Scientific Studies of Reading

ISSN: 1088-8438 (Print) 1532-799X (Online) Journal homepage: https://www.tandfonline.com/loi/hssr20

Visual-Syntactic Text Format: Improving Adolescent Literacy

Tamara P. Tate, Penelope Collins, Ying Xu, Joanna C. Yau, Jenell Krishnan, Yenda Prado, George Farkas & Mark Warschauer

To cite this article: Tamara P. Tate, Penelope Collins, Ying Xu, Joanna C. Yau, Jenell Krishnan, Yenda Prado, George Farkas & Mark Warschauer (2019) Visual-Syntactic Text Format: Improving Adolescent Literacy, Scientific Studies of Reading, 23:4, 287-304, DOI: 10.1080/10888438.2018.1561700

To link to this article: https://doi.org/10.1080/10888438.2018.1561700

SCIENTIFIC STUDIES OF READING
2019, VOL. 23, NO. 4, 287–304
https://doi.org/10.1080/10888438.2018.1561700

Visual-Syntactic Text Format: Improving Adolescent Literacy

Tamara P. Tate, Penelope Collins, Ying Xu, Joanna C. Yau, Jenell Krishnan, Yenda Prado, George Farkas, and Mark Warschauer

University of California, Irvine

ABSTRACT

Seventh- and 8th-grade students in a within-teacher randomized control study read from visual-syntactic formatted text for 44 min per week over the course of 1 year. On the annual state assessment, we found small statistically significant improvements on the overall English Language Arts scaled score (ES = 0.05, $p < .05$) and the writing assessment (ES = 0.07, $p < .01$) for the treatment group compared to the control group. We found no interactions between gifted, special education, or English learner classification and treatment status on the effect on overall English Language Arts score, but our categorical and subgroup analyses showed that the use of visual-syntactic text formatting provided a modest benefit to middle school students who were near or at grade level in the prior school year.

For most typically achieving students entering middle school, continued reading at grade level depends on more than their automatized decoding skills. Instead, successful performance in middle school reading increasingly depends on students' effective processing of language and text structures. At the same time, the linguistic complexity of the texts students must read in seventh and eighth grades increases dramatically. One way of scaffolding students' understanding of complex texts and ability to write critically about those texts is to modify the formatting of the text to make its underlying structure more visible and encourage students to synthesize their understanding of the content more frequently. Research studies have shown that texts that have been modified to highlight prosodic cues and syntactic structures, such as phrase boundaries, can facilitate linguistic processing and reading comprehension (e.g., Hirotani, Frazier, & Rayner, 2006; Jandreau & Bever, 1992). Visual-syntactic text formatting (VSTF) is produced by natural language processing techniques that parse text to highlight phrase and clause boundaries, thereby scaffolding students in processing complex syntactic structures while leaving the content, vocabulary, and syntax of the text unchanged. This article presents the findings of a within-teacher randomized control trial of a digital intervention in seventh- and eighth-grade English Language Arts (ELA) classes designed to improve students' reading and writing performance. The intervention was implemented over 1 school year using VSTF, digital devices in classrooms, and integrated professional development designed to foster pedagogically sound use of the devices and formatted text.

Adolescent literacy

According to the Reading Systems Framework (Perfetti & Stafura, 2014), reading comprehension is accomplished through linkages between the word identification and the comprehension system, with readers updating their representations of the text as they decode the text. More specifically, the word identification processes yields semantic representations of the meanings of words and phrases. These

CONTACT Tamara P. Tate tatet@uci.edu School of Education, University of California, 3200 Education, Irvine, CA 92697-5500.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hssr.
Supplemental data for this article can be accessed here.

semantic representations are the linkage between the word identification and word-to-text integration systems, and they directly influence and are influenced by sentence representations. The sentence representations, in turn, shape and are shaped by both the reader's text models, or representations of the propositions expressed in the text, and situation models, or mental representations of the scenarios described by the text (Perfetti & Stafura, 2014). Reading comprehension is a dynamic process, as readers update their mental model online, integrating new word-level information with text and situation models (Kintsch, 2005; Perfetti & Stafura, 2014). If readers are slower to update their mental model of the text, there can be consequences for maintaining coherence across sentences, as memory resources are limited (Perfetti & Stafura, 2014). This updating is part of "wrap-up processes," whereby individuals pause at clause and sentence boundaries for meaning construction and integration (Stein-Morrow, Shake, Miles, Lee, Gao & McConkie, 2010; Tiffin-Richards & Schroeder, 2018). Proficient readers parse text into phrases or clauses (Fuchs, Fuchs, Hosp, & Jenkins, 2001). Thus, critical comprehension skills include sensitivity to the prosodic features of text and an understanding of complicated syntactic and organizational text structures (Cain, 2007; Fitzgerald & Shanahan, 2000).

Although most middle school students, particularly those who are reading near or at grade level, will have automatized their word identification skills, students still need support in processing the complex language structures they face as texts become progressively more challenging throughout secondary school (Fang, Schleppegrell, & Cox, 2006; Williams, Fitzgerald, & Stenner, 2013). These complex language and text structures may pose barriers to students' ability to construct coherent mental representations of the texts (Biancarosa & Snow, 2004; Faggella-Luby & Deshler, 2008; Lipka & Siegel, 2012; Mokhtari & Thompson, 2006; Torgesen et al., 2007; Veenendaal, Groen, & Verhoeven, 2014). These challenges have come more into focus since adoption of the Common Core State Standards, with their emphasis on the use of informational and complex texts (CCSSI, 2017) and on high-level reading and writing skills and mastery of language structures (Porter, McMaken, Hwang, & Yang, 2011). Thus, middle school students are positioned to benefit from scaffolding designed to aid in the processing of text structures if they are sufficiently adept at word identification (i.e., reading and writing near or at grade level) because the texts in the curriculum are becoming sufficiently complex to require the support (as compared to, e.g., simpler texts found in elementary grades). The skills of reading and writing are closely linked (Fitzgerald & Shanahan, 2000; Shanahan, 2006) and share many of the same orthographic, phonological, and working memory subprocesses (Berninger, Cartwright, Yates, Swanson, & Abbott, 1994).

Prior interventions

Although text complexity may be thought of as the interactions among text features, reader characteristics, and task demands, a common pedagogical approach to making texts accessible is to simplify texts to match students' reading levels (Bunch, Walqui, & Pearson, 2014; Oh, 2001). However, simplified texts may not be an ideal solution to supporting struggling readers, as they may limit the overall coherence of the original texts, reduce students' exposure to important vocabulary and text structures that students will eventually need, yet remain syntactically complex texts (Crossley, Louwerse, McCarthy, & McNamara, 2007).

Another approach to making texts more accessible for diverse and struggling readers is by adjusting the print itself. These alterations have taken the form of simple changes in letters, spacing between lines and sentences, or capitalization (Johnson, Bui, & Schmitt, 2018), or the insertion of additional spaces between phrases (Bever, Jandreau, Burwell, Kaplan, & Zaenen, 1990; Hirotani et al., 2006; Jandreau & Bever, 1992). Although a number of studies found that print modifications may support English learners' understanding of English texts (Lee, 2007; Lee & Huang, 2008; Shook, 1994; Simard, 2009), others reported that simple print modifications, such as underlined and bold letters, may have little or no effect on language learning (Leow, 1997; Overstreet, 1998).

VSTF is a form of textual organization that uses natural language processing techniques to automatically parse text to highlight sentence and phrase structures. Specifically, VSTF breaks sentences up at salient clause and phrase boundaries, fits each row of text into one or two fixation eye spans, uses a cascading pattern to denote syntactic hierarchies, and creates visual clusters across multiple rows that help readers retain and integrate multiphrase images in their mind. VSTF also renders active verbs in colored font to further highlight meaning. The end result is a streamlined column of text that allows more efficient eye movement and syntactic processing (see Figure 1 for an illustration of VSTF text compared to traditional block text without formatting). Thus, VSTF alters the presentation of texts' print to scaffold readers in processing complex syntactic structures while leaving the content, vocabulary, and syntax of the passage unchanged. In this way, VSTF supports syntactic processing and reading comprehension by facilitating the word-to-text integration processes by making sentence representations more salient.

