


Examining the Effects of a Tier 2 Reading Comprehension Intervention Aligned to Tier 1 Instruction for Fourth-Grade Struggling Readers

Exceptional Children
2020, Vol. 86(4) 430–448
© The Author(s) 2020
DOI: 10.1177/0014402919893710
journals.sagepub.com/home/ecx


Elizabeth A. Stevens¹ , Sharon Vaughn²,
Elizabeth Swanson², and Nancy Scammacca²

Abstract

This article presents findings from a quasiexperimental study examining the effects of a Tier 2 intervention aligned to Tier 1 instruction, a nonaligned Tier 2 intervention, and a business-as-usual (BAU) comparison on the content knowledge, vocabulary, and reading outcomes of fourth-grade struggling readers. In the aligned condition, teachers were trained to provide content-area reading practices during social studies, and struggling readers from these classes received small-group intervention aligned to those practices, allowing for additional practice opportunities with feedback. Struggling readers in the nonaligned condition received the same small-group intervention, though they were not provided the comprehension practices during their Tier 1 social studies instruction. Students in the BAU received typical social studies instruction and typical intervention. Results yielded statistically significant, positive effects in favor of the aligned condition on proximal measures of content knowledge and vocabulary but no significant differences on standardized measures of reading comprehension and vocabulary.

Improving reading comprehension has been a somewhat elusive goal for students with reading difficulties. In 2010, the Institute of Education Sciences (IES) issued a request for proposals to improve reading for understanding (RFU), investing \$100 million in the development of reading comprehension assessments and interventions for students from preK through 12th grade. This initiative engaged over 130 scholars across numerous institutions in a cooperative path to improve what we know and understand about reading comprehension. The goal of the IES RFU initiative was to “aggressively attack and derive solutions for enabling students to understand what they read” (IES, n.d.). This work yielded significant scientific findings that are reported in numerous scholarly outlets; however, the findings from the comprehension intervention studies yielded generally small-to-negligible effects

(e.g., Connor et al., 2018; Goldman et al., 2016; Lonigan et al., 2018; Scammacca et al., 2016; Vaughn et al., 2013; Wanzek et al., 2013).

There are many possible explanations for these unimpressive effects, including system-level barriers as well as school-level and classroom-level challenges that are offered to explain small effects on reading comprehension. One factor that the research teams reported that they observed consistently was incoherent application of comprehension practices within and across classroom instruction (Swanson et al.,

¹Georgia State University

²University of Texas at Austin

Corresponding Author:

Elizabeth A. Stevens, Department of Learning Sciences,
College of Education and Human Development, Georgia
State University, Suite 750, 30 Pryor St. SW, Atlanta,
GA 30303.

Email: estevens11@gsu.edu

2016). For example, across social studies and English language arts classrooms in Grades 7 through 12, teachers (a) rarely provided high-quality vocabulary instruction that supports vocabulary acquisition and reading comprehension; (b) frequently built background knowledge and previewed text prior to reading, though this instruction was of low quality and consumed the majority of the class time; and (c) rarely engaged students in text reading, thus limiting the opportunities for explicit comprehension instruction and text-based discussions in both content areas. We wondered what the additional influence would be if we aligned instructional practices between core instruction and intervention so that students with reading difficulties had a systematic and coherent approach to applying reading comprehension practices across instructional settings.

Many students with reading difficulties are provided with supplemental (Tier 2 type) reading intervention for approximately 30 min a day (Foorman et al., 2016; Kamil et al., 2008). We were interested in the effect of aligning this supplemental reading intervention with the Tier 1 content-area reading instruction provided in the classroom. Aligning Tier 1 and Tier 2 instruction allowed us to examine the effect of additional practice with feedback in a small-group Tier 2 intervention setting on students' content acquisition and use of reading comprehension and vocabulary practices. There are few studies at the primary grades examining aligned Tier 1 and Tier 2 approaches (e.g., Fien et al., 2015; Scanlon et al., 2008). For example, Fien and colleagues (2015) implemented an efficacy study aligning Tier 1 and Tier 2 reading practices for first-grade students at risk for reading difficulty, yielding statistically significant, positive outcomes on nonsense word reading and oral reading fluency. Another study examined the effects of (a) professional development (PD) in the interactive strategies approach for Tier 1 teachers, (b) small-group Tier 2 reading intervention using the Interactive Strategies Approach, and (c) the combination of Tier 1 PD and Tier 2 reading intervention on the prevention of early reading difficulties for at-risk kindergarteners (Scanlon et al., 2008). However, we were unable to identify efficacy studies that exam-

ined the coordinated implementation of comprehension and vocabulary practices in Tier 1 and Tier 2 instruction specifically for struggling readers in Grades 4 through 12.

We wondered what the additional influence would be if we aligned instructional practices between core instruction and intervention so that students with reading difficulties had a systematic and coherent approach to applying reading comprehension practices across instructional settings.

Researchers have acknowledged the challenges to successfully implementing reading comprehension instruction during English language arts classes (Wanzek et al., 2013). Providing these instructional practices solely within English language arts classes does not adequately provide students with a model for using effective practices across all reading opportunities. In addition, there is an increased focus in progressive state standards (Common Core State Standards; National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010) on teaching reading comprehension using informational text, and social studies provides a prime opportunity for integrating reading comprehension and vocabulary during informational-text reading. Therefore, we targeted the integration of reading and vocabulary instructional practices within content-area classes, such as social studies and history (Swanson et al., 2019).

Instructional Practices

In response to the need to increase the use of evidence-based vocabulary and comprehension practices using informational text in upper-elementary classrooms, we utilized STRIVE (Strategies for Reading Information and Vocabulary Effectively; Simmons et al., 2010). STRIVE features a set of evidence-based vocabulary and reading comprehension instructional practices delivered by classroom teachers within fourth-grade social studies classes. The PD for implementing the instructional practices was co-designed with upper-elementary school

teachers and a research team (Simmons et al., 2010). STRIVE features key components (Desimone, 2009; Gersten et al., 2010) to build teacher knowledge and improve students' vocabulary and comprehension within content learning of social studies. STRIVE PD begins with a workshop led by researchers that focuses on building teachers' knowledge of a set of instructional practices aimed at improving vocabulary and comprehension within social studies texts. Follow-up PD opportunities include teacher study teams that allow teachers to practice new skills, address areas of need in their delivery of STRIVE, and identify ways in which STRIVE may address student needs.

The PD for implementing the instructional practices was co-designed with upper-elementary school teachers and a research team.

This distributed PD model features a set of integrated STRIVE instructional practices that align with the social studies texts (see Swanson et al., 2019, for more detail). Teachers use vocabulary semantic maps addressing key content words to facilitate discussion of the relationships among words (Kamil et al., 2008) and the use of context clues to facilitate students' independent word learning. Lessons include comprehension practices that build background knowledge prior to reading, teach students to identify the main idea of short sections of text during reading (e.g., Klingner et al., 1998), and combine main-idea statements to generate paragraph-long summaries after reading. Teachers also lead students in text-based discussion using a variety of question types, and students learn how to use text evidence to support their answers. Teachers in the aligned Tier 1 and Tier 2 (T1-T2) condition implemented these STRIVE practices.

As described in more detail in the Method section, the research team hired and trained tutors to provide Tier 2 intensified STRIVE to groups of five or six students with reading difficulties with the goal of providing structured additional practice with feedback for students. Students received the same vocabulary and comprehension instructional practices in the

Tier 2 intensified STRIVE sessions as during the Tier 1 social studies, with several adaptations designed to increase the intensity of the instruction (detailed in the Method section).

