Original Research

The Differential Efficacy of a Professional Development Model on Reading Outcomes for Students With and Without Disabilities

Elizabeth Swanson[1] (iD)**, Sharon Vaughn**[1]**, Anna-Maria Fall**[1]**, Elizabeth A. Stevens**[2] (iD)**, Alicia A. Stewart**[3]**, Philip Capin**[1]**, and Greg Roberts**[1]

Abstract

This study examined the differential effects of Strategies for Reading Informational Text and Vocabulary Effectively (STRIVE) on the vocabulary, reading comprehension, and content learning among 4,757 fourth-grade students with and without disabilities. Schools were assigned to one of three conditions: researcher-supported professional development (RPD), school-supported professional development (SPD), or business as usual (BAU). Treatment teachers received 12 hr of professional development distributed over time and delivered a set of instructional practices for teaching vocabulary and comprehension embedded within 36 social studies lessons (45 min per lesson). Findings revealed that students with and without disabilities in classes assigned to STRIVE treatment (SPD and RPD) outperformed their peers in classes assigned to the BAU condition at a statistically significant level on measures of content knowledge and content vocabulary and a distal measure of vocabulary. STRIVE did not affect students with disabilities (SWDs) and without disabilities differentially on the measures of content knowledge and distal vocabulary, though effects sizes for non-SWDs in SPD and RPD were larger than those for SWDs.

Within general education elementary school classrooms, student reading comprehension skills can range from one to more than seven grade levels (Firmender et al., 2013; Gilmour et al., 2019), and as students get older, the range increases (Gagné, 2005). This academic diversity produces a substantial challenge for teachers and requires them to plan carefully to meet the literacy needs of all students. One way to meet student needs is to differentiate instruction. When teachers use multiple grouping practices combined with differentiated instruction and content, students benefit. However, this form of differentiation is often implemented inconsistently (Firmender et al., 2013; Reis et al., 2004), requiring considerable time for planning and requiring extensive teacher knowledge, skill, and practice.

At the elementary level in particular, large numbers of students with disabilities (SWDs) are included in the general education classroom (McFarland et al., 2019). With a reading gap between students with and without disabilities as wide as 3.3 years (Gilmour et al., 2019), providing successful pathways for SWDs to access the general education curriculum is essential. In an effort to effectively address the diverse

[1]University of Texas at Austin
[2]Georgia State University
[3]Central Connecticut State University

Corresponding Author:
Elizabeth Swanson, Meadows Center for Preventing Educational Risk, University of Texas at Austin, 1912 Speedway, D4900, Austin, TX 78712, USA.
Email: easwanson@austin.utexas.edu

academic needs of their students and to provide SWDs meaningful access to the general education curriculum, teachers seek ways to leverage instructional practices that have broad effects across a range of learners within the general education classroom.

> *Teachers seek ways to leverage instructional practices that have broad effects across a range of learners within the general education classroom.*

We identify social studies as a particularly useful content area for improving both literacy and content knowledge for a range of learners (Lee & Sprately, 2010; National Governors Association Center for Best Practices & Council of Chief State School Officers, 2012; Scruggs, 2012; Swanson et al., 2015). One reason is that a variety of text types are available. For example, primary and secondary sources provide readers with access to a wide range of perspectives, historical context (National Council for the Social Studies, 2013), and rich vocabulary applicable across content areas. Biographies and first-person accounts often read like stories that bring social studies alive and help elementary school students relate to historical accounts.