Reading in VSTF is also thought to free up cognitive resources, such as working memory, which would otherwise be involved in making sense of texts. Studies on the relevant cognitive process have demonstrated that meaningful unit-based segmentation of information has beneficial effects on learning, either by reducing repeated attention-switching between old and new information (Barrouillet, Gavens, Gaillard, & Camos, 2009) or by reducing the cognitive load that learners need to identify salient boundaries between information (Schwan, Garsoffky, & Hesse, 2000; Wouters, Paas, & van Merriënboer, 2008). VSTF text, with its explicit phrasal segmentation, may help reduce readers' cognitive loads by cueing when wrap-up processes should occur (Hirotani et al., 2006). Facilitating wrap-up processes at phrase and clause boundaries enables readers to integrate the meanings of the phrases so that readers may construct richer and more lasting mental models of the text (Just & Carpenter, 1980).

Initial studies by the developers of VSTF reported gains in reading comprehension, reading speed, retention, and proficiency among high school and middle school students (Walker et al., 2007; Walker, Schloss, Fletcher, Vogel, & Walker, 2005; Walker & Vogel, 2005). Independent examination of the effects of VSTF on young adolescents' literacy achievement revealed generalizable gains in reading comprehension for students in sixth grade but not in fourth grade (Park & Warschauer, 2016; Park, Warschauer, Collins, Hwang, & Vogel, 2013; Park, Zheng, Lawrence, & Warschauer, 2012), perhaps suggesting that a baseline level of literacy skill is needed or that a certain level of text complexity is required to make the scaffolding salient. Further, the effects of VSTF for sixth-grade students were found not just for the composite state ELA assessment score ($b = .07$, $p < .10$) but also in three subcategories: word analysis ($b = .11$, $p < .05$), written conventions ($b = .12$, $p < .05$), and writing strategies ($b = .13$, $p < .05$; Park, Xu, Collins, Farkas, & Warschauer, 2018). The study was

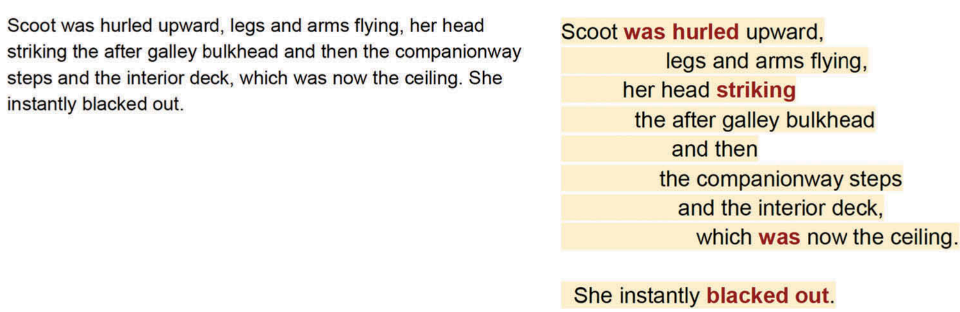

Figure 1. Traditional block-formatted text on the left and the same text in visual-syntactic formatted text on the right.

limited by the class-level randomization, which made it difficult to isolate teacher and class effects from the effects of the intervention.

The current study

The current study was designed to provide a more rigorous examination of VSTF's effectiveness for improving middle school student literacy. First, we randomly assigned classes to the treatment and control conditions within teachers, thereby limiting differences in teacher quality across conditions. Acknowledging the potential that different digital devices might have different affordances, we also examined the effects of using VSTF on the digital devices widely used in schools—Chromebooks and iPads. Finally, although past studies provided teachers with minimal professional development, this was one of the first studies to provide teachers with robust professional development opportunities to guide their use of formatted texts.

This study focused on answering the following research questions:

RQ1: To what extent does students' reading of texts in VSTF improve academic outcomes on standardized state assessment measures of reading and writing compared with students in control classes who read the same texts in the usual block-text format?

RQ2: Are there differential effects of VSTF for students based on (a) subgroups, such as English learners, gifted students, and students with disabilities; or (b) prior academic performance?[1]

Method

Participants

District context

The study took place in an urban school district that enrolled more than 45,000 students during the 2015–2016 school year. Of these students, 68% qualified for free and reduced-price lunch and 39% were English learners (California Department of Education, n.d.). All 10 of the district's middle schools participated and shared a common curriculum textbook series, pacing guide, and quarterly benchmark tests.

Teacher participants

All 93 middle school ELA teachers were invited to participate in the study, and 59 teachers initially agreed to participate. Each participating teacher received a study information sheet, gave informed consent, and was paid at the overtime rate to complete all research activities. Over the course of the year, seven teachers discontinued their participation (two left due to medical leave for part of the year, and five chose to discontinue their participation). Thus, our final sample consisted of 54 teachers (including four teachers who cotaught) and their 170 seventh- or eighth-grade ELA classes. Each teacher received a technology cart equipped with a classroom set of either iPads with keyboards or Chromebook computers at the end of the previous school year or the beginning of the intervention school year.

Student participants

Our final analytic sample included the 3,453 students (1,400 in seventh grade, 2,053 in eighth grade) who remained in the same class with the same teacher for the full academic year, with 1,780 students in the treatment condition and 1,673 students in the control condition.[2]

[1]We found no statistical difference between devices. Please see the supplementary materials online for additional detail.
[2]See the supplemental materials for additional details regarding our sample.

With the exception of more students receiving special education services in control classes (χ^2 = 9.843, p = .002), treatment and control students are not significantly different demographically or in their achievement on the previous year's state ELA assessment (Table S-1 shows the treatment and control group characteristics of the final sample). Parents/guardians were informed of the study in a letter sent home, which was translated into Spanish and Vietnamese. Implied consent was granted if parents did not object to their child participating.

Measures

We collected student demographic and achievement data. Demographic data included the student's grade, gender, ethnicity, whether the student qualified for free/reduced lunch, whether the student was an English language learner, and whether the student was identified for GATE or special education. Achievement data were drawn from the annual state assessment, the Smarter Balanced assessment's overall ELA score, and the reading and writing subtests. The overall ELA score is a continuous scaled score from approximately 2,258 to 2,769 in Grades 7–8 in 2017, with a mean of 2,542.4 in Grade 7 and 2,558.7 in Grade 8 (California Department of Education, n.d.). These data also included the state assessment's ELA subscales for reading, writing, listening, and research and inquiry, which were scored 1 (below standards), 2 (near standards), or 3 (above standards). Although the categorical subscores were less sensitive measures than the overall ELA score, they allowed us to examine differences in effect on specific components of the ELA test, particularly reading and writing. The Smarter Balanced assessment is administered each May online, with scores from the previous year used as baseline and from the current year as the outcome. The Smarter Balanced assessment does not use VSTF; thus, all treatment and control students completed this assessment using traditional block formatting.

For each teacher, we observed both a control and a treatment class in the fall, winter, and spring. Graduate student researchers scored class activities every 5 min to gather information on the types of activities that students were engaged in and the use of VSTF or block text reading materials. Teachers also completed weekly reports on their study experiences, including a report of the minutes spent in each class engaging with either VSTF or block text.

Study design

We used a randomized controlled trial study design in which classes were randomly assigned to the treatment (VSTF texts) or control (traditional block-formatted texts) condition within grades, schools, and teachers.[3] Thus, each teacher taught both treatment and control classes, ensuring that teacher characteristics were similar across conditions. This design reduces the likelihood of confounding teacher knowledge and teacher quality with the treatment. Ethical review and oversight was provided by the University of California Irvine Institutional Review Board, which confirmed that the research project was exempt research in accordance with 45 CFR 46.101, and conforms to U.S. Federal Policy for the Protection of Human Subjects.