Purpose of the Study

This study examined the relative effects of a set of vocabulary and comprehension practices—aimed at improving reading comprehension—provided to students during T1-T2 instruction, Tier 2 alone, or business as usual (BAU). We implemented a three-arm quasiexperimental study that included the following conditions: (a) Tier 1 instruction with an aligned Tier 2 intervention (i.e., STRIVE Tier 1 + intensified STRIVE Tier 2), (b) a non-aligned Tier 2 intervention (i.e., BAU Tier 1 + intensified STRIVE Tier 2), and (c) BAU (i.e., BAU Tier 1 instruction and BAU Tier 2 intervention). These three conditions allowed us to address the following research questions for a sample of fourth-grade struggling readers:

1. What is the effect of Tier 1 content-area reading instruction in social studies combined with a highly aligned Tier 2 reading intervention compared to a BAU condition on students' vocabulary, content knowledge, and reading comprehension?
2. What is the effect of a Tier 2 reading intervention that is *not* aligned to Tier 1 content-area instruction compared with a BAU condition on students' vocabulary, content knowledge, and reading comprehension?

Method

Research Design

Twelve schools from one school district in a southwestern state participated in the study. We utilized a quasiexperimental design, assigning 13 schools to one of three conditions described previously. This design allowed us to test the efficacy of the intervention aligned with Tier 1, the intervention without alignment to Tier 1, and the third condition of neither Tier 1 nor Tier 2 instruction as specified by the research team.

The pool of potential schools participated in a larger randomized control trial (RCT) involving multiple, nonoverlapping cohorts; in that study, schools from the same district were randomly assigned to receive PD on implementing STRIVE Tier 1 literacy practices during social studies instruction or to deliver BAU social studies instruction (Swanson et al., 2019). From that pool, we identified five schools in the district whose teachers previously received training in STRIVE Tier 1 practices (i.e., as part of a previous cohort that received STRIVE PD from the larger efficacy study; Swanson et al., 2019) and assigned those schools to the aligned T1-T2 condition, meaning the teachers provided Tier 1 STRIVE instruction and the students received the aligned, intensified STRIVE Tier 2 intervention from trained tutors. To strengthen the quality of the quasiexperimental research design, we implemented propensity score matching prospectively using sampling without replacement, a tolerance of .50, and matching on the following school variables: total enrollment; 2017 state accountability rating; attendance rate; percentage enrollment of White, economically disadvantaged, English language learner, and special education students; mobility rate; average fourth-grade class size; total operating expenditures per student; percentage of students who met passing criterion on the state assessment; and school performance index ratings for student achievement, student progress, closing the performance gap, and postsecondary readiness. Using the obtained propensity scores, we matched the aligned T1-T2 schools with schools in the same district (i.e., these schools participated in the BAU condition in the previously conducted larger RCT, so teachers had not been exposed to the Tier 1 STRIVE practices) with the closest propensity scores. We assigned these schools to the nonaligned Tier 2 condition, meaning teachers provided Tier 1 instruction and the students received the same intensified STRIVE Tier 2 intervention used in the aligned T1-T2 condition. Finally, we assigned three schools to the BAU condition, meaning students received typical Tier 1 social studies instruction and typical Tier 2 intervention. We selected these schools because they also served as the BAU sites in the final cohort of the larger efficacy trial that occurred simultaneously (Swan-

son et al., 2019). One school withdrew from the aligned T1-T2 condition prior to the start of the study because the school transitioned to new leadership and was unable to proceed with the project. In order to maintain the original covariates used in the propensity score matching, we did not replace this school and proceeded with the four schools in this condition.

Participants and Setting

All fourth-grade teachers ($N = 36$; Table 1) from the 12 participating schools provided consent to participate in the study and received an honorarium for their participation. We identified struggling readers who did not pass the previous year's high-stakes reading assessment (i.e., end of third grade). We included students with attention difficulties, those with dyslexia, and those receiving special education services as long as those students participated in social studies instruction in general education. Students who did not pass the reading assessment yielded a potential pool of 305 participants; of those students, 195 returned parental consent providing permission to participate in the study. We identified up to 12 participants at each school for the aligned T1-T2 condition ($n = 4$ schools), up to 10 participants at each school for the nonaligned Tier 2 condition ($n = 5$ schools), and up to 17 students at each school for the BAU condition ($n = 3$ schools). We limited the number of participating students to 150 (i.e., 50 per condition) because this number represented the largest sample our resources would allow us to serve. If there were more eligible students than spots available at a given school, we randomly selected participants from the list of consented students. At the start of the study, there were 48 students in the aligned T1-T2 condition, 49 students in the nonaligned Tier 2 condition, and 50 in the BAU condition. Three students left the nonaligned Tier 2 condition and eight left the BAU condition due to withdrawing from their respective schools; no students left the aligned T1-T2 condition. Table 2 presents the demographic characteristics for the final sample. There were no statistically significant differences between the conditions on race or ethnicity, free or reduced-price lunch, limited-English-proficient status, special education status, or dyslexia and 504 status.

Table 1. Teacher Demographic Information.

Characteristic	Aligned T1-T2 (n = 11)		Nonaligned T2 (n = 15)		BAU (n = 10)	
	n or M	%	n or M	%	n or M	%
Gender						
Female	11	100.0	14	93.3	9	90.0
Male			1	6.7	1	10.0
Mean years teaching	10.83		9.60		13.80	
Certification						
Elementary	11	100.0	15	100.0	10	100.0
Secondary	1	9.1	2	13.3	2	20.0
ESL	0		1	6.7	1	10.0
Special education	0		1	6.7	0	
Bilingual	4	36.4	3	20.0	2	20.0
Degree						
Bachelor's	6	54.5	9	60.0	7	70.0
Master's	5	45.5	6	40.0	3	30.0

Note. T1 = Tier 1; T2 = Tier 2; BAU = business as usual; ESL = English for speakers of other languages.

Procedures

The Tier 1 comprehension instruction occurred during social studies classes. All schools in all conditions utilized the state-adopted social studies curriculum. This included using the same fourth-grade progressive state standards and the same state-developed timeline for delivering social studies. Subsequently, we describe what occurred in each condition across the tiers of instruction (see Figure 1).

Tier 1 instruction in each condition. In the non-aligned Tier 2 and BAU conditions, teachers provided typical social studies lesson throughout the course of the project; teachers did not receive any PD or lesson support for Tier 1 social studies instruction. In the aligned T1-T2 condition, teachers implemented STRIVE (Swanson et al., 2019), a set of Tier 1 content-area comprehension practices across three 6-week units of social studies instruction. Previous research found the STRIVE practices effective for improving students' reading comprehension and vocabulary knowledge when reading informational texts (Hairrell et al., 2011; Simmons et al., 2010). Each instructional unit consisted of 12, 45-min lessons (i.e., 36

lessons total). Teachers delivered lessons on Monday and Wednesday of each week; the lessons included instructional practices implemented before, during, and after reading informational text. The lessons included reading grade-level informational text that addressed the state-adopted social studies curriculum.

Before-reading practices. Prior to reading text, teachers implemented three instructional practices: (1) building background knowledge, (2) providing explicit vocabulary instruction, and (3) posing the comprehension purpose question. To build background knowledge, teachers stated the big idea of the unit (e.g., "Remember that we're learning about the first civilizations, which included Paleoamericans, American Indians, and European settlers"), connected the new text to prior learning (e.g., "Last time, you learned about the Spanish explorers' journey to America; today, you will learn what happened when they arrived and the hardships they encountered"), and explained important background knowledge using a visual of key content. Teacher-led discussion about the illustrations were designed to prompt students to make connections between prior knowledge and new content.

Table 2. Student Demographic Information.