Although many teachers include vocabulary and reading comprehension practices during social studies, they are often misaligned with the evidence base and lack explicit instruction and strategy instruction (Capin et al., 2020; Swanson et al., 2016). In addition, teachers often rely on lecture and relatively narrow vocabulary instruction (e.g., providing a student-friendly definition and using examples and nonexamples). However, findings from multiple meta-analyses provide evidence that leveraging interesting content and text variety available in social studies provides key opportunities to teachers to meet the literacy needs of all students. For example, content enhancement techniques, like graphic organizers and explicit vocabulary instruction, taught within the content area of social studies have a moderate-to-large effect on reading comprehension (Gajria et al., 2007) and content knowledge (Scruggs et al., 2010) among SWDs. This aligns with meta-analytic findings that report reading interventions using social studies content have a large effect on content and comprehension outcomes for students with learning disabilities in particular (Swanson et al., 2014). A recent series of studies that complement these meta-analytic findings examined the efficacy of a set of social studies lessons implemented in 6-week units that addressed explicit vocabulary instruction combined with text-based discussion and team-based learning for middle-grade students. Across multiple studies, authors reported statistically significant effects in favor of the treatment group on reading comprehension, content knowledge, and informational text comprehension among struggling comprehenders (Vaughn et al., 2019), SWDs (Swanson et al., 2015; Wanzek et al., 2016), and students without disabilities (Vaughn et al., 2013, 2014).

In a recent examination exploring the idea of academic diversity (Wanzek et al., 2019), authors reported that students with varying levels of background knowledge and reading achievement benefited equivalently when their teachers provided a set of effective instructional practices. Wanzek et al.'s (2019) study provides additional evidence that infusing social studies with an evidence-based set of vocabulary and reading comprehension practices can influence student reading outcomes no matter the variation in prior background knowledge and reading achievement. These meta-analytic and study findings indicate that teaching vocabulary and reading comprehension practices within general education social studies classrooms can have a positive effect on students both with and without disabilities. Unfortunately, elementary school teachers often report an unfamiliarity about how to teach comprehension practices using informational text—much like those used in social studies (Marinak & Gambrell, 2008)—requiring careful consideration of not only the instructional practices but also the professional development (PD) model used to influence classroom implementation.

> *Elementary school teachers often report an unfamiliarity about how to teach comprehension practices using informational text.*

STRIVE

The most commonplace type of PD provided to teachers is a "one-shot" session where they learn a new set of instructional practices and are expected to use them with students. Converging evidence indicates that this type of PD often does not result in much change in classroom practice (Darling-Hammond et al., 2017; Richardson, 2003). However, when an initial PD session is coupled with meaningful follow-up opportunities focused on how to use lessons, procedures, and materials, there is meaningful effect on teachers' knowledge, skills, and classroom practices (Garet et al., 2001, 2008) as well as students' reading comprehension outcomes (Klingner et al., 2004; Porche et al., 2012). The STRIVE PD model reflects this sociocultural theoretical perspective that knowledge acquisition is not a solitary action and instead depends heavily on others' perspectives and the setting for which the knowledge is intended (Tudge & Scrimsher, 2003). Therefore, STRIVE combines features from several efficacious PD designs, namely, an initial workshop-style PD session focused on evidence supporting the use of a small number of instructional practices (Klingner et al., 2004; Porche et al., 2012) combined with small-group teacher study-team meetings focused on examining the curriculum materials and lessons (Gersten et al., 2010).

STRIVE features a set of vocabulary and reading comprehension practices that are distributed over time so that both teachers and students are not "bombarded" with all practices at once. Instead, instructional practices are added over time to allow for plenty of practice with a feasible number of new instructional practices (see the online supplement). During development (Simmons et al., 2010), researchers and teachers worked together with guidance from multiple practice guides (e.g., Gersten et al., 2007; Pashler et al., 2007; Shanahan et al., 2010) to enhance the features of social studies and literacy instruction to meet the needs of a range of learners, including SWDs. For example, vocabulary instruction centers on a graphic organizer that provides multiple opportunities for explicit instruction on both word meanings and strategies for student-driven word learning (i.e., a context clue strategy; Gersten et al., 2007). Vocabulary instruction, in particular, connects abstract ideas and concrete representations of key terms (Gersten et al., 2007; Pashler et al., 2007). Instruction is spaced over time and combines graphics with opportunities for discourse surrounding images and text (Pashler et al., 2007). During reading, teachers engage students in discussion rooted in deep questions (Pashler et al., 2007; Shanahan et al., 2010). See the online supplement for more on STRIVE instructional practices.

