension scores showed consistent progress toward grade-level expectations. Students in School 1 started with 28.7% at or above grade level in comprehension and finished with 58.4%; students in School 2 started with 32.7% at or above grade level in comprehension and finished with 54.3%; students in School 3 started with 28.7% at or above grade level in comprehension and finished with 58.4%.

Comparison data contextualized these growth trends and highlighted the role of PD in initiating and sustaining the *Bookworms* curriculum. All schools had the same free access to the curriculum and online PD materials. The partner district and Comparison District A added face-to-face PD at a lower dosage while Comparison District B did not schedule any face-to-face PD for *Bookworms*. Teachers at both districts with face-to-face PD described the components of the curriculum consistent with its design. The district without face-to-face PD described their instruction containing routines inconsistent with the curriculum. Both districts with face-to-face PD reported the same frequency of implementation, but the district without face-to-face PD reported less frequent use of the grade-level components of the curriculum and similar implementation of differentiation.

Fluency status at the beginning of second grade and growth from second grade to third grade at the three partner schools was better than in the comparison schools. It may be that increased PD for the partner schools and/or the wide reading component influenced early primary achievement and growth. That same trend was repeated in two of the three partner schools for students beginning in grade 3 and moving to grade 4. For students moving from grade 4 to grade 5, the school with half the amount of PD as the partner schools began and ended with the highest fluency; two of the partner schools

ended with very similar fluency. All three of the schools with no face-to-face PD and low frequency of grade-level instruction ended with weaker fluency scores.

Overall comprehension at the beginning of third grade and at the end of fourth grade at the three partner schools was higher than the comparison schools with one exception, the district with a low dose of PD. All three of the schools that did not receive PD experienced lower ending comprehension scores. Trends from the beginning of fourth grade through the end of fifth grade are more difficult to attribute to curriculum and PD differences. Two of the three schools with no face-to-face PD had the weakest outcomes, but one evidenced exceptional growth. Outcomes in all of the partner schools, one of the schools in the district with a low dose of PD and one of the schools with no face-to-face PD were very similar at this grade. More information about the quality of instruction beyond frequency of implementation might be necessary to explain these differences.

Over the course of this project, the researchers and the partner district established the feasibility of *Bookworms* and negotiated changes to make future implementation easier. They demonstrated and reported strong implementation across both years of the project. At their request, the research team linked evidence-based writing instruction to the existing evidence-based reading instruction. The classroom libraries added through this project enabled students to increase the number of books students read independently. Online PD offerings were examined, and those aspects teachers in partner schools found most useful were videos of classroom implementation. Student achievement in both fluency and comprehension increased consistently over time but was not yet sufficient for all students to reach proficiency on the state's literacy assessment.

Examination of implementation and achievement in comparison districts also added value to understanding the potential for student achievement growth in this high-volume evidence-based reading curriculum. The district with a low dose of PD demonstrated that maintenance of teacher implementation and continuation of student achievement growth is feasible with less intense PD over time. In contrast, the district with no face-to-face PD reported inconsistent implementation, and its students experienced less growth in achievement. These trends reveal the importance of future investigations of the amount of PD necessary for both strong implementation of the curriculum and growth in student literacy achievement.

References

- Al Otaiba, S., Connor, C. M., Folsom, J. S., Greulich, L., Meadows, J., & Li, Z. (2011). Assessment data-informed guidance to individualize kindergarten reading instruction: Findings from a cluster-randomized control field trial. *Elementary School Journal, 111*, 535-560.
- Baker, L. (2003). The role of parents in motivating struggling readers. *Reading & Writing Quarterly, 19*, 87-106.
- Connor, C. M., Morrison, F., & Katch, L. E. (2004). Beyond the reading wars: Exploring the effect of child-instruction interactions on growth in early reading. *Scientific Studies of Reading, 8*, 305-336.
- Council of Chief State School Officers (CCSSO) & the National Governors Association Center (NGAC). (2010). *Common core state standards for English language arts and literacy in history/social studies, science, and technical subjects*. Washington, D.C.: National Governors Association. Retrieved from <http://www.corestandards.org/ELA-Literacy/>
- Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher, 38*, 181-199.
- Dillon, D. R., O'Brien, D. G., & Heilman, E. E. (2000). Literacy research in the next millennium: From paradigms to pragmatism and practicality. *Reading Research Quarterly, 35*, 10-26.