Characteristic	Aligned T1-T2 (n = 48)		Nonaligned T2 (n = 46)		BAU (n = 42)	
	n or M (SD)	%	n or M (SD)	%	n or M (SD)	%
Gender						
Female	25	52.1	24	49.0	20	40.0
Male	23	47.9	22	44.9	22	44.0
Nonbinary	0		0		0	
Mean age	9.13 (0.33)		9.24 (0.43)		9.14 (0.35)	
Ethnicity						
Native American	0		0		0	
African American	3	6.3	4	8.2	0	
Asian	0		0		0	
Hispanic	45	93.8	42	85.7	41	82.0
Caucasian	0		0		0	
Two or more	0		0		1	2.0
LEP	12	25.0	7	14.3	14	28.0
FRPL	48	100	45	97.8	42	84.0
SPED						
Learning disability	2	4.2	4	8.2	6	12.0
Speech or language impairment	2	4.2	5	10.2	0	
Other health impairment	1	2.1	1	2.0	0	
Intellectual disability	1	2.1	0		2	4.0
Traumatic brain injury	0		0		0	
Emotional disturbance	0		1	2.0	0	
Hearing impairment	0		0		0	
Other	4	8.3	6	12.2	2	4.0
Dyslexia/504	7	14.6	11	22.4	5	11.9
Mean score on 2018 state reading test	1277.68 (83.46)		1255.02 (69.80)		1248.02 (75.20)	

Note. T1 = Tier 1; T2 = Tier 2; BAU = business as usual; LEP = limited-English-proficient designation; FRPL = receiving free or reduced-price lunch; SPED = receiving special education services.

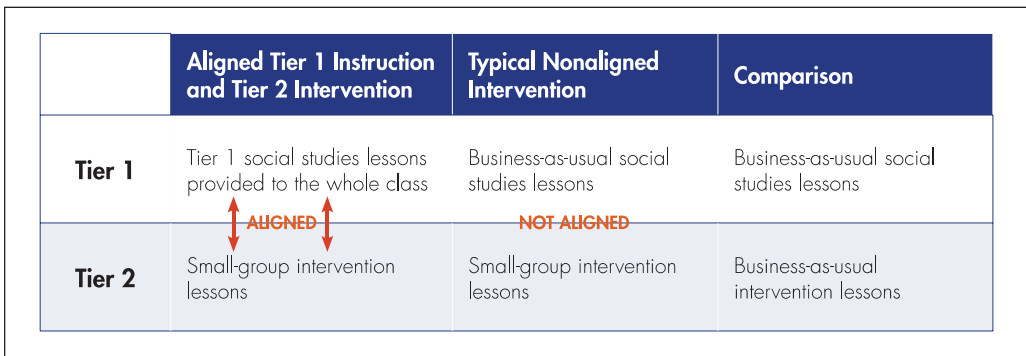


Figure 1. Instruction that occurred in each condition across Tiers 1 and 2.

After building background knowledge, teachers used a researcher-developed semantic map

to provide explicit vocabulary instruction for two high-utility vocabulary words (i.e., words

identified as crucial to understanding the text and used across content domains). Teachers introduced each word using a student-friendly definition, led a discussion about a visual representation of the word (i.e., “How does this photograph help you remember the meaning of the word?”), and provided examples of the word in the appropriate context. Finally, teachers posed a researcher-provided comprehension purpose question prior to reading the text, establishing a purpose for reading and directing students’ attention to the main idea of the passage (e.g., “What contributed to the demise, or end, of the explorers and their expeditions?”). After reading, teachers returned to the comprehension purpose question and led students in a discussion to answer the question.

During-reading practices. Teachers learned three text-reading routines to implement based on the reading proficiency level of their students (i.e., silent reading, partner reading, or teacher read-aloud). While reading lesson passages, teachers stopped to ask questions using various question types (i.e., who, what, when, where, why, and how), which encouraged literal and inferential thinking and prompted student discussion of content. Students learned the answers corresponding to each question type (e.g., a place or a location answers a “where” question; a reason or cause answers a “why” question). Understanding these response types allowed teachers to scaffold student responses. For example, if a student had difficulty answering a “why” question, the teacher reminded the student that a “why” question is answered with a reason or cause and prompted the student to reread the passage to identify a reason or cause. Teachers instructed students to use a paraphrasing process, Get the Gist (Klingner et al., 1998), to identify main ideas for specific sections of text. The process included three steps: (1) Who or what is the passage about? (2) What is the most important idea about the “who” or “what”? and (3) Combine Steps 1 and 2 to write the gist. Teachers modeled the process and guided students through the steps to compose main-idea statements written in the students’ own words for specific subsections of text.

After-reading practices. After reading, students returned to the semantic maps using and applying the words in context, identifying words that were and were not associated with the vocabulary word (i.e., related words), writing a sentence or phrase that appropriately used the term, and engaging in “turn-and-talk” using the word in paired, brief discussions. The turn-and-talk prompted students to apply their understanding of words (e.g., “If you could go on an ‘expedition,’ where would you go and why?” “How would you avoid having ‘debt’?”). In Unit 2, teachers showed students how to use gist statements across subsections of the text to write a summary of the entire passage. Teachers modeled (1) writing an introductory sentence, (2) using previously written gists for the body of the summary, and (3) writing a concluding sentence. Teachers used a graphic organizer to illustrate how the previously generated paragraph-level gists informed the summary.

Tier 1 PD for the aligned T1-T2 condition. Classroom teachers in the aligned T1-T2 condition participated in a distributed PD model preparing them to provide the whole-class Tier 1 social studies aligned with the Tier 2 intervention lessons. We used the distributed PD model used in prior STRIVE studies (Hairrell et al., 2011; Simmons et al., 2010; Swanson et al., 2019). Teachers attended an initial 8-hr PD session led by the research team. The session included an overview of the instructional practices with modeling and practice opportunities for Unit 1 (i.e., the first 6 weeks of instruction) practices. At the conclusion of the training, teachers completed a PD fidelity form evaluating (a) how well researchers addressed each component on a scale of 1 to 3 (i.e., 1 being *not addressed* and 3 being *addressed*), and (b) preparedness to teach the practices on a scale of 1 to 4 (i.e., 1 being *not prepared* and 4 being *prepared*). Teachers reported each component was addressed ($M = 3.0$), and they were prepared to teach each practice (M range = 3.80–3.90). Teachers also attended two follow-up 2-hr sessions prior to Unit 2 and Unit 3 instruction to learn

about the upcoming instructional practices in those units.

Tier 2 intervention in each condition. In the BAU condition, the struggling readers identified at each school received the typical intervention provided by the school (see Tier 2 fidelity results for more information on the interventions provided in the BAU condition). Students did not receive any intervention from the research team. In the aligned T1-T2 condition, identified struggling readers received the small-group reading intervention three times per week (i.e., 54 intervention lessons) provided by tutors hired by the research team. In the nonaligned Tier 2 condition, identified struggling readers received the same small-group reading intervention provided in the aligned T1-T2 condition. The instruction in Tier 1 differed across the two treatment conditions. In the aligned T1-T2 condition, students also received the aligned Tier 1 STRIVE lessons described previously; in the nonaligned Tier 2 condition, students received typical Tier 1 social studies lessons provided by their classroom teachers.

Alignment between Tier 2 intervention and Tier 1 STRIVE instruction. We aligned the Tier 2 intervention to the Tier 1 STRIVE instruction. The intervention included the same instructional practices used in the whole-class STRIVE instruction (i.e., building background knowledge, explicit vocabulary instruction, text reading with questioning, and main-idea generation using Get the Gist); however, the intervention did not address summary writing due to time constraints. The lessons addressed the same social studies topics used in the whole-class STRIVE lessons.