Professional development and teacher support

All teachers participated in the professional development, consisting of a discussion of the research motivation behind the study, facilitation of technology use, and explicit instruction in teaching close reading strategies (see Figure 2 for the first professional development meeting agenda). In addition to being supplied with a technology cart equipped with a classroom set of iPads or Chromebook computers, teachers were paid at their extra-duty rate as an incentive to attend professional development sessions, complete weekly reflection forms, and supply writing assessments. To increase teachers' support and minimize attrition, funding was provided for a teacher-on-special-assignment

[3]We complied with one teacher's request that her first period be a control class.

and for a staff member to provide technical support for the use of iPads or Chromebooks during classroom instruction.

Treatment and control conditions

Teachers remained responsible for creating the curriculum for their classes in both conditions. In the treatment classes, students were to engage with the standard texts in class for 50 min each week using VSTF, whether the reading was done digitally with study-based e-books, Google documents, or dialectical journals, or with paper-based printouts of the texts. The 50 min could be in one or more class periods, solely in the teachers' discretion. In the control classes, teachers taught the same material, but the texts were *not* reformatted into VSTF. Researchers observed each teacher three times during the year, attending both a control and treatment class on the same day generally. We not only confirmed fidelity of implementation during these observations, as noted next, but also saw that teachers were generally using the same presentation slides and pedagogy in each class, simply exchanging VSTF and block text examples of the text being discussed that day, and similarly the students read either digitally on a GoogleDoc or on paper in the appropriate format, but classes were largely doing the same reading activity (e.g., a first read of a text, or looking for new vocabulary in a text) in the same mode (paper-based or digital) regardless of condition.

Fidelity of implementation measures

We used two measures to assess fidelity of implementation: teacher self-reports through a weekly reflection form and a modification of the Pathway Observation Measure (Olson et al., 2012) completed during our observations. On average, teachers completed 29 reflection forms (an 85.3% completion rate). Data from completed forms indicated that, on average, teachers used the study's reading materials approximately ten minutes more in their VSTF classes (M = 44.11, SD = 15.47) than in their control classes (M = 33.67, SD = 17.56), $t(50)$ = 2.26, p = .03. However, there was great variability in the use of the study materials across teachers (ranging from a weekly mean of 12 min to 112 min). Our observations found no use of VSTF in control classes.

Data analytic strategies

We used several regression approaches to examine the impact of the VSTF intervention on students' annual assessment scores and categorical change in the reading and writing domain scores, respectively.

Effects on annual assessment scores

We used fixed effects (teachers) regression and clustered standard errors (classes) to analyze the treatment effect of VSTF on overall ELA, reading, and writing scores. Our analyses were restricted to students who had had a full year of either the VSTF intervention or a full year of the control condition and who had spent the entire year in the same class, with the same teacher. Given the limited nature of the intervention and the time generally required to fully implement any technology-based practice, we felt that a full year of the intervention was necessary. However, for comparative purposes, we also ran the same analyses on the full sample based on placement in treatment or control in Semester 1 for all students, regardless of their condition status in Semester 2.

To estimate the program's effect on student academic outcomes (overall ELA, reading, and writing scores) we analyzed data at the student unit of analysis. These include tests involving the gender, ethnic, racial, and linguistic diversity of our population and allow us to analyze heterogeneous effects, one of our research questions. A variety of regression techniques were used to increase the precision of our estimates and provide correct standard errors accounting for the clustering of students in classes. Because our design involved random assignment within each teacher, we used teacher fixed effects estimation to guarantee that outcomes for treated students were compared only to those for control students who had the same teacher. Standard errors were

Thursday

Topic	Intended Results	Time
Welcome and Setting the Stage	Introduce key personnel from the Live Ink Team and today's objectives.	**8:00AM-8:35AM**
Live Ink Foundations	Dr. Warschauer will share Live Ink history, research, and purpose. Dr. Schleppegrell will preview lesson on the affordances of Live Ink format for reading comprehension .	**8:35 AM -9:10AM**
Teacher Commitments	Dr. Jenkins, Dr. Collins, and the Live Ink team will discuss 2016-2017 Live Ink teacher commitments.	**9:10AM-10:00AM**
	BREAK	**10:00AM-10:10AM**
Introduction to eBook Readers	Teachers will receive an introduction and interact with the eBook Reader.	**10:10AM-11:00AM**
	LUNCH	**11:00AM-12:00PM**
Spend-A-Sticker Brainstorm	Teachers will collaborate on how to engage students in 50 minutes of interaction with Live Ink Text.	**12:00PM-12:50PM**
Web Clip Read Introduction	Teachers will interact with Web Clip Read interface as a tool that can be used in the classroom.	**12:50PM-1:15PM**
Survey	Teachers will complete a survey to gather data on their beliefs about teaching ELA	**1:15PM-1:35PM**
	BREAK	**1:35PM-1:45PM**
Reflection Q and A	Teachers will discuss Live Ink Foundations, Teacher Commitments, eBook Readers, and Web Clip Read.	**1:45PM-2:05PM**
Review	Teachers will participate in a Live Ink review game via Socrative.	**2:05PM-2:30PM**
Next Steps	Teachers will review the day's objective and receive Day 2 information.	**2:30PM -3:00PM**

Friday

Topic	Intended Results	Time
Welcome	Recap of Day #1 Introduce Mary Schleppegrell and Purpose of the Day	**8:00AM-8:05AM**
Model Lesson	Teachers will participate in a model lesson given by Dr. Schleppegrell using the 7th grade text, "Rogue Wave."	**8:05AM-9:15AM**
Lesson TOTD and Reflection	Teachers will reflect on the model lesson and complete a TOTD using a Google Form.	**9:15AM-9:45AM**
	Break	**9:45AM-10:00AM**
Break Out Session: Basics and Digital Text Annotation	Teachers will participate in a break out session outlining the basics of either Notability or Google Documents and practice using digital text annotations.	**10:00AM-11:00AM**
	Lunch	**11:00AM- 12:00PM**
Team Collaboration	Teachers will collaborate on daily Live Ink specific lessons utilize tools, skills, and information shared during the presentations.	**12:00PM - 1:30PM**
	Break	**1:30PM-1:40PM**
Team Collaboration	Teachers will collaborate on daily Live Ink specific lessons utilize tools, skills, information shared during the presentations.	**1:40PM-2:45PM**
Closing	Teachers will generate and share lessons on Google Drive and complete Evaluation Forms.	**2:45PM - 3:00PM**

Figure 2. Agenda for 2-day professional development.

clustered at the class level to account for the nonrandom assignment of students to classes. We used the standardized overall ELA score, reading score, and writing score as outcome variables. We controlled for race/ethnicity, gender, free or reduced lunch status, English learner status, gifted designation, and special education status (students with an Individualized Education Plan). We also controlled for prior achievement levels in each of the ELA domain areas (reading, writing, listening, and research/inquiry). Using a fixed effects model is appropriate due to our design, which randomly assigned classrooms within each teacher. This assignment and modeling eliminates any effects of teacher-level variables from the analysis, ruling out "between" variation so that fixed, unobserved differences in our higher level unit (teachers) do not bias our results. Although HLM was considered, a fixed effects regression with clustered standard errors makes a smaller number of assumptions and produces results that are equally reliable as those that would be obtained from HLM. In particular, the clustered standard errors methodology is to be preferred over HLM when, as is the case here, the goals of the study do not involve a decomposition of the residual variance within and between aggregate units (McNeish, Stapleton, & Silverman, 2017).

We also tested for moderating variables by including interactions with time spent using VSTF as reported by teachers, device type (iPads or Chromebooks), demographic subgroups, and prior achievement levels. Finally, we conducted subgroup analyses on students identified as gifted, English learners, or in special education, as well as students at each of the levels on the prior reading and writing scores, by running the same analyses solely on students in each of these groups separately.