STRIVE (i.e., distributed PD featuring evidence-based instructional practices) efficacy has been investigated through two randomized controlled trials (RCTs). The first RCT provided evidence that students whose teachers participated in STRIVE and implemented the instructional practices outperformed those who did not on measures of social studies content knowledge and vocabulary (Simmons et al., 2010). In the most recent investigation (Swanson et al., 2021), the scope was broadened to examine a sustainable PD delivery format that transfers ownership of PD and instructional practices to school-based leaders, a feature of PD theorized as related to uptake and sustainability (Coburn, 2003). Teachers were randomly assigned to a researcher-supported PD (RPD) condition, a school-supported PD (SPD) condition, or a business-as-usual (BAU) condition. Effects were statistically significant on measures of content knowledge (effect size [ES] = 0.51–0.55) content vocabulary (ES = 0.49), social studies reading comprehension (ES = 0.16–0.26), and general vocabulary (ES = 0.07). Groups performed equally well on a distal measure of reading comprehension (ES = 0.04–0.06). Findings provided evidence that STRIVE had a causal effect on student reading outcomes, and a more sustainable form of PD delivered by school-based leaders was equally efficacious as PD delivered by researchers.

Although prior RCTs provide evidence that STRIVE is efficacious among fourth graders in general, neither investigated the extent to which STRIVE was effective for SWDs or the extent to which STRIVE effected students with and without disabilities differentially. Addressing these questions may further our understanding of the benefit of providing teachers with a set of instructional practices that support students with a wide range of abilities. The purpose of this study was to examine whether STRIVE differentially effects outcomes among students with and without disabilities. The following research questions guided our work: (a) What is the effect of STRIVE on reading outcomes for SWDs included in the fourth-grade general education classroom? (b) What is the differential response to STRIVE among students with and without disabilities included in the fourth-grade general education classroom?

Method

Data for this study were drawn from a multi-cohort RCT examining the efficacy of STRIVE in general education fourth-grade classrooms (Swanson et al., 2021). The current work represents secondary analysis of the data set. Secondary analysis is particularly useful when investigating research questions in addition to those explored in the initial analysis (Payne & Payne, 2004). The current study extends prior work by examining the effect of STRIVE on outcomes for SWDs. Secondary analyses depend on the quality of the original research (Payne & Payne, 2004). In this case, the original study was a multicohort RCT with randomization conducted at the school level that meets the What Works Clearinghouse (WWC; 2020) group design standards without reservations. This article highlights information pertinent to the secondary analysis of the STRIVE data set, and the authors invite readers to refer to the previously published article to read specifics about the prior study (i.e., Swanson et al., 2021).

Human subject interaction was approved by the institutional review board. A total of 81

schools were blocked on district and randomly assigned to one of three conditions: RPD, SPD, or BAU. Teachers in schools assigned to the RPD and SPD conditions received STRIVE that included an initial daylong PD session delivered by researchers followed by a series of two teacher study-team meetings provided by either a member of the research team (RPD condition) or a school leader (SPD condition). During PD, teachers learned a set of vocabulary and reading comprehension instructional practices that were delivered to their fourth-grade students over the course of 36 lessons delivered twice per week over an 18-week time period.

Controlling for Content, Scope, and Sequence

As a means of providing an equivalent content base across all conditions, all schools used the state-adopted fourth-grade social studies curriculum. This means that every school used a state-adopted textbook and followed the same state-developed standards and state-developed scope and sequence for teaching fourth-grade social studies content. The units of study during the 18-week period across all conditions were (a) Native American groups prior to European exploration, (b) European exploration and colonization, and (c) revolution and annexation of the state. The only difference for the treatment conditions was STRIVE—teacher participation in PD and implementation of the instructional practices during social studies.