Intensified instruction in Tier 2. We intensified instruction by providing students with intervention in small groups of five or six students, allowing struggling readers to have more opportunities for practice with high-quality feedback. Tutors received training on providing specific feedback (i.e., “What is correct about the student’s response?”). Second, we identified skills taught in the whole-class lessons (e.g., generating main-idea statements)

that posed challenges for students with reading difficulties. Tutors scaffolded the instruction for the students by providing more explicit instruction that included extensive modeling and guided practice. For example, in the whole-class lesson, students received instruction on the six question types (i.e., who, what, when, where, why, and how) in one lesson. In the intervention, students received instruction on each question stem in a separate lesson, beginning with the literal questions (e.g., who, when, where) and increasing in difficulty to the inferential question types (e.g., why and how). As students learned additional question types, tutors interleaved these question stems during reading. For example, students learned “what” and “where” questions in Lesson 1. During reading, tutors asked only “what” and “where” questions and then integrated “when” questions in Lesson 2, “who” questions in Lesson 3, and so on. Third, we began the intervention lessons 1 week ahead of the whole-class STRIVE lessons, providing students with the opportunity to learn the skills and social studies content prior to classroom instruction. We intended for this sequence to frontload the skills and content such that students would be better able to engage in, understand, and keep up with the pace of the Tier 1 instruction. As part of this, we incorporated goal setting as a way to support generalization of students’ skill and transfer across instructional settings. At the end of each lesson, students set a goal aimed toward applying or sharing their learning outside of the intervention group (e.g., to raise their hand to read aloud in class, to explain the meaning of a key vocabulary term to the class, to raise their hand to answer a “why” or “how” question during class). Fourth, we incorporated a more intentional repeated-reading routine to support students with accessing grade-level text and content using evidence-based approaches to repeated reading (Stevens et al., 2017). Teachers modeled fluently reading the text aloud, the students and the teacher chorally read the text, and then students read the text independently. Teachers used this routine for each section of text, asking questions to prompt discussion and check for understanding after each section; students generated gist statements for specific sections of text, as well.

Tutor training. Of the six tutors providing the Tier 2 intervention, five held master's degrees. The lead researcher provided an 8-hr training to the tutors on the use of the materials and the intervention practices. Tutors practiced providing instruction in pairs and received feedback from the lead researcher.

Implementation Fidelity

Tier 1 and Tier 2 fidelity coding procedure. The fidelity code sheets for the Tier 1 instruction and Tier 2 intervention contained two sections that addressed procedural fidelity and overall quality (i.e., instruction, classroom management). For procedural fidelity, coders rated each instructional practice on a 4-point Likert-type scale ranging from 1 (*low fidelity*) to 4 (*high fidelity*). The global ratings for quality of instruction and classroom management were rated on a 5-point scale with 1 being *lowest quality* and 5 being *highest quality*. One coder participated in two 2-hr trainings (i.e., one each for the Tier 1 and Tier 2 coding) conducted by the first author. The coder has a master's degree in education, 22 years of experience in the field of education, 15 years of experience working on various research projects, and previous experience with fidelity coding in other reading-related intervention studies. The trainings included learning the fidelity form codebook for each tier and practicing assigning codes using sample audio recordings. Following the Tier 1 training, the first author and the coders independently scored one Tier 1 audiotape; interrater agreement was 93%. Following the Tier 2 fidelity training, the first author and the coders independently scored one Tier 2 audiotape; interrater agreement was 95%.

Sampling plan for coding fidelity. The Tier 1 teachers in the aligned T1-T2 condition audio-recorded all STRIVE lessons across the 18 weeks of implementation. We randomly selected one unit of instruction and then randomly selected one lesson from within that unit. The Tier 1 teachers from schools assigned to the nonaligned Tier 2 and the BAU conditions (i.e., teachers who provided typical social studies instruction in Tier 1) recorded 1 week

of BAU social studies instruction per 6-week period. We randomly selected 1 week of instruction and then randomly selected one lesson provided within that week. For the schools assigned to the aligned T1-T2 and nonaligned Tier 2 conditions, we randomly selected two intervention lessons from each unit per tutoring group. The first author double-coded 10% of the Tier 1 ($n = 4$) and Tier 2 lessons ($n = 12$), achieving 95% agreement for both tiers.

Tier 1 fidelity results. We used the Tier 1 fidelity data to identify (a) the extent to which the teachers in the aligned T1-T2 condition implemented the practices as intended and (b) the extent to which the Tier 1 practices in the aligned T1-T2 condition occurred during Tier 1 social studies instruction in the nonaligned Tier 2 condition and the BAU condition. Teachers implemented most of the Tier 1 practices in the aligned T1-T2 condition with high fidelity (M ranging from 3.36 to 4.00; Table 3). Building background knowledge ($M = 2.73$), posing the comprehension purpose question prior to reading ($M = 2.36$), and summary writing ($M = 2.50$) occurred with medium fidelity. Teachers infrequently provided lesson closure at the end of the lesson as they often forgot this step in the lessons ($M = 1.45$). As we expected, the ratings for the nonaligned Tier 2 condition suggest little to no alignment between the Tier 1 and Tier 2 practices. Within this condition, the typical Tier 1 social studies instruction contained medium-level implementation for building background knowledge and text reading with questioning but low implementation for the remaining practices. Finally, teachers in the BAU condition provided little to no implementation of the Tier 1 practices during whole-class social studies lessons (M ranging from 1 to 2.30).

Tier 2 fidelity results. We used the Tier 2 fidelity data to identify the extent to which the tutors in the aligned T1-T2 and nonaligned Tier 2 conditions implemented the practices as intended. Across both conditions, tutors implemented most practices with a high level of fidelity (M ranging from 3.50 to 3.92 in the aligned T1-T2 condition; M ranging from

Table 3. Tier I Implementation Fidelity.

	Range	Aligned T1-T2 (<i>n</i> = 11)			Nonaligned T2 (<i>n</i> = 15)			BAU (<i>n</i> = 10)		
		<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
Procedural fidelity of lesson components										
Build background knowledge	1-4	11	2.73	0.91	15	2.47	0.74	10	1.90	0.88
Explicit vocabulary instruction before reading	1-4	11	4.00	0.00	15	2.00	1.00	10	2.30	0.95
Comprehension purpose question	1-4	11	2.36	0.92	15	1.20	0.41	10	1.00	0.00
Text reading and questioning to check for understanding	1-4	11	3.91	0.30	15	2.53	1.06	10	1.50	1.08
Gist statements ^a	1-4	7	3.71	0.76	15	1.07	0.26	10	1.10	0.32
Explicit vocabulary instruction after reading	1-4	11	3.36	1.03	15	1.20	0.56	10	1.00	0.00
Summary writing ^a	1-4	4	2.50	1.73	15	1.00	0.00	10	1.00	0.00
Lesson closure	1-4	11	1.45	1.04	15	1.07	0.26	10	1.10	0.32
Quality of implementation: Global observation										
Global instruction	1-5	11	4.27	0.65	15	3.80	0.41	10	3.50	0.53
Global classroom management	1-5	11	4.73	0.65	15	4.73	0.46	10	4.90	0.32
Fidelity of dosage										
Mean instructional minutes provided across all sessions		11	61.00	24.6	15	31.27	10.57	10	35.40	14.72

Note. T1 = Tier 1; T2 = Tier 2; BAU = business as usual; *n* = number of observations.

^aGist statements and summary writing were not included in every Strategies for Reading Information and Vocabulary Effectively lesson for teachers in the aligned T1-T2 condition, so the *n* for those components does not sum to the total audios coded.