Effects on categorical change in reading and writing proficiency levels

As a robustness check, we used multinomial regression to test the treatment effect on students' categorical improvement in the reading and writing domains. Specifically, we divided the student sample into three groups by their prior state assessment scores in the same domain (i.e., Levels 1, 2, and 3, corresponding to below, near, and above standards) and examined whether and to what extent receiving the intervention would enhance the probability of having an improvement in their proficiency level based on prior proficiency level. We conducted a series of multinomial logistic regressions separately for students in each of the prior proficiency levels. For example, for students whose proficiency was at Level 2 in the pretreatment assessment, we investigated their relative probability of performing at Level 1 or Level 3 if they had or had not received the intervention. The estimation equation was the following:

$$\Pr(y = m|X) = \frac{e^{x\beta m|b}}{\sum_{j=1}^{J} X\beta_{m|b}}$$

We are estimating the probability that the posttreatment categorical proficiency level results in any $J = 3$ outcomes m, where $m = \{1, 2, 3\}$. The subscript on β indicates that a separate vector of coefficients is estimated for each of the outcome categories in reference to the base category b. As the purpose of this analysis is to examine students' improvement, we always use students' prior proficiency level as the reference group. In this example, Level 2 is the base group. Thus, the estimate of this equation results in two vectors of coefficients, one corresponding to the probability that a student performed at Level 3 (relative to Level 2) and one corresponding to the probability that a student performed at Level 1 (relative to Level 2). Student-level covariates and teacher fixed effects were included, as in the ordinary least squares model. Standard errors were also clustered at the classroom level.

The same analyses were carried out six times, for two outcome variables and three student groups, respectively. However, for students who had a pretreatment proficiency at Level 1, few of them managed to improve two levels to Level 3. For this analysis, we thus collapsed Level 2 and Level 3 in the outcome category. Similarly, few students who were at Level 3 at pretreatment performed at Level 1 in the treatment year, and thus Level 1 and Level 2 were collapsed. Therefore, for students at Level 1 or Level 3, logistic regression analyses were conducted to estimate the probability of changing to Level 2 or Level 3 (relative to Level 1, for Level 1 students) or the probability of changing to Level 1 or Level 2 (relative to Level 3, for Level 3 students).

Results

Descriptive information

We found that teachers did not average the intended 50 min of classroom use of VSTF, but rather treatment teachers reported an average of 44 min of use. Control classes were reported to use 10 min less of block-formatted text, but given that all traditional text is effectively in the "control" format, we believe that time to be underreported, with teachers reporting only the use of the online or digital texts that corresponded to the texts formatted in VSTF for their treatment classes. We suspect that all classes read traditional block texts (e.g., in their books) for additional minutes each week, but teachers found it salient to retroactively report only the time spent with consciously formatted text. Because teachers were observed conducting essentially the same class in each condition, we are hesitant to ascribe much to the differences in minutes reported between conditions. We tested our models to determine whether reported time was a significant predictor of improvement and found that reported time was neither practically nor statistically significant and, thus, dropped it from our analyses.

Descriptive information on mean student achievement (and standard deviation) on the state annual assessment in the year prior to our intervention and in the spring of our intervention by grade, condition, and device are set out in Table S-2. The district routinely performs well overall compared to similarly situated districts on the annual state assessments. In some respects, this reduced the opportunity to improve students' scores as the large number of students in the highest proficiency level created a ceiling effect.

Analyses of treatment effects

Effects on annual assessment scores

Table 1 shows the results of estimating VSTF effects on overall ELA, reading, and writing test scores. For each outcome, the first model shows a regression with the dichotomous treatment variable as the only predictor. This is equivalent to a difference in means between the treatment and control groups. Models 1, 3, and 5 indicate that students exposed to the VSTF intervention gained more than control students by, respectively, .08 *SD* on the overall ELA assessment, .05 *SD* on the reading subtest, and .10 *SD* on the writing subtest, although none of these are statistically significant.

The second model shows the regression with a full set of controls, including the four subtest scores of the ELA from the prior year, and includes teacher fixed effects (dummy variables for all but one of the teachers). Because these added predictors decrease the standard errors of the treatment coefficients, the treatment now achieves statistical significance for two of the three outcomes: the overall ELA assessment (ES = .05, p = .026) and the writing assessment (ES = .07, p = .005). The .04 effect on reading was not statistically significant, with p = .118.

With respect to heterogeneous effects, we found no interactions with gifted, special education, or English learner classification on the effect on overall ELA, reading, or writing scores (see Table S-3). However, our subgroup analyses (see Tables S-4, S-5, and S-6) show that there was no statistically significant treatment effect for students identified as gifted or in special education, and the treatment effect for English learners on the writing subscore (ES =.08, p = .068) did not reach statistical significance at the $p < .05$ level. We found no statistically significant interaction with prior reading or writing scores (Tables S-7 and S-8). Our subgroup analysis based on prior reading and writing scores showed only one statistically significant effect, students previously scoring a 2 on writing received a significant treatment effect on the current year writing score (ES =.14, p = .000; see Figure 3). Our categorical results, discussed next, show that there was, however, a *nonlinear* interaction based on student performance levels.

Table 1. Regressions Predicting Overall ELA, Reading, and Writing Scores on Analytic Sample.

	(1) ELA	(2) ELA	(3) Reading	(4) Reading	(5) Writing	(6) Writing
Treatment	0.08	0.05*	0.05	0.04	0.10	0.07**
	(0.10)	(0.02)	(0.08)	(0.02)	(0.08)	(0.03)
Hispanic		−0.01		−0.03		−0.02
		(0.03)		(0.05)		(0.04)
Asian		0.29***		0.18***		0.28***
		(0.04)		(0.05)		(0.04)
Black		−0.19		−0.13		−0.02
		(0.12)		(0.18)		(0.17)
Filipino		0.15		0.14		0.11
		(0.09)		(0.11)		(0.13)
Male		−0.07***		0.02		−0.12***
		(0.02)		(0.03)		(0.03)
Socioeconomically disadvantaged		−0.08**		−0.07		−0.09**
		(0.02)		(0.04)		(0.03)
English learner		−0.25***		−0.19***		−0.24***
		(0.03)		(0.04)		(0.04)
Gifted education		0.35***		0.25***		0.24***
		(0.03)		(0.04)		(0.04)
Special education		−0.17**		−0.16*		−0.05
		(0.06)		(0.06)		(0.07)
Prior reading		0.23***		0.21***		0.15***
		(0.01)		(0.02)		(0.02)
Prior writing		0.21***		0.19***		0.22***
		(0.02)		(0.02)		(0.02)
Prior listening		0.09***		0.08***		0.07***
		(0.01)		(0.02)		(0.01)
Prior research/Inquiry		0.20***		0.17***		0.15***
		(0.01)		(0.02)		(0.02)
Constant	−0.04	−0.07	−0.05	0.10	−0.06	0.10
	(0.07)	(0.06)	(0.06)	(0.09)	(0.06)	(0.06)
N	3,258	3,125	3,257	3,125	3,257	3,125
R^2	.002	.703	.001	.498	.002	.485

Note. Teacher variables for fixed effects not shown in table, but included in analyses with controls. Standard errors are in parentheses. ELA = English Language Arts.
*$p < .05$. **$p < .01$. ***$p < .001$.

Given the attrition from students changing teachers, classes, or conditions midyear, we also ran the same analyses using the entire sample to gauge the effect of being placed in the treatment condition Semester 1 regardless of condition placement in Semester 2. The results are similar to those found with the more limited analytic sample (see Table S-9).

Effects on categorical change in reading and writing proficiency levels

Table 2 presents a descriptive breakdown of students' posttreatment reading proficiency level by pretreatment reading proficiency level, for treatment and control samples, respectively. Row totals indicate the total number of students in each proficiency level before the intervention, and column totals indicate the number of students in each proficiency level after the intervention. Each cell represents the number of students with specific pretreatment and posttreatment proficiency levels. The numbers of students in these cells are percentaged across the rows in order to show, separately for treatment and control students, year-to-year movement across the proficiency levels.