Setting

The study took place in 81 schools located in six school districts in one state located in the southwestern United States. Two schools that were randomized to participate did not complete the study. One school assigned to the BAU condition declined to participate after a new principal was appointed. No data were collected at this school. A second school, assigned to the RPD condition, sent teachers to the initial PD session and allowed student data to be collected at pretest and after the first

unit (of three total units) of content. At that time, the school withdrew from STRIVE in order to focus on improving the school's state standardized test performance rating. This produced an overall attrition of 2.5%, well within the limits of attrition deemed tolerable by the WWC (2020). We determined not to include data from these schools in the analyses because it likely was not missing at random. Within these 79 schools, a variety of services were available to SWDs. Resource room services and co-teaching were provided in math and English language arts. However, all SWDs in this sample were included in the general education classroom for social studies.

Participants

All 235 fourth-grade teachers (210 female and 25 male) from the 79 urban and near-urban schools consented to participate in the study; teachers received compensation for attending PD meetings and completing project tasks outside of their regularly scheduled workday. In all conditions, teachers averaged 10 years of teaching experience, and all were certified teachers (see Table 1). There were no statistically significant differences between teachers in the three conditions in average years of teaching experience, $F(1, 34) = 0.583$, $p = .446$, or the percentage of teachers who held master's degrees, $\chi^2(2) = 1.607$, $p = .448$.

We asked the principal at each school to identify a coordinator for the campus who served as the liaison between the campus and the research team. Coordinators in schools assigned to the SPD condition also led the teacher study-team meetings at their schools. Demographic information is in Table 1. All coordinators were female with more than 12 years of teaching experience and held either teacher or instructional coach positions at the school. Coordinators received a small stipend for completing these extra duties.

Parents of 4,757 students who received social studies in the general education setting provided consent for their children to participate in data collection efforts (77% return rate). Of these, 485 were SWDs and 3,897

were students without a disability. Student demographic information is provided in Table 2. Learning disability was the most commonly identified disability category (40% to 45.2% of SWDs), followed by speech impairments (16.4% to 28% of SWDs). Across SWDs and students without disabilities, most were of Hispanic or Caucasian ethnicity, and a majority of the students qualified to receive free or reduced lunch.

Independent Variable

The independent variable in this study is STRIVE—a distributed PD model featuring a set of evidence-based vocabulary and reading comprehension practices embedded within social studies content. In intervention research, the intervention being examined is often yoked to PD, particularly when the intervention is delivered by schoolteachers. In a recent meta-analysis and synthesis of 37 studies conducted between 2000 and 2015 investigating the efficacy of evidence-based practices delivered in the general education classroom (Swanson et al., 2017), 34 of the 37 studies described PD provided to teachers so that they could master using the instructional practices being studied. In these studies, authors chose to focus on the efficacy of the intervention with less attention to the role that PD played in the independent variable. The current study shifts that focus to include within the independent variable a full model we call STRIVE—a distributed PD model featuring a set of evidence-based vocabulary and reading comprehension practices embedded within social studies content. In so doing, we acknowledge the potency of a package that contains an evidence-based PD model used to show teachers how to implement the evidence-based instructional practices. Therefore, in the following sections, we describe both the PD and the instructional practices combined into what we call STRIVE.

STRIVE PD. The PD was distributed over time and was delivered across the course of three units of study. Teachers in both treatment conditions participated in a PD session prior to the

Table 1. Teacher and School Coordinator Demographics.