3.61 to 3.93 in the nonaligned Tier 2 condition; Table 4). The lesson opening ($M = 1.98$, aligned T1-T2 condition; $M = 1.52$, nonaligned Tier 2 condition) and lesson closing ($M = 1.50$, aligned T1-T2 condition; $M = 2.19$, nonaligned Tier 2 condition) occurred less consistently, perhaps due to the fact that tutors skipped these components or forgot to do them due to time constraints. Fidelity coding resulted in mid-high to high ratings for global instruction ($M = 3.88$, aligned T1-T2 condition; $M = 4.08$, nonaligned Tier 2 condition) and management ($M = 3.96$, aligned T1-T2 condition; $M = 4.27$, nonaligned Tier 2 condition), and students received an average of 43 min of intervention in the aligned T1-T2 condition and 40 min in the nonaligned Tier 2 condition.

We were unable to collect Tier 2 audio recordings for students in the BAU condition.

Instead, we collected survey information from students' teachers in order to characterize the intensity of intervention provided. For students in the BAU condition, 74% ($n = 37$) received supplemental reading intervention in addition to their whole-class English language arts block. The majority of those students received intervention in a small-group setting (i.e., two to six students; $n = 33$), three students received one-on-one intervention, and one student received large-group intervention (i.e., seven to 10 students). Most students received intervention three times per week ($n = 19$), three students received intervention twice per week, five students received intervention four times per week, and nine students received intervention daily. Intervention sessions lasted 15 min for nine students, 30 min for 25 students, 45 min for two students, and 1 hr for one student.

Table 4. Intervention Implementation Fidelity.

	Range	Aligned T1-T2 (n = 48)			Nonaligned T2 (n = 60)		
		N	M	SD	N	M	SD
Procedural fidelity of lesson components							
Lesson opening: Goal setting check-in	1-4	47	1.98	1.24	60	1.52	0.83
Build background knowledge	1-4	48	3.92	0.28	60	3.93	0.31
Explicit vocabulary instruction using the vocabulary map	3-4	35	3.60	0.74	44	3.61	0.62
Repeated reading with questions to check for understanding	3-4	48	3.69	0.55	60	3.68	0.60
Get the Gist to identify main ideas	1-4	42	3.50	0.94	53	3.77	0.51
Lesson closure: Goal setting	1-4	46	1.50	0.94	58	2.19	1.29
Quality of implementation: Global observation							
Global instruction	1-5	48	3.88	0.73	60	4.08	0.79
Global classroom management	1-5	48	3.96	0.99	60	4.27	0.78
Fidelity of dosage							
Mean instructional minutes provided across all sessions		48	42.9	4.07	60	40.00	6.03

Note. T1 = Tier 1; T2 = Tier 2; BAU = business as usual; n = number of observations. Not all components were included in each Tier 2 intervention lesson, so the n for each component varies within condition.

Data Collection and Measures

Group Reading Assessment and Diagnostic Evaluation (GRADE) Passage Comprehension Subtest (Pearson, 2001). We administered the GRADE passage comprehension subtest at pretest and posttest as a group-administered, 25-min timed assessment of reading comprehension. The assessment consists of informational and narrative passages ranging in length from two to six paragraphs. Students read each passage silently and answer three to six multiple-choice questions related to the passage (i.e., questioning, predicting, clarifying, and summarizing). Median reliability coefficients for internal reliability, alternate form reliability, and test-retest reliability are .96, .89, and .90, respectively. Raw scores are reported as we were unable to administer the Sentence Comprehension subtest, required for calculation of standard scores, due to testing-time limitations in the district.

Gates-MacGinitie Reading Tests (4th ed.; GMRT; MacGinitie et al., 2000) Vocabulary Subtest. The GMRT Vocabulary subtest is a group-administered, 45-item, 20-min timed assessment of vocabulary knowledge administered at pretest

and posttest. Each item presents a word in brief context followed by five word-meaning choices. The Kuder Richardson 20 reliability for this measure is .90 to .92.

Unit Tests of Content Knowledge. Students in all three conditions responded to three content knowledge assessments administered at the end of each content unit. Each unit content test consists of 20 or 21 items; each item includes a brief sentence stem followed by four answer choices. The tests took approximately 30 min to complete. Item-total correlations were lower than desired, ranging from .33 to .41. Given these low item-total correlations, internal consistency reliabilities also were lower than desired, ranging from .77 to .84 across the three content tests. We refer the readers to Swanson et al. (2019) for more information on the development of these content knowledge assessments.

Unit Tests of Vocabulary Knowledge. Swanson et al. (2019) developed three vocabulary-matching tests of knowledge of key terms from each of the three STRIVE units. The tests consist of 24 items for Unit 1, 16 items for Unit 2,

and 19 items for Unit 3. For each assessment, students match each word with a brief definition. Students completed the measure in approximately 20 min. Item-total correlations ranged from .54 to .58 with internal consistency reliabilities from .89 to .93. The reliability estimates may be inflated because matching tests use a common set of response choices.

Content Reading Comprehension. A content reading comprehension assessment (Swanson et al., 2019) assessed students' understanding of expository texts containing social studies content. Students read five passages and responded to six multiple-choice items following each passage. Passages ranged from 197 to 233 words and conformed to a fourth-grade reading level with a Lexile range of 700L to 900L. The average item-total correlation was .43, and internal consistency reliability for the assessment was .89.

Strategy Use Measure (SUM). The development of the SUM was informed by the approaches used in studies of theoretical models of comprehension (e.g., Cromley & Azevedo, 2007). In Part 1, students read three leveled passages and then answer open-ended questions that assess their knowledge and use of two specific comprehension strategies: question generation and main-idea generation. In Part 2, students read the same passages and select the best possible main-idea statement from four choices. It takes approximately 20 min to administer with a total of 21 possible points for both sections. Scammacca (2017) examined the reliability and validity of the SUM with interrater reliability of over 95% agreement. A confirmatory factor analysis indicated the measure consisted of two factors (i.e., question generation and main-idea identification) that were moderately correlated ($r = .71$). The validation study reported positive construct validity for the SUM measuring students' question generation and main-idea identification.

Data Analysis

We conducted the current pilot study within the context of a larger RCT, so the multilevel structure of the data was complex. Assignment to

Tier 1 instruction for the RCT had already occurred at the school level. In addition to being nested within schools, the students who received the intervention were nested in teachers (i.e., for Tier 1 instruction) and tutors (i.e., for Tier 2 instruction), and the nesting of students in the BAU condition for Tier 2 instruction was unknown. The tutoring groups contained students from multiple teachers, making data from students in these groups cross-classified. Given the relatively small sample size (about 50 per condition) and the number of students within each unit at each level (i.e., some teachers had only one or two students in Tier 2 intervention, and there were 10 to 12 students per school), computing an analytical model that accurately represented the nesting and cross-classification present in the data would be prohibitive, and it is likely the model would not converge. Given this rationale and the fact that we present this work as a pilot study, we decided to take a simpler approach and completed the analysis at the student level.

We analyzed data from participants in the three conditions using repeated-measures analysis of variance (RM-ANOVA) for measures administered at pretest and posttest (GMRT Vocabulary, GRADE Passage Comprehension, SUM). Data from the unit tests of content and vocabulary knowledge and the content reading comprehension test were analyzed using ANOVA because they were administered at one time point only. Pairwise comparisons were made for all conditions when statistically significant differences were evident in the omnibus test to determine which groups differed. Given the number of hypothesis tests conducted, we implemented the Benjamini-Hochberg procedure for controlling the false discovery rate (Benjamini & Hochberg, 1995), resulting in a critical p value of .031. Effect sizes for all contrasts and their associated standard errors were computed as Hedges' g . For measures administered at pretest and posttest, we utilized the correlation between scores at the two time points in the calculation of g .