- Ellis, E. S., & Graves, A. W. (1990). Teaching rural students with learning disabilities: A paraphrasing strategy to increase comprehension of main ideas. *Rural Special Education Quarterly, 10*(2), 2–10.
- Fuchs, D., & Fuchs, L. S. (1998). Researchers and teachers working together to adapt instruction for diverse learners. *Learning Disabilities Research and Practice, 13*, 126-137.
- Good, R. H., Kaminski, R. A., Cummings, K., Dufour-Martel, C., Petersen, K., Powell-Smith, K., . . . Wallin, J. (2011). *DIBELS Next assessment manual*. Eugene, OR: Dynamic Measurement Group, Inc.
- Graham, S., and Hebert, M. A. (2010). *Writing to read: Evidence for how writing can improve reading. A Carnegie Corporation Time to Act Report*. Washington, DC: Alliance for Excellent Education.
- Graham, S., & Perin, D. (2007). *Writing next: Effective strategies to improve writing of adolescents in middle and high schools – A report to Carnegie Corporation of New York*. Washington, DC: Alliance for Excellent Education.
- Graham, S., Bollinger, A., Booth Olson, C., D’Aoust, C., MacArthur, C., McCutchen, D., & Olinghouse, N. (2012). *Teaching elementary school students to be effective writers: A practice guide* (NCEE 2012- 4058). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Gregoire, M. (2003). Is it a challenge or a threat? A dual-process model of teachers’ cognition and appraisal processes during conceptual change. *Educational Psychology Review, 15*, 147-179.

- Guskey, T., & Yoon, K. (2009). What works in professional development? *Phi Delta Kappan*, *90*, 495-500.
- High, P., Hopmann, M., LaGasse, L., & Linn H. (1998). Evaluation of a clinic-based program to promote book sharing and bedtime routines among low-income urban families with young children. *Archives of Pediatrics and Adolescent Medicine*, *15*, 459-465.
- Juel, C., & Minden-Cupp, C. (2000). Learning to read words: Linguistic units and instructional strategies. *Reading Research Quarterly*, *35*, 458-92.
- Kane, T. J., Owens, A. M., Marinell, W. H., Thal, D. R. C., & Staiger, D. O. (2016). *Teaching higher: Educators' perspectives on Common Core implementation*. Retrieved from <https://cepr.harvard.edu/files/cepr/files/teaching-higher-report.pdf>
- Mathes, P. G., & Babyak, A. E. (2001). The effects of Peer-Assisted Literacy Strategies for first-grade readers with and without additional mini-skills lessons. *Learning Disabilities Research & Practice*, *16*(1), 28-44.
- McMaster, K. L., Fuchs, D., Fuchs, L. S., & Compton, D. L. (2005). Responding to nonresponders: An experimental field trial of identification and intervention methods. *Exceptional Children*, *71*(4), 445-463.
- Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly*, *40*, 184-202.
- Scholastic Inc. (2009). *Scholastic reading inventory technical guide*. New York, NY: Author.
- Shanahan, T., Callison, K., Carriere, C., Duke, N. K., Pearson, P. D., Schatschneider, C., & Torgesen, J. (2010). *Improving reading comprehension in kindergarten*

- through 3rd grade: A practice guide* (NCEE 2010-4038). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Stahl, S. A., & Heubach, K. M. (2005). Fluency-oriented reading instruction. *Journal of Literacy Research, 37*, 25-60.
- Stein, M. L., Berends, M., Fuchs, D., McMaster, K., Sáenz, L., Yen, L., & Compton, D. L. (2008). Scaling up an early reading program: Relationships among teacher support, fidelity of implementation, and student performance across different sites and years. *Educational Evaluation and Policy Analysis, 30*(4), 368–388.
- Taylor, J. E. (2008). Instructional coaching: The state of the art. In M. M. Mangin & S. R. Stoelinga (Eds.), *Effective teacher leadership: Using research to inform and reform* (pp. 10-35). New York, NY: Teachers College Press.
- U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2006). *Early Childhood Education intervention report: Phonological awareness training plus letter knowledge training*. Retrieved from <http://whatworks.ed.gov>
- U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2015). *Early Childhood Education intervention report: Shared book reading*. Retrieved from <http://whatworks.ed.gov>
- Walpole, S. (2016). Bookworms lessons [Video files]. Retrieved from <https://www.youtube.com/c/bookwormslessons>

Walpole, S., McKenna, M. C., Amendum, S., Pasquarella, A., & Strong, J. Z. (2017). The promise of a literacy reform effort in the upper elementary grades. *Elementary School Journal, 118*, 258-280.

Wexler, J., Vaughn, S., Roberts, G., & Denton, C. A. (2010). The efficacy of repeated reading and wide reading practice for high school students with severe reading disabilities. *Learning Disabilities Research & Practice, 25*(1), 2–10.