Variable	School-supported PD ($n = 28$)	Researcher-supported PD ($n = 25$)	BAU comparison ($n = 26$)
Teachers			
n	80	77	78
Gender			
Male	9	11	5
Female	71	66	73
Years teaching, *M*	10.90	9.68	10.20
Certification			
Elementary	78	76	73
Secondary	13	14	9
Special education	8	2	5
ESL	22	12	14
Bilingual	14	18	23
Degree			
Bachelor's	54	49	45
Master's	25	27	29
Coordinators			
n	29	25	25
Gender			
Male	0	0	1
Female	29	25	24
Years teaching, *M*	13.90	13.60	12.20
Job title			
Teacher	15	7	13
Assistant principal	0	1	1
Principal	0	1	0
Implement specialist	0	1	0
Instructional coach	14	15	11

Note. One teacher in the researcher-supported condition, one teacher in the school-supported condition, and four teachers in the BAU condition did not report degree status. PD = professional development; BAU = business as usual; ESL = English as a second language.

school year. They also participated in teacher study-team meetings prior to Unit 2 and Unit 3.

During the initial 8-hr PD session led by researchers, teachers received an overview of all STRIVE instructional practices. They also watched videos of teachers modeling the practices and in-person models of Unit 1 practices followed by time to practice with their peers and plan for implementation in classrooms. Teachers also learned about the importance of treatment adherence to support student outcomes. After attending the initial training, teachers implemented Unit 1 lessons in their classrooms. Prior to Unit 2 and again prior to Unit 3, teachers participated in a teacher study-team meeting that lasted approximately 2 hr after school. They were led by members of the STRIVE research team (in the RPD condition) or a school coordinator (in the SPD condition). To promote collaboration, meetings included an average of four teacher participants. Researchers and school coordinators followed the same meeting agenda, consisting of three components: (a) reflect on instruction from the previous unit, (b) introduce new practices for the coming unit, and (c) set one major instructional goal for the coming unit.

STRIVE Instructional Practices. Teachers in schools assigned to both treatment groups implemented the instructional practices across

Table 2. Student Demographics.

Variable	SPD Non-SWDs		SPD SWDs		RPD Non-SWDs		RPD SWDs		BAU Non-SWDs		BAU SWDs	
	n	%	n	%	n	%	n	%	n	%	n	%
Gender												
Male	740	49.3	99	55.3	667	49.8	113	63.8	484	47.3	75	62
Female	760	50.7	80	44.7	673	50.2	64	36.2	540	52.7	46	38
Ethnicity												
Caucasian	355	23.4	42	23.2	262	19.5	30	16.8	136	13.3	15	12.2
African American	70	4.6			63	4.7	7	3.9	48	4.7	7	5.7
Hispanic	1042	68.6	124	68.5	999	74.2	137	76.5	825	80.7	100	81.3
Asian	20	1.3	3	1.7	4	0.3	0	0	1	0.1	0	0
Native American or Pacific Islander	1	0.1	8	4.4	4	0.3	0	0	1	0.1	0	0
Two or more	31	2	4	2.2	15	1.1	5	2.8	11	1.1	1	0.8
Economically disadvantaged	970	63.9	117	64.6	869	64.6	123	68.7	762	73.9	97	77.6
Limited English proficient	270	17.8	30	16.9	261	19.4	34	19	246	24	17	13.7
Disability category												
Auditory impairment			3	1.6			4	2.2			0	0
Autism			12	6.5			16	8.7			5	4
Deaf-blind			0	0			0	0			1	0.8
Emotional disturbance			7	3.8			16	8.7			12	9.6
Intellectual disability			8	4.3			8	4.4			6	4.8
Learning disability			84	45.2			80	43.7			50	40
Other health impairment			23	12.4			24	13.1			13	10.4
Speech impairment			39	21			30	16.4			35	28
Traumatic brain injury			0	0			1	0.5			0	0
Visual impairment			1	0.5			0	0			2	1.6

Note. There is no disability category information for students non-SWDs. SPD = school-supported professional development; RPD = researcher-supported professional development; BAU = business as usual; SWDs = students with disabilities.

three 6-week units of study. Each instructional unit consisted of twelve 45-min social studies lessons (36 lessons total). Teachers delivered two lessons per week.

Two instructional practices took place before text reading. First, teachers used illustrations of the text content to lead discussions prompting students to make connections between prior knowledge and new content. Second, explicit vocabulary instruction included introducing four words per week using a student-friendly definition, leading a discussion guided by a visual representation of the word, and providing examples of the word in the appropriate context. In Unit 3, teachers taught students a context clue strategy to derive the meaning of words from the context, rather than providing a student-friendly definition.