Missing Data

As noted in Participants and Setting, eight students withdrew from BAU schools and three

students withdrew from nonaligned Tier 2 schools during the study. There were no statistically differences at pretest within or between conditions when students who withdrew and those who did not were compared in an ANOVA. We included pretest data for students who later withdrew in the assessment of baseline equivalence. At posttest and for the unit tests, data were missing for the withdrawn students after the point of their withdrawal and for a small number of students in the BAU condition. Researchers made every effort to test students in the BAU condition absent from school on testing days. However, due to limitations on when BAU students could be tested, not all students in this condition could be tested on all measures at all time points in cases where they were absent on testing days. We included all available data for each measure at each time point under an intent-to-treat framework.

Results

Baseline Equivalence

Research questions were addressed in the same analyses in which the effects of the aligned T1-T2 condition, the nonaligned Tier 2 condition, and the BAU condition were contrasted. Results of ANOVAs on each pretest measure indicated that there were no statistically significant differences between students in the three conditions at pretest. Effect sizes for the comparison of the T1-T2 and the BAU conditions and the nonaligned Tier 2 and BAU conditions reflected higher scores for the BAU group on the GRADE and GMRT Vocabulary measures. These effect sizes ranged from -0.22 to -0.06 . Effect sizes at pretest for the SUM were -0.02 and 0.02 . Given the effect sizes for the differences at pretest on the GRADE and GMRT Vocabulary assessments, we used RM-ANOVA to compare growth from pretest to posttest across conditions in order to address our research questions regarding treatment effects for the aligned T1-T2 condition versus the nonaligned Tier 2 condition.

Effects on Vocabulary

The results of the RM-ANOVA that tested for treatment effects on the GMRT Vocabulary

subtest indicated there were no statistically significant differences between conditions for change in scores from pretest to posttest, $F(2, 133) = 0.19, p = .83$. The ANOVAs for scores on the unit tests of social studies vocabulary indicated statistically significant differences between groups for Unit 1, $F(2, 132) = 7.51, p = .001$; Unit 2, $F(2, 133) = 7.63, p = .001$; and Unit 3, $F(2, 130) = 2.85, p = .02$. Post hoc comparisons showed that on the Unit 1 test, the aligned T1-T2 group scored significantly higher than both the nonaligned Tier 2 ($p < .001$) and BAU ($p = .03$) groups. On the Unit 2 test, the aligned T1-T2 group also scored significantly higher than both the nonaligned T2 ($p = .001$) and BAU ($p = .001$) groups. On the Unit 3 test, the aligned T1-T2 group scored significantly higher than the BAU group only ($p = .02$). See Table 5 for descriptive statistics and effect sizes.

Effects on Content Knowledge

The ANOVAs for the tests of social studies content knowledge indicated statistically significant differences between groups for all three units: Unit 1, $F(2, 137) = 9.13, p < .001$; Unit 2, $F(2, 135) = 13.23, p < .001$; Unit 3, $F(2, 133) = 3.61, p = .03$. Post hoc comparisons showed that the aligned T1-T2 group scored significantly higher than the BAU group on content knowledge for all three units (Units 1 and 2, $p < .001$; Unit 3, $p = .015$) and significantly higher than the nonaligned Tier 2 group for Units 1 ($p = .001$) and 2 ($p < .001$). See Table 5 for descriptive statistics and effect sizes.

Effects on Reading Comprehension

The RM-ANOVA results indicated no statistically significant difference between groups for the GRADE, $F(2, 133) = 2.20, p = .11$, or the SUM, $F(2, 127) = 2.37, p = .10$. For the content reading comprehension assessment, the ANOVA showed statistically significant group differences, $F(2, 132) = 3.56, p = .031$. In the post hoc comparisons, results indicated that the aligned T1-T2 group scored significantly higher than the nonaligned Tier 2 group ($p = .009$). No statistically significant differ-

Table 5. Descriptive Statistics by Condition.

Measure	Aligned T1-T2		Nonaligned T2			BAU			Aligned T1-T2 vs. BAU			Nonaligned T2 vs. BAU			
	M	SD	n	M	SD	n	M	SD	n	g	SE	g	SE	g	SE
GRADE Passage Comprehension pretest	5.58	4.08	48	6.00	3.78	49	6.37	3.15	49	-0.22	0.20	-0.11	0.20		
GRADE Passage Comprehension posttest	8.02	3.92	48	7.87	3.89	46	6.64	4.48	42	0.42	0.21	0.31	0.21		
GMRT Vocabulary pretest ^a	409.75	35.01	48	414.29	29.77	48	416.18	30.29	50	-0.20	0.20	-0.06	0.20		
GMRT Vocabulary posttest ^a	420.71	39.11	48	428.43	28.62	46	428.38	29.53	42	0.00	0.21	0.12	0.21		
Unit 1 content knowledge	8.67	3.52	48	6.54	2.68	48	6.09	3.09	44	0.77*	0.21	0.15	0.21		
Unit 2 content knowledge	8.85	3.21	48	6.13	3.33	48	6.00	2.41	42	0.99*	0.22	0.04	0.21		
Unit 3 content knowledge	9.25	4.29	48	7.63	3.45	46	7.33	3.17	42	0.50*	0.21	0.09	0.21		
Unit 1 vocabulary	9.98	5.20	48	6.49	4.10	49	7.87	3.84	38	0.45*	0.22	-0.34	0.22		
Unit 2 vocabulary	6.27	3.27	48	4.26	2.49	47	4.24	2.83	41	0.65*	0.22	0.01	0.21		
Unit 3 vocabulary	8.46	4.76	48	7.17	4.20	46	6.28	3.74	39	0.50*	0.22	0.22	0.22		
Unit 3 content reading comprehension	13.60	4.68	48	11.26	3.63	46	12.66	4.45	41	0.20	0.21	-0.34	0.21		
SUM pretest	3.54	3.22	48	3.40	3.70	45	3.47	3.70	45	0.02	0.21	-0.02	0.21		
SUM posttest	6.65	4.83	48	5.26	4.51	43	7.23	4.97	39	-0.19	0.21	-0.47	0.22		

Note. T1 = Tier 1; T2 = Tier 2; BAU = business as usual; GRADE = Group Reading Assessment and Diagnostic Evaluation; GMRT = Gates-MacGinitie Reading Tests; SUM = Strategy Use Measure.

^aStandardized, vertical scale scores are presented for the GMRT Vocabulary subtest; all other means represent the raw scores on each assessment.

* $p < .031$.

ence was found between the aligned T1-T2 and BAU groups ($p = .30$). See Table 5 for descriptive statistics and effect sizes.

Discussion

In contrast to early reading intervention results for at-risk readers, comprehension intervention studies show a consistent pattern of null or small effects on the reading comprehension outcomes of struggling readers in the middle grades (i.e., Grades 4 to 8). In fact, some studies suggest that students need long-term, intensive intervention provided over several years in order to maintain existing performance levels (e.g., Vaughn et al., 2012). These findings prompt us to consider ways to develop and implement effective instructional practices that could yield a larger effect on the reading comprehension of middle-grade students. Although calls for more intensive interventions are a reasonable pathway for addressing reading difficulties (Fuchs et al., 2014), many schools and educational personnel struggle to identify adequate time during the school day to provide these more intensive interventions (Zumeta Edmonds et al., 2019).