In the first row of Table 2 we see that, among students at the first reading proficiency in the prior year, the treatment and control groups had very similar distributions across the reading proficiency levels after program implementation. By contrast we do find meaningful treatment-control differences for students at the second reading proficiency in the prior year: In the treatment

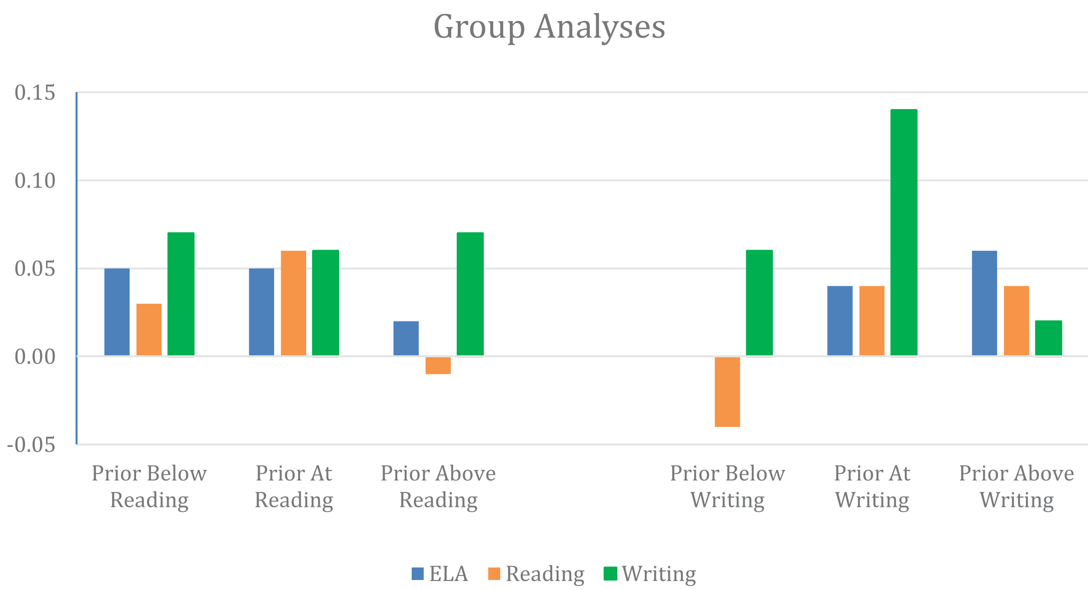

Figure 3. Group analyses, indicating effect size of impact on treatment on overall ELA score, reading score, and writing score, for students previously scoring below, at, and above standard on the prior year's reading or writing scores. Only the effect of treatment on the posttreatment writing score for students previously scoring at or near grade-level standard for writing is significant (ES = .14, p = .000).

Table 2. Prior and Current Year Reading Proficiency Level by Treatment Condition.

	Current Year			
Prior Year	Level 1	Level 2	Level 3	Total
Treatment sample				
Level 1	183	162	12	357
	51.26%	45.38%	3.36%	100%
Level 2	111	491	235	837
	13.26%	58.66%	28.08%	100%
Level 3	2	143	343	488
	0.41%	29.30%	70.29%	100%
Total	296	796	590	1,682
	17.60%	47.32%	35.08%	100%
Control sample				
Level 1	172	162	8	342
	50.29%	47.37%	2.34%	100%
Level 2	115	492	201	808
	14.23%	60.89%	24.88%	100%
Level 3	8	120	300	428
	1.87%	28.04%	70.09%	100%
Total	295	774	509	1,578
	18.69%	49.05%	32.26%	100%

Note. Reading proficiency level was scored 1 (below standards), 2 (near standards), or 3 (above standards).

group, 28.08% moved up from Reading Proficiency Level 2 to 3, whereas in the control group only 24.88% moved up.

Table 3 presents the same statistics as Table 2 but for the writing domain. The patterns are similar to those for reading. Among students at Writing Proficiency Level 2 in the prior year, differences between control and treatment groups were found: 25.49% of the treatment students but only 20.55% of the control students moved up from Writing Proficiency Level 2 to 3. There was also

Table 3. Prior and Current Year Writing Proficiency Level by Treatment Condition.

	Current Year			
Prior Year	Level 1	Level 2	Level 3	Total
Treatment sample				
Level 1	96	130	6	232
	41.38%	56.03%	2.59%	100%
Level 2	88	523	209	820
	10.73%	63.78%	25.49%	100%
Level 3	8	161	461	630
	1.27%	25.56%	73.17%	100%
Total	192	814	676	1,682
	11.41%	48.39%	40.19%	100%
Control sample				
Level 1	96	128	6	230
	41.74%	55.65%	2.61%	100%
Level 2	111	500	158	769
	14.43%	65.02%	20.55%	100%
Level 3	5	168	406	579
	0.86%	29.02%	70.12%	100%
Total	212	796	570	1,578
	13.43%	50.44%	36.12%	100%

Note. Writing proficiency level was scored 1 (below standards), 2 (near standards), or 3 (above standards).

a difference in the rate at which Writing Proficiency Level 2 students fell back to Level 1: 14.43% for control group students but only 10.73% for treatment students.

To further explore the patterns of gains described in the cross tabs in Tables 2 and 3, we estimated regressions to test the significance of those effects while using a full set of controls. Table 4 presents the treatment effects on reading proficiency level for students at Levels 1, 2, and 3 in the preintervention assessment, respectively. All covariates and teacher fixed effects were included. Relative risk ratios are reported. Significant treatment effects were found for students performing at Reading Proficiency Level 2 in the previous year. Specifically, treatment students had a significantly higher probability of improving their reading proficiency to Level 3 (odds ratio [*OR*] = 1.307, p = .020). However, no significant effects were found for students at Reading Proficiency Level 1 or 3 at preintervention, suggesting that receiving the intervention yielded neither positive effects to students who were Level 1 nor negative effects to students who were at Level 3 in reading proficiency.

Table 5 presents the treatment effects on writing proficiency level for students at Levels 1, 2, and 3 in the preintervention assessment. Similarly, significant positive treatment effects were revealed for students at Writing Proficiency Level 2 in the pretreatment assessment. Students who received the intervention were approximately 43% more likely than control students to improve their writing proficiency from Levels 2 to 3 (*OR* = 1.427, p = .011). In addition, treatment students were 29% less likely than control students to exhibit a worsening of their writing proficiency from Level 2 to Level 1 (*OR* = 0.690, p = .007). Similar with the reading proficiency outcomes, the intervention did not manifest significant changes in writing proficiency levels for students at Writing Proficiency Level 1 or Level 3 at the preintervention assessment.

We conclude that the VSTF intervention significantly affected the reading and writing skills of students at the midrange (Level 2) of proficiency—students who were near or at grade level at the end of the prior year. For reading, the intervention increased the upward mobility of these students. For writing, it increased their upward mobility and decreased their downward mobility. However, for students who were Level 1 or Level 3 in the prior year, the effects of VSTF on change across proficiency levels were not significant. Recall that the regression estimates testing for interactions between the treatment and prior achievement levels found no significant interactions. However, the tabular analyses in Tables 3 and 4 found a stronger effect for students whose prior score was at the middle level. This would not be detected from the linear regression analyses because it is, in fact, a *non*linear interaction. Thus, the results are not in conflict with one another. Instead, the extra detail provided by the tabular analyses in Tables 3 and 4 allowed us to see

Table 4. Logistic and Multinomial Logistic Regression Analyses of Treatment Effects by Student Prior Proficiency Levels in Reading.