During text reading, teachers led a text-based discussion framed by various question types to encourage literal and inferential thinking. They also taught students how to use the Get the Gist strategy several times during passage reading, a method for writing main-idea statements that support content comprehension (Klingner et al., 2012).

After text reading, teachers returned to explicit vocabulary instruction and asked students to evaluate a list of four words to identify the ones related to the target word. Students also wrote a sentence using the word and engaged in a turn-and-talk activity to apply their understanding of words in a way that connected to their own lives (e.g., If you could go on an *expedition*, where would you go and why?). A word-building activity asked students to add prefixes or suffixes to the target word to create new words. In Unit 2, teachers showed students how to use gist statements from sections of the text to write a summary of the entire passage.

Table 3. Professional Development Fidelity.

Preparedness to teach each component	Initial PD		Teacher Study Team 1		Teacher Study Team 2	
	RPD *M (SD)*	SPD *M (SD)*	RPD *M (SD)*	SPD *M (SD)*	RPD *M (SD)*	SPD *M (SD)*
Before reading						
Background knowledge	3.74 (0.47)	3.84 (0.40)	NA	NA	NA	NA
Explicit vocabulary instruction	3.86 (0.35)	3.87 (0.37)	NA	NA	NA	NA
During reading						
Ask and answer questions	3.75 (0.49)	3.81 (0.42)	NA	NA	NA	NA
Get the Gist main idea statements	3.71 (0.51)	3.68 (0.55)	NA	NA	NA	NA
Get the Gist routine in collaborative learning pairs	NA	NA	3.64 (0.61)	3.56 (0.55)	NA	NA
After reading						
Gist to summary writing	NA	NA	3.66 (0.53)	3.48 (0.60)	NA	NA
Comprehension purpose question	3.75 (0.51)	3.81 (0.40)	NA	NA	NA	NA
Vocabulary maps in collaborative learning pairs	3.70 (0.54)	3.87 (0.34)	3.72 (0.51)	3.78 (0.45)	NA	NA
Overall quality	3.96 (0.21)	3.90 (0.30)	3.77 (0.45)	3.82 (0.39)	3.88 (0.33)	3.72 (0.57)

Note. Teachers rated how prepared they felt to teach each practice after the professional development sessions, with 1 being *not prepared* and 4 being *prepared*. Overall quality was rated on a 4-point scale. PD = professional development; RPD = researcher-supported PD; SPD = school-supported PD; NA = not applicable because the instructional practice was not introduced.

Fidelity

Strive PD Fidelity. All teachers attended the initial PD session and teacher study-team meetings. After each, teachers completed a PD fidelity form. In the first section, teachers rated the extent to which they felt prepared to teach each instructional component. In the second section, teachers rated PD quality. Items in both sections were rated on a scale of 1 (*strongly disagree*) to 4 (*strongly agree*). Table 3 shows that teachers in both treatment conditions felt prepared to teach each instructional component across all PD sessions. Teachers rated the PD quality as high, indicating that the PD was engaging, the lesson components were well described, materials were useful, and opportunities to practice were helpful.

STRIVE Implementation Fidelity. The fidelity code sheet was adapted from a previous study (Simmons et al., 2010) to align with STRIVE practices. Instructional adherence items aligned were coded on a 4-point Likert-type rating scale ranging from 1 (*low alignment with intended method*) to 4 (*high alignment with intended method*). The research team

asked teachers randomized to the RPD and SPD conditions to audio record all STRIVE lessons. Teachers assigned to the BAU condition recorded 1 week of BAU social studies instruction per 6-week period. A total of 207 lessons (RPD, 64 lessons; SPD, 70 lessons; BAU, 73 lessons) were randomly selected across conditions for coding. Seven members of the research team participated in a 4-hr training conducted by the principal investigator, who has extensive experience in observation coding (e.g., Swanson et al., 2012, 2016), and met a 90% interrater agreement threshold prior to coding audios. To maintain agreement, one third of audio recordings were double coded, all of which met 90% agreement. Further information about fidelity training can be found in Swanson et al. (2021).