In this study, we considered another approach to improving outcomes for students with reading difficulties in the middle grades. Realizing that many schools provide approximately 30-min Tier 2-type interventions for students, we considered whether aligning this intervention with enhanced Tier 1 instruction might provide greater influence than Tier 1 and Tier 2 that are not aligned. Our intent was to alter the typically compartmentalized nature of reading instruction across tiered systems of support. Tier 2-type interventions are often nonoverlapping with the pedagogical focus of Tier 1 instruction and frequently address different skills and content than those provided in Tier 1 and may include different language or terms. Multitiered systems of support are intended to provide increasingly intensified instruction, but the approach often lacks the type of cohesion across the tiers that could maximize intensity of instruction. Alignment may provide a more cohesive and integrated instructional model, particularly for students with reading difficulties who have the most challenging time integrating this variation in instructional approaches.

Our intent was to alter the typically compartmentalized nature of reading instruction across tiered systems of support. Tier 2-type interventions are often nonoverlapping with the pedagogical focus of Tier 1 instruction and frequently address different skills and content than those provided in Tier 1 and may include different language or terms.

In this study, we investigated this issue directly by contrasting an aligned intervention with a nonaligned intervention and also included a BAU contrast condition. By aligning Tiers 1 and 2, we aimed to provide more comprehensive, connected reading instruction in which students learned the same reading comprehension and vocabulary practices in both Tier 1 and Tier 2, thus allowing for extended practice opportunities. Students in the aligned condition outperformed students in the nonaligned intervention condition and those in the BAU condition on the reading comprehension, content knowledge, and vocabulary outcomes. These findings supported the hypothesis that students in the aligned condition would perform significantly better than those in BAU on the unit content knowledge and vocabulary measures with medium to large effect sizes (effect size range of 0.45 to 0.99). There were no significant differences at posttest on the content reading comprehension (effect size of 0.20 for aligned T1-T2 vs. BAU) or the standardized vocabulary and reading comprehension measures, though the effect size of 0.42 on the standardized passage comprehension subtest is potentially promising.

Students in the aligned condition outperformed students in the nonaligned intervention condition and those in the BAU condition on the reading comprehension, content knowledge, and vocabulary outcomes.

We interpret these findings as indicating preliminary support for the benefits of aligning Tier 1 and Tier 2 instruction for students with reading difficulty in the middle grades. One possible explanation is that front-loading instruction through Tier 2 prior to being taught in Tier 1 can potentially improve students' access to the Tier 1 instruction. Our rationale for this interpretation is that students benefited significantly from the combined Tier 1 and Tier 2 instruction, whereas when the same Tier 2 instruction was provided but without alignment with the Tier 1 instruction, students' performance was not significantly influenced. These findings confirm results from previous studies (Chambers et al., 2011; Fien et al., 2015) in which aligned multitiered intervention supported at-risk readers in the primary grades. To our knowledge, the study reported here is the only one conducted at the upper-elementary grades examining the effect of aligned T1-T2 instruction. However, another possible explanation worth mentioning is that students in the aligned T1-T2 condition received nearly twice the amount of instructional time in the Tier 1 setting than students in the nonaligned and BAU settings. Future research to investigate the effects of aligned T1-T2 instruction is warranted in order to identify the crucial elements related to student outcomes.

We also expected the nonaligned intervention students to outperform the BAU group at posttest but not as strongly as those in the aligned condition. This expectation was not supported by the results as there were no significant differences between the nonaligned intervention students and the BAU students at the end of the study. One possible explanation for this finding is that the aligned condition allowed for preteaching practices such that students were more prepared to receive that instruction when it was introduced in Tier 1, thus increasing students' engagement and participation with intentional, structured practice opportunities in Tier 1. It may be that students in the nonaligned condition simply did not have enough practice opportunities across the intervention without the added benefit of the same types of practice opportunities in Tier 1. Fien et al. (2015) reported an interaction between condition and practice opportunities

on the oral reading fluency outcomes of at-risk first graders in their alignment study; students in the treatment condition outperformed students in the comparison group not because they had more practice opportunities but because those opportunities were more intentional and structured and included immediate feedback. In other words, the alignment provided in Tier 1 and Tier 2 provided structured opportunities for students to engage in the same practice (e.g., main-idea generation), using the same approach (i.e., Get the Gist), with the same language or terms (e.g., a main idea is referred to as a "gist" in both settings). The connection among instruction across the tiers may have required less guesswork for students, making the time in both instructional settings more meaningful. Students had opportunities to practice the skill in more depth with specific feedback while in intervention and then to apply that practice in the Tier 1 setting.

Another possible explanation for the similar outcomes at posttest among the nonaligned intervention and BAU conditions is that the intervention simply was not intensive enough for students in the nonaligned condition. The extended scale scores on the GMRT Vocabulary subtest suggest this sample was performing between the 2nd and 5th percentiles in vocabulary understanding at pretest. Students received intervention three times per week in groups of five or six for an average of 40 min per session. It may be that these students need intervention provided more frequently, for a longer duration of time (i.e., more than 54 sessions), and in smaller groups (i.e., three or four students).

In other words, the alignment provided in Tier 1 and Tier 2 provided structured opportunities for students to engage in the same practice (e.g., main-idea generation), using the same approach (i.e., Get the Gist), with the same language or terms (e.g., a main idea is referred to as a "gist" in both settings).

Limitations and Future Directions

Given that this is an initial study examining the effect of aligned instruction for struggling readers in the upper-elementary grades, we offer our recommendations regarding the implementation of aligned instruction cautiously—recognizing that additional research with a larger sample is needed. We interpret the results of this study as promising and warranting of future research. In particular, given the limitations the sample size posed for our analytical approach, an RCT with a sufficient sample to have adequate power to model the data using a multilevel approach and account for cross-classification is a crucial next step in understanding the efficacy of aligning Tier 1 and Tier 2 instruction.

One might argue that the conceptualization of aligned Tier 1 instruction and Tier 2 intervention is meaningful, and the results of this study suggest promise; however, there are practical challenges to implementation, including inadequate personnel to support Tier 2 interventions and inadequate instructional time. We can appreciate this concern and recognize the importance of identifying interventions that can be implemented in existing school frameworks. However, considering the long-standing pattern of null findings or small effects in reading comprehension research for upper-elementary and middle-grade struggling readers, examining alternative models for supporting struggling readers beyond typical interventions that are loosely related to Tier 1 are necessary.

Hill and colleagues (2012) noted the lack of implementation fidelity data reported for Tier 1 when authors evaluated Tier 2 interventions. In order to understand the effect of Tier 2-type interventions, it may be important for future research to include information about Tier 1 so that we can better understand the extent to which the tiers align and the ways in which this may contextualize the findings. We propose collecting additional fidelity data during Tier 1 instruction in future studies to address this issue. It may be useful to reconsider the way in which we think about multitiered systems of support, viewing Tiers 1 and 2 as connected and intentionally planning instructional delivery across the school day (e.g., What is this student receiving in Tiers 1 and 2?) rather

than as separate pieces of the puzzle (e.g., What intervention is this student receiving?). Further research investigating the effects of aligning whole-class instruction and intervention for adolescent readers will help to identify the added benefit of aligned instruction and the extent to which this can be implemented in schools.

It may be useful to reconsider the way in which we think about multitiered systems of support, viewing Tiers 1 and 2 as connected and intentionally planning instructional delivery across the school day (e.g., What is this student receiving in Tiers 1 and 2?) rather than as separate pieces of the puzzle (e.g., What intervention is this student receiving?).

References

- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*, *57*, 289–300. <https://doi.org/10.1111/j.1467-9868.2010.00746.x>
- Chambers, B., Slavin, R. E., Madden, N. A., Abrami, P., Logan, M. K., & Gifford, R. (2011). Small-group, computer-assisted tutoring to improve reading outcomes for struggling first and second graders. *Elementary School Journal*, *111*, 625–640. <https://doi.org/10.1086/659035>
- Connor, C. M., Phillips, B. M., Kim, Y. S. G., Lonigan, C. J., Kaschak, M. P., Crowe, E., . . . Al Otaiba, S. (2018). Examining the efficacy of targeted component interventions on language and literacy for third and fourth graders who are at risk of comprehension difficulties. *Scientific Studies of Reading*, *22*, 462–484. <https://doi.org/10.1080/10888438.2018.1481409>
- Cromley, J. G., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, *99*, 311–325. <https://doi.org/10.1037/022-0663.99.2.311>
- Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures.