	Prior Level 1	Prior Level 2		Prior Level 3
	Base Category = Level 1	Base Category = Level 2		Base Category = Level 3
Outcome	Level 2 or Level 3	Level 1	Level 3	Level 1 or Level 2
Treatment	0.985	0.933	1.307*	1.136
	(−0.09)	(−0.54)	(2.32)	(0.87)
Hispanic	1.554	0.989	0.866	1.424
	(1.58)	(−0.04)	(−0.59)	(1.31)
Asian	2.274*	0.284***	1.695*	0.592*
	(2.27)	(−3.69)	(2.40)	(−1.99)
Black	—	1.783	0.356	—
	—	(0.71)	(−0.94)	—
Filipino	4.098	0.560	1.363	0.527
	(0.69)	(−0.67)	(0.60)	(−0.78)
Male	0.595**	1.161	1.273	0.875
	(−3.06)	(0.93)	(1.80)	(−0.78)
Socioeconomically disadvantaged	0.927	1.098	0.546***	0.955
	(−0.26)	(0.33)	(−3.37)	(−0.26)
English learner	0.478***	3.117***	0.470***	1.613
	(−3.91)	(6.49)	(−3.79)	(1.00)
Gifted education	—	0.000***	4.395***	0.319***
	—	(−36.41)	(4.36)	(−5.15)
Special education	0.718	2.369*	0.348*	1.168
	(−1.10)	(2.47)	(−2.28)	(0.23)
Observations	689	1,637		902

Note. Exponentiated coefficients; *t* statistics are in parentheses. Teacher fixed effects are included in all models. Covariates included are identical with Table 5. Reading proficiency level was scored 1 (below standards), 2 (near standards), or 3 (above standards). In model with sample of Prior Level 1 students, Black and GATE were omitted due to collinearity. In model with sample of Prior Level 3 students, Black was omitted due to collinearity.
**p* < .05. ***p* < .01. ****p* < .001.

Table 5. Logistic and Multinomial Logistic Regression Analyses of Treatment Effects by Student Prior Proficiency Levels in Writing.

	Prior Level 1	Prior Level 2		Prior Level 3
	Base Category = Level 1	Base Category = Level 2		Base Category = Level 3
Outcome	Level 2 or Level 3	Level 1	Level 3	Level 1 or Level 2
Treatment	1.148	0.690**	1.427*	0.874
	(0.66)	(−2.72)	(2.55)	(−0.79)
Hispanic	2.319	1.786	0.997	1.042
	(1.55)	(1.85)	(−0.01)	(0.18)
Asian	4.427**	0.397*	2.568***	0.357***
	(2.65)	(−2.53)	(3.86)	(−4.32)
Black	2.883	2.406	0.911	1.007
	(0.65)	(0.98)	(−0.13)	(0.01)
Filipino	—	0.452	1.388	0.536
	—	(−0.82)	(0.57)	(−0.96)
Male	0.459**	1.655**	0.954	1.485**
	(−3.06)	(2.89)	(−0.37)	(2.66)
Socioeconomically disadvantaged	0.396	0.789	0.513***	1.312
	(−1.90)	(−0.96)	(−3.30)	(1.58)
English learner	0.500*	2.120***	0.311***	2.508**
	(−2.31)	(4.08)	(−6.25)	(2.77)
Gifted education	—	0.000***	2.902*	0.225***
	—	(−41.43)	(2.44)	(−6.44)
Special education	1.113	1.572	0.807	1.282
	(0.27)	(1.42)	(−0.58)	(0.42)
Observations	449	1,584		1,203

Note. Exponentiated coefficients; *t* statistics are in parentheses. Teacher fixed effects are included in all models. Covariates included are identical with Table 5. Writing proficiency level was scored 1 (below standards), 2 (near standards), or 3 (above standards). In model with sample of Prior Level 1 students, Filipino and GATE were omitted due to collinearity.
**p* < .05. ***p* < .01. ****p* < .001.

that the treatment effects were largely confined to students who demonstrated near or at-grade level ability in reading and writing on the annual state assessment at the end of the prior year. Similarly, our subgroup analyses showed no effects on students classified as gifted or qualified for special education and suggestive, but not statistically significant, results for English learners.

Discussion

The main goal of our study was to explore whether an intervention that modifies the presentation, but not the content, of text would support middle school students' literacy skills. Our design was particularly rigorous, as the randomization within teachers ensured that all students experienced the benefits of teachers who had received the professional development, access to technology, and the same instructional materials. The sole difference between treatment and control classrooms was the use of VSTF to present texts and instructional materials. Overall, we found that students who read texts formatted in VSTF for fewer than 50 min a week showed small yet significant gains in their overall ELA achievement and in their writing skills. Given that empirical benchmarks suggest that normative growth in the middle school grades is slower than the elementary school grades, these effect sizes are of practical value (relative to years of expected growth; see Bloom, Hill, Black, & Lipsey, 2008; Hill, Bloom, Black, & Lipsey, 2008). What is particularly noteworthy about these gains is that they were found on a fairly broad standardized assessment (see Hill et al., 2008, noting mean effect sizes decreasing from .44 on a specialized test, to .23 on a narrow standardized test, to .07 or below on a broad standardized test). In addition, both treatment and control students completed the Smarter Balanced assessment using traditional block texting, suggesting that the benefits of VSTF were not limited to supporting online processing of challenging texts, but also led to changes in the word-to-text integration processes underlying reading comprehension (Perfetti & Stafura, 2014). Although the design of our study makes us unable to determine which or how many components of word-to-text integration processes (e.g., sentence representations, text models, or situation models) were enhanced through ongoing use of VSTF, the finding that the effects of using VSTF transferred to unformatted texts suggests that long-term use of this formatting improved students' skill at processing and comprehending challenging texts.

The intervention's larger effect on writing compared to reading scores is consistent with prior studies of VSTF (Park, Warschauer et al., 2013; Park & Warschauer, 2016) and a recent meta-analysis of the impact of reading interventions on writing (Graham et al., 2018). Reading instruction designed to increase students' knowledge about the functions and purposes of text gives students important metaknowledge about written language (Graham et al., 2018). In addition, we note that the writing subtest of the Smarter Balanced assessment required students to write an essay using multiples texts as sources of evidence. Consequently, success on the writing subtest was, in part, dependent on heavy reading comprehension demands. In contrast, the prompts and required responses for the reading subtest were less complex, shorter, and generally limited to a single text. Thus, the writing assessment may have had greater sensitivity to reveal the gains in comprehension processes and academic language.

Finally, we wished to determine if the use of VSTF would yield differential effects for students based on subgroup characteristics, such as students who are gifted, students with disabilities, and students who are English learners. Whereas our fixed effect regression analyses suggested that treatment effects did not interact with gifted, special education, or English learner classification, the subgroup and categorical analyses suggested that VSTF did produce differential effects. Although on the surface these may appear to be contradictory, our fixed effects regression assumed linear relationships, whereas the subgroup and categorical analyses did not. Indeed, the interaction uncovered by the categorical analysis is nonlinear, indicating that the treatment effect occurs only for students whose prior performance was at the middle level—at or near grade level, which may account for why the linear interaction was not significant.

There are a variety of reasons why the effects of VSTF were limited to students whose baseline was at the middle range, or near grade level at the end of the prior school year. First, almost twice as many students had middle-level baseline scores as students whose baseline scores were at the highest and lowest proficiency levels. Thus, the power to detect change may have been sufficient for midlevel students but insufficient for students at the top and bottom levels of performance. Second, the 3-point scale of the reading and writing subtests may have limited sensitivity and produced ceiling effects for the students whose baseline scores were at the highest level. Other reasons may reflect potential mismatches between the intervention and the students. For example, students with the highest baseline scores, and those identified as gifted, may not have been sufficiently challenged by the reading level or complexity of the texts and, thus, did not need the scaffolding provided by VSTF. Students below grade level may need support with more fundamental reading skills, such as work identification, which were not supported by the intervention. The limited effects of VSTF for students with disabilities may reflect the heterogeneity of the population covered by the special education designation (any student with an individualized education plan, ranging from physical disabilities to mental disabilities of all degress), and thus we identify this as an area for further study. In contrast, there was suggestive evidence that English learners, who tend to struggle with comprehension rather than word identification processes (Nakamoto, Lindsey, & Manis, 2007), may benefit from VSTF scaffolding. Whereas English learners did not show a benefit from VSTF on their overall ELA or reading scores, they did show an effect on writing scores that, although not statistically significant at the $p < .05$ level, was of a similar magnitude to the overall writing results. Thus, although VSTF may hold promise as an intervention for English learners, replication studies are needed to test this.