We used fidelity data to identify (a) the extent to which instruction in the treatment conditions aligned with the instructional practices as intended and (b) the extent to which the instructional practices were observed in the BAU condition (Table 4). Scores for the RPD group ranged from 2.75 (building background knowledge) to 3.68 (explicit vocabulary instruction before reading). Scores for the SPD group ranged from

Table 4. Instructional Fidelity.

STRIVE component	Researcher-supported		School-supported		BAU	
	M (SD)	Times observed	M (SD)	Times observed	M (SD)	Times observed
Background knowledge	2.75 (0.95)	55	2.42 (1.18)	60	2.28 (0.90)	57
Explicit vocabulary instruction						
Before reading	3.68 (0.71)	59	3.61 (0.73)	62	1.88 (0.97)	40
After reading[a]	3.26 (0.94)	42	3.44 (0.91)	39	1.33 (0.58)	3
Questions to prompt text-based discussion						
Pose comprehension purpose question	3.14 (1.31)	59	2.45 (1.48)	62	2.82 (1.17)	11
Questions during text reading	3.25 (0.97)	60	3.28 (1.0)	67	2.77 (1.01)	53
Discuss comprehension purpose question	2.84 (1.32)	37	2.87 (1.32)	39	2.33 (1.53)	3
Gist statements	3.22 (0.84)	46	3.00 (1.03)	54	2.14 (0.90)	7
Summaries	3.04 (1.07)	23	2.89 (1.19)	27	1.25 (0.50)	4
Lesson closure	3.00 (1.11)	30	2.77 (1.22)	30	1.92 (1.02)	24

Note. Implementation of components was rated on a 4-point scale (1 lowest and 4 highest). Total observations collected from each condition were as follows: researcher-supported professional development, 67; school-supported professional development, 74; business as usual, 87.
[a]In Unit 3, explicit vocabulary instruction included the context clue strategy.

2.42 (building background knowledge) to 3.61 (explicit vocabulary instruction before reading). Fidelity ratings for the RPD and the SPD conditions were mid-high to high for most practices, indicating that teachers implemented the instructional practices as intended.

We were also interested in the extent to which the instructional practices were observed in the BAU condition. Recorded BAU social studies lessons averaged 34 min ($SD = 19$) in length. The median recorded lesson lasted 26 min. Building background knowledge ($M = 2.28$), explicit vocabulary instruction ($M = 1.88$), and asking questions during reading ($M = 2.77$) were observed in more than 50% of lessons, with alignment to the treatment instructional practices in the low-to-medium-high range. Lesson closure was observed in almost one third of BAU lessons ($M = 1.92$). Explicit vocabulary instruction after reading, discussing the comprehension purpose question, gist statements, and summaries were observed in fewer than 10% of BAU observations. Although there was some overlap of practices, most practices were either mis-aligned (e.g., vocabulary instruction) or not present in the BAU condition. These data provide evidence of differentiation between treatment and comparison groups.

Measures

The Gates-MacGinitie Reading Tests Reading Comprehension subtest (GMRC) and Vocabulary subtest (GMV) were administered by trained testers to students within 2 weeks prior to and 2 weeks immediately following STRIVE implementation. Unit tests of content knowledge and vocabulary knowledge were administered after each unit. The Content Reading Comprehension measure was administered at posttest. A correlation matrix is provided in the online supplement.