- Educational Researcher*, 38, 181–199. <https://doi.org/10.3102/0013189X08331140>
- Fien, H., Smith, J. L. M., Smolkowski, K., Baker, S. K., Nelson, N. J., & Chaparro, E. (2015). An examination of the efficacy of a multi-tiered intervention on early reading outcomes for first grade students at risk for reading difficulties. *Journal of Learning Disabilities*, 48, 602–621. <https://doi.org/10.1177/0022219414521664>
- Foorman, B., Beyler, N., Borradaile, K., Coyne, M., Denton, C. A., Dimino, J., . . . Keating, B. (2016). *Foundational skills to support reading for understanding in kindergarten through 3rd grade: Educator's practice guide* (NCEE 2016-4008). National Center for Education Evaluation and Regional Assistance.
- Fuchs, D., Fuchs, L. S., & Vaughn, S. (2014). What is intensive instruction and why is it important? *TEACHING Exceptional Children*, 46(4), 13–18. <https://doi.org/10.1177/0040059914522966>
- Gersten, R., Dimino, J., Jayanthi, M., Kim, J., & Santoro, L. (2010). Teacher study group: Impact of the professional development model on reading instruction and student outcomes in first grade classrooms. *American Educational Research Journal*, 47, 694–739. <https://doi.org/10.3102/0002831209361208>
- Goldman, S. R., Snow, C., & Vaughn, S. (2016). Common themes in teaching reading for understanding: Lessons from three projects. *Journal of Adolescent & Adult Literacy*, 60, 255–264. <https://doi.org/10.1002/jaal.586>
- Hairrell, A., Rupley, W. H., Edmonds, M., Larsen, R., Simmons, D., Willson, V., . . . Vaughn, S. (2011). Examining the impact of teacher quality on fourth-grade students' comprehension and content-area achievement. *Reading & Writing Quarterly*, 27, 239–260. <https://doi.org/10.1080/10573569.2011.560486>
- Hill, D. R., King, S. A., Lemons, C. J., & Partanen, J. N. (2012). Fidelity of implementation and instructional alignment in response to intervention research. *Learning Disabilities Research & Practice*, 27, 116–124. <https://doi.org/10.1111/j.1540-5826.2012.00357.x>
- Institute of Education Sciences. (n.d.). *Reading for understanding research initiative*. <https://ies.ed.gov/ncer/projects/program.asp?ProgID=62>
- Kamil, M. L., Borman, G. D., Dole, J., Kral, C. C., Salinger, T., & Torgesen, J. (2008). *Improving adolescent literacy: Effective classroom and intervention practices: A practice guide* (NCEE 2008-4027). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Klingner, J., Vaughn, S., & Schumm, J. (1998). Collaborative strategic reading during social studies in heterogeneous fourth-grade classrooms. *Elementary School Journal*, 99, 3–22. <https://doi.org/10.1086/461914>
- Lonigan, C. J., Burgess, S. R., & Schatschneider, C. (2018). Examining the simple view of reading with elementary school children: Still simple after all these years. *Remedial and Special Education*, 39, 260–273. <https://doi.org/10.1177/0741932518764833>
- MacGinitie, W. H., MacGinitie, R. K., Maria, K., Dreyer, L. G., & Hughes, K. E. (2000). *Gates-MacGinitie Reading Tests* (4th ed.). Houghton Mifflin Harcourt.
- National Governors Association Center for Best Practices & Council of Chief State School Officers (2010). *Common Core State Standards English Language Arts*. Washington, DC: Authors.
- Pearson (2001). *Group Reading Assessment and Diagnostic Evaluation*. Pearson.
- Scammacca, N. (2017). *Technical summary: Strategy Use Measure*. Meadows Center for Preventing Educational Risk.
- Scammacca, N. K., Roberts, G. J., Cho, E., Williams, K. J., Roberts, G., Vaughn, S. R., & Carroll, M. (2016). A century of progress: Reading interventions for students in Grades 4–12, 1914–2014. *Review of Educational Research*, 86, 756–800. <https://doi.org/10.3102/0034654316652942>
- Scanlon, D., Gelzheiser, L., Vellutino, F., Schatschneider, C., & Sweeney, J. (2008). Reducing the incidence of early reading difficulties: Professional development for classroom teachers versus direct interventions for children. *Learning and Individual Differences*, 18, 346–359. <https://doi.org/10.1016/j.lindif.2008.05.002>
- Simmons, D., Hairrell, A., Edmonds, M., Vaughn, S., Larsen, R., Rillson, V., . . . Byrns, G. (2010). A comparison of multiple-strategy methods: Effects on fourth-grade students' general and content-specific reading comprehension and vocabulary development. *Journal of Research on Educational Effectiveness*, 3, 121–156. <https://doi.org/10.1080/19345741003596890>
- Stevens, E. A., Walker, M. A., & Vaughn, S. (2017). The effects of reading fluency interventions on the reading fluency and reading comprehension performance of elementary students with learning disabilities: A synthesis

- of the research from 2001–2014. *Journal of Learning Disabilities*, 50, 576–590. <https://doi.org/10.1177/0022219416638028>
- Swanson, E., Stevens, E. A., Capin, P., Stewart, A. A., Scammacca, N., & Vaughn, S. (2019). *Effects of distributed professional development featuring vocabulary and comprehension instruction on students' reading outcomes*. Manuscript in preparation.
- Swanson, E., Wanzek, J., McCulley, L., Stillman-Spisak, S., Vaughn, S., Simmons, D., . . . Hairrell, A. (2016). Literacy and text reading in middle and high school social studies and English language arts classrooms. *Reading & Writing Quarterly*, 32, 199–222. <https://doi.org/10.1080/10573569.2014.910718>
- Vaughn, S., Swanson, E. A., Roberts, G., Wanzek, J., Stillman-Spisak, S. J., Solis, M., & Simmons, D. (2013). Improving reading comprehension and social studies knowledge in middle school. *Reading Research Quarterly*, 48, 77–93. <https://doi.org/10.1002/rrq.039>
- Vaughn, S., Wexler, J., Leroux, A., Roberts, G., Denton, C., Barth, A., & Fletcher, J. (2012). Effects of intensive reading intervention for eighth-grade students with persistently inadequate response to intervention. *Journal of Learning Disabilities*, 45, 515–525. <https://doi.org/10.1177/0022219411402692>
- Wanzek, J., Vaughn, S., Scammacca, N. K., Metz, K., Murray, C. S., Roberts, G., & Danielson, L. (2013). Extensive reading interventions for students with reading difficulties after Grade 3. *Review of Educational Research*, 83, 163–195. <https://doi.org/10.3102/0034654313477212>
- Zumeta Edmonds, R., Gandhi, A. G., & Danielson, L. (2019). *Essentials of intensive intervention*. Guilford Press.

Authors' Note

This research was supported in part by the Meadows Center for Preventing Educational Risk at the University of Texas at Austin and the Institute of Education Sciences, U.S. Department of Education, through Grant R305A150407 to the University of Texas at Austin. The opinions expressed are those of the authors and do not represent views of the Meadows Center, the Institute of Education Sciences, or the U.S. Department of Education.

ORCID iD

Elizabeth A. Stevens  <https://orcid.org/0000-0002-8412-1111>

Manuscript submitted July 2019; accepted October 2019.