Limitations and future research

This study was conducted only in general education ELA classes. Because special education classes, resource rooms, and basic English language development classes were excluded, our findings cannot be generalized to students who lacked the prerequisite skills to participate in the general middle school curriculum.

Unique factors complicated implementation of this study. Although teachers used the same lessons for their treatment and control classes, preparing these lessons with VSTF and unformatted texts presented an additional burden to teachers. This was exacerbated by the introduction of a new ELA curriculum, including both a new digital textbook and pacing guide, which was being developed well into the first semester of the school year. Finally, most of the devices were not provided to teachers prior to the intervention; teachers and students typically become more adept with new technology in the second and subsequent years of use (Means & Olson, 1995; Sandholtz & Reilly, 2004). For all these reasons, we would expect the benefits of VSTF use to expand if teachers continue to use it in instruction beyond the constraints of the study.

Our achievement measures were less sensitive than we would have liked (with scores of 1, 2, or 3 for reading and writing) and given before the end of the school year. We had planned to use district assessments for more proximal measures over each quarter but found the measures were not aligned to the new curriculum, unlike the end-of-the-year state assessment. Future studies with a norm-referenced reading comprehension and writing component at the beginning and end of the intervention would provide more fine-grained detail on changes due to the intervention.

Conclusions

Despite these limitations, VSTF shows promise as a means of supporting the word-to-text integration processes critical for reading comprehension, particularly among typically achieving middle-school students. Use of VSTF provided a modest benefit to middle school students who were near or at grade level in the prior school year, making it appropriate for use as a tool in general education classrooms, particularly if optional.

Acknowledgments

We thank the participating educators and students who made this research possible and Randy Walker, MD who created LiveInk, the VSTF used in this intervention (publicly available at LiveInk.com).

Disclosure statement

There is no potential conflict of interest reported by any of the authors.

Funding

This research was supported by a grant from the US Department of Education, Institute of Education Sciences Award R305A150429, PI: Warschauer, Co-PIs: Collins & Farkas, "Digital Scaffolding for English Language Arts."

ORCID

Tamara P. Tate http://orcid.org/0000-0002-1753-8435

References

*References marked with asterisks are present only in supplementary material.

*Aaltonen, M., Mannonen, P., Nieminen, S., & Nieminen, M. (2011). Usability and compatibility of e-book readers in an academic environment: A collaborative study. *IFLA Journal*, *37*(1), 16–27. doi:10.1177/0340035210396775

*Applebee, A. N., & Langer, J. A. (2011). A snapshot of writing instruction in middle schools and high schools. *The English Journal*, *100*(6), 14–27.

*Armitage, U., Wilson, S., & Sharp, H. (2004). Navigation and ownership for learning in electronic texts: An experimental study. *Electronic Journal of E-Learning*, *2*(1), 19–30.

Barrouillet, P., Gavens, N., Vergauwe, E., Gaillard, V., & Camos, V. (2009). Working memory span development: A time-based resource-sharing model account. *Developmental Psychology*, *45*, 477–490. doi:10.1037/a0014615

Berninger, V., Cartwright, A., Yates, C., Swanson, H. L., & Abbott, R. D. (1994). Developmental skills related to writing and reading acquisition in the intermediate grades: Shared and unique functional systems. *Reading and Writing*, *6*, 161–196. doi:10.1007/BF01026911

Bever, T. G., Jandreau, S., Burwell, R., Kaplan, R., & Zaenen, A. (1990). Spacing printed text to isolate major phrases improves readability. *Visible Language*, *25*, 74–87.

Biancarosa, G., & Snow, C. E. (2004). *Reading next—A vision for action and research in middle and high school literacy: A report from Carnegie Corporation of New York*. Washington, DC: Alliance for Excellent Education.

Bloom, H., Hill, C., Black, A., & Lipsey, M. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Research on Educational Effectiveness*, *1*(4), 289–328. doi:10.1080/19345740802400072

Bunch, G. C., Walqui, A., & Pearson, P. D. (2014). Complex text and new common standards in the United States: Pedagogical implications for English learners. *TESOL Quarterly*, *48*, 533–559. doi:10.1002/tesq.175

Cain, K. (2007). Syntactic awareness and reading ability: Is there any evidence for a special relationship? *Applied Psycholinguistics*, *28*, 679–694. doi:10.1017/S0142716407070361

California Department of Education (n.d.). Retrieved on September 7, 2018 from https://caaspp.cde.ca.gov/sb2017/ViewReport?ps=true&lstTestYear=&lstTestType=B&lstCounty=00&lstDistrict=00000&lstSchool=0000000

*Collins, P., Hwang, J. K., Zheng, B., & Warschauer, M. (2014). Writing with laptops: A quasi-experimental study. *Writing & Pedagogy*, *5*(2), 203–230. doi:10.1558/wap.v5i2.203

Common Core State Standards Initiative, (CCSSI 2017). California common core state standards. Retrieved October 07, 2017, from https://www.cde.ca.gov/be/st/ss/documents/finalelaccssstandards.pdf

Crossley, S. A., Louwerse, M. M., McCarthy, P. M., & McNamara, D. S. (2007). A linguistic analysis of simplified and authentic texts. *The Modern Language Journal*, *91*, 15–30. doi:10.1111/j.1540-4781.2007.00507.x

*Evans, M. A., Charland, A. R., & Saint-Aubin, J. (2009). A new look at an old format: Eye-tracking studies of shared book reading and implications for eBook and eBook research. *Multimedia and Literacy Development*, 89–111.

Faggella-Luby, M. N., & Deshler, D. D. (2008). Reading comprehension in adolescents with LD: What We Know; What We Need to Learn. *Learning Disabilities Research & Practice*, *23*, 70–78. doi:10.1111/j.1540-5826.2008.00265.x

Fang, Z., Schleppegrell, M. J., & Cox, B. E. (2006). Understanding the language demands of schooling: Nouns in academic registers. *Journal of Literacy Research*, *38*, 247–273. doi:10.1207/s15548430jlr3803_1

Fitzgerald, J., & Shanahan, T. (2000). Reading and writing relations and their development. *Educational Psychology*, *93*, 3–22. doi:10.1037/0022-0663.92.1.3

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, *5*(3), 239–256. doi:10.1207/S1532799XSSR0503_3

Graham, S., Liu, X., Bartlett, B., Ng, C., Harris, K., Aitken, A., … Talukdar, J. (2018). Reading for writing: A meta-analysis of the impact of reading interventions on writing. *Review of Educational Research*, *88*(2), 243–284. doi:10.3102/0034654317746927

*Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology*, *99*(3), 445–476. doi:10.1037/0022-0663.99.3.445

Hill, C., Bloom, H., Black, A., & Lipsey, M. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives*, *2*(3), 172–177. doi:10.1111/cdep.2008.2.issue-3

Hirotani, M., Frazier, L., & Rayner, K. (2006). Punctuation and intonation effects on clause and sentence wrap-up: Evidence from eye movements. *Journal of Memory and Language*, *54*, 425–443. doi:10.1016/j.jml.2005.12.001

Jandreau, S., & Bever, T. G. (1992). Phrase-spaced formats improve comprehension in average readers. *Journal of Applied Psychology*, *77*(2), 143–146. doi:10.1037/0021-9010.77.2.143

Johnson, R. L., Bui, B., & Schmitt, L. L. (2018). Are two spaces better than one? The effect of spacing following periods and commas during reading. *Attention, Perception & Psychophysics*, *80*, 1504–1511. doi:10.3758/s13414-018-1527-6

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*, 329–354. doi:10.1037/0033-295X.87.4.329

Kintsch, W. (2005). An overview of top-down and bottom-up effects in comprehension: The CI perspective. *Discourse Processes*, *39*(2–3), 125–128. doi:10.1080/0163853X.2005.9651676

Lee, S. (2007). Effects of textual enhancement and topic familiarity on Korean EFL students' reading comprehension and learning of passive form. *Language Learning*, *57*(1), 87–118. doi:10.1111/lang.2007.57.issue-1

Lee, S. K., & Huang, H. T. (2008). Visual input enhancement and grammar learning: A metaanalytic review. *Studies in Second Language Acquisition*, *30*, 307–331. doi:10.1017/S0272263108080479

Leow, R. (1997). The effects of input enhancement and text length on adult L2 readers' comprehension and intake in second language acquisition. *Applied Language Learning*, *82*, 151–182.

Lipka, O., & Siegel, L. S. (2012). The development of reading comprehension skills in children learning English as a second language. *Reading & Writing*, *25*, 1873–1898. doi:10.1007/s11145-011-9309-8

*Mangen, A., Walgermo, B. R., & Brønnick, K. (2013). Reading linear texts on paper versus computer screen: Effects on reading comprehension. *International Journal of Educational Research*, *58*, 61–68. doi:10.1016/j.ijer.2012.12.002

*Margolin, S. J., Driscoll, C., Toland, M. J., & Kegler, J. L. (2013). E-readers, computer screens, or paper: Does reading comprehension change across media platforms? *Applied Cognitive Psychology*, *27*, 512–519. doi:10.1002/acp.2930

McNeish, D., Stapleton, L. M., & Silverman, R. D. (2017). On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods*, *22*(1), 114. doi:10.1037/met0000078

Means, B., & Olson, K. (1995). *Technology's role in education reform: Findings from a National Study of Innovating Schools*. Washington, DC: Office of Educational Research and Improvement, US Department of Education. Retrieved on September 5, 2018 from, https://www2.ed.gov/PDFDocs/techrole.pdf

Mokhtari, K., & Thompson, H. B. (2006). How problems of reading fluency and comprehension are related to difficulties in syntactic awareness skills among fifth graders. *Reading Research and Instruction*, *46*, 73–94. doi:10.1080/19388070609558461

Nakamoto, J., Lindsey, K. A., & Manis, F. R. (2007). A longitudinal analysis of English language learners' word decoding and reading comprehension. *Reading and Writing*, *20*, 691–719. doi:10.1007/s11145-006-9045-7

Oh, S. Y. (2001). Two types of input modification and EFL reading comprehension: Simplification versus elaboration. *TESOL Quarterly*, *35*(1), 69–96. doi:10.2307/3587860

Olson, C. B., Kim, J. S., Scarcella, R., Kramer, J., Pearson, M., van Dyk, D. A., … Land, R. E. (2012). Enhancing the interpretive reading and analytical writing of mainstreamed English learners in secondary school: Results from a randomized field trial using a cognitive strategies approach. *American Educational Research Journal*, *49*(2), 323–355. doi:10.3102/0002831212439434

Overstreet, M. (1998). Text enhancement and content familiarity: The focus of learner attention. *Spanish Applied Linguistics*, *2*, 229–258.

Park, Y., & Warschauer, M. (2016). Syntactic enhancement and second language literacy: An experimental study. *Language Learning & Technology*, *20*(3), 180–199.

Park, Y., Warschauer, M., Collins, P., Hwang, J. K., & Vogel, C. (2013). Building awareness of language structures with visual-syntactic text formatting. In K. E. Pytash & R. E. Ferdig (Eds.), *Exploring technology for writing and writing instruction* (pp. 640–656). Hershey, PA: IGI Global.

Park, Y., Xu, Y., Collins, P., Farkas, G., & Warschauer, M. (2018). Scaffolding learning of language structures with visual-syntactic text formatting. *British Journal of Educational Technology*. doi:10.1111/bjet.12689

Park, Y., Zheng, B., Lawrence, J., & Warschauer, M. (2012). Technology-enhanced reading environments. In M. Thomas, H. Reinders, & M. Warschauer (Eds.), *Contemporary computer-assisted language learning*. London & New York: Continuum.

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*(1), 22–37. doi:10.1080/10888438.2013.827687

Porter, A., McMaken, J., Hwang, J., & Yang, R. (2011). Common core standards: The new U.S. intended curriculum. *Educational Researcher*, *40*, 103–116. doi:10.3102/0013189X11405038

Sandholtz, J. H., & Reilly, B. (2004). Teachers, not technicians: Rethinking technical expectations for teachers. *Teachers College Record*, *106*(3), 487–512. doi:10.1111/tcre.2004.106.issue-3

Schwan, S., Garsoffky, B., & Hesse, F. (2000). Do film cuts facilitate the perceptual and cognitive organization of activity sequences? *Memory & Cognition*, *28*, 214–223. doi:10.3758/bf03213801

Shanahan, T. (2006). Relations among oral language, reading, and writing development. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research*. New York, NY: Guilford.

Shook, D. (1994). FL/L2 reading, grammatical information, and the input-to-intake phenomenon. *Applied Language Learning*, *52*, 57–93.

Simard, D. (2009). Differential effects of textual enhancement formats on intake. *System*, *37*, 124–135. doi:10.1016/j.system.2008.06.005

Stine-Morrow, E. A., Shake, M. C., Miles, J. R., Lee, K., Gao, X., & McConkie, G. (2010). Pay now or pay later: Aging and the role of boundary salience in self-regulation of conceptual integration in sentence processing. *Psychology and Aging*, *25*, 168–176. doi:10.1037/a0018127

Tiffin-Richards, S. P., & Schroeder, S. (2018). The development of wrap-up processes in text reading: A study of children's eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(7), 1051–1063. doi:10.1037/xlm0000506

Torgesen, J. K., Houston, D., Rissman, L., Decker, S., Roberts, G., Vaughn, S., … Lesaux, N. (2007). *Academic literacy instruction for adolescents: A guidance document from the Center on Instruction*. Portsmouth, NH: RMC Research Corporation, Center on Instruction. https://www.carnegie.org/media/filer_public/a7/9b/a79bee13-b82e-47bd-ab63-9190baa31975/ccny_report_2007_guidance.pdf

Veenendaal, N. J., Groen, M. A., & Verhoeven, L. (2014). What oral text reading fluency can reveal about reading comprehension. *Journal of Research in Reading*. doi:10.1111/1467-9817.12024

Walker, R., & Vogel, C. (2005). VSTF: Brain-based text formatting raises standardized test scores. Paper presented at the National Educational Computing Conference, Philadelphia.

Walker, R. C., Gordon, A. S., Schloss, P., Fletcher, C. R., Vogel, C., & Walker, S. (2007). Visual-syntactic text formatting: Theoretical basis and empirical evidence for impact on human reading. Paper presented at the IEEE Professional Communication Conference, 2007, Seattle. 1–14. Retrieved from http://ieeexplore.ieee.org

Walker, S., Schloss, P., Fletcher, C. R., Vogel, C. A., & Walker, R. C. (2005, May/June). Visual-syntactic text Formatting: A new method to enhance online reading. *Reading Online*, *8*(6).

*Wertsch, J. V. (1991). *Voices of the mind: A sociocultural approach to mediated action*. Cambridge, MA: Harvard University Press.

What Works Clearinghouse (n.d.). *Assessing attrition bias*. Institute of Education Sciences. Retrieved on September 7, 2018 from https://ies.ed.gov/ncee/wwc/Document/243

Williams, G. L., Fitzgerald, J., & Stenner, A. J. (2013). The common core state standards' quantitative text complexity trajectory: Figuring out how much complexity is enough. *Educational Researcher*, *42*(2), 59–69. doi:10.3102/0013189X12466695

Wouters, P., Paas, F., & van Merriënboer, J. J. (2008). How to optimize learning from animated models: A review of guidelines based on cognitive load. *Review of Educational Research*, *78*(3), 645–675. doi:10.3102/0034654308320320