GMRC (MacGinitie et al., 2000). The GMRC subtest is a group-administered, 35-min timed assessment of reading comprehension. The assessment consists of informational and narrative passages ranging in length from three to 15 sentences. Students read each passage silently and answer three to six multiple-choice questions related to the passage. Schools in one district placed limits on testing time. As a result, 2,344 students were not administered this measure. Analysis of outcomes on the GMRC was conducted using data from 37 schools in districts that did not place limits on testing time. Internal consistency reliability for the reading comprehension subtest ranges from .91 to .93, and alternative form reliability is .80 to .87.

GMV (MacGinitie et al., 2000). The GMV subtest is a group-administered, 45-item, 20-min timed assessment of vocabulary knowledge. Each item presents a word in brief context followed by five word-meaning choices. Items consist of age-appropriate and frequently encountered vocabulary words. The Kuder Richardson 20 reliability for this measure is .90 to .92.

Content Measures. Three curriculum-based measures tested student content knowledge, vocabulary knowledge, and comprehension of social studies text. Content knowledge and vocabulary knowledge were measured at the end of each 6-week unit. Content reading comprehension was assessed at posttest. Because these measures were used to assess curriculum mastery and many students in the sample were novice English speakers, Spanish versions of the three curriculum-based measures were made available to teachers in all three conditions. None of the teachers in the BAU condition opted to use the Spanish version. As a result, the students in the STRIVE PD conditions who responded to the Spanish version of the content measures were dropped before analyses of outcomes were conducted.

Unit tests of content knowledge. Students in all three conditions responded to three content knowledge assessments. Each unit content test consisted of 20 or 21 items; each item included 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 ranged from .77 to .84 across the three content tests, exceeding the WWC (2020) standards for reliability.

Unit tests of vocabulary knowledge. Using Espin and colleagues' (2001) procedures, Swanson et al. (2018) developed three vocabulary-matching tests of knowledge of terms taught during each of the three units. The tests consisted of 24 items for Unit 1 (score range = 0–24), 16 items for Unit 2 (score range = 0–16), and 19 items for Unit 3 (score range = 0–19). For each assessment,

students matched each word with a brief definition. Students completed the measure in approximately 20 min. We utilized data from students included in Cohort 1 to evaluate the item properties of the unit test of vocabulary knowledge. None of the items on the vocabulary measure were determined to be too difficult for students (based on a criterion of 40% or fewer students responding correctly). Within our sample, item-total correlations ranged from .54 to .58. Internal consistency reliabilities ranged from .89 to .93, exceeding WWC (2020) standards. However, the reliability estimates may be inflated because matching tests use a common set of response choices (i.e., all response choices are considered for each test item).

Content reading comprehension. A content reading comprehension assessment aligned with the state standards also was developed by researchers (see Swanson et al., 2021) to test students' ability to understand expository texts containing social studies content. Students read five passages not previously read and responded to six multiple-choice items following each passage. Fourth-grade level (Lexile 700–900) passages ranged from 197 to 233 words. Internal consistency reliability for the assessment was .89, exceeding WWC (2020) standards.

Data Analysis Plan

To determine the effect of STRIVE among fourth graders with disabilities included in the general education setting, we fit a series of three-level models (Hoffman, 2015), which accounted for how students were nested in teachers within schools. These regression models, based on a path-analytic framework (Preacher & Hayes, 2008), are extensions of previous analyses that estimated the effect of STRIVE reported in Swanson et al. (2021). The statistical model was expanded to include disability status as a predictor of differential response and its interaction with the treatment effect (see Equations 1 and 2). We have separate equations for the content measures and the Gates-MacGinitie Reading Tests because

the content measures, including the unit tests of content knowledge and vocabulary knowledge and the content reading comprehension measure, were administered at posttest only. The GMRC and GMV subtests (MacGinitie et al., 2000), on the other hand, were administered to students within 2 weeks prior to and 2 weeks immediately following treatment implementation.

Content

$$\begin{aligned} \text{Measures}_{ijk} &= \gamma_{000} + \gamma_{001}\left(\text{SPD}\right)_k \\ &+ \gamma_{002}\left(\text{RPD}\right)_k + \gamma_{100}\left(\text{SWD}\right)_{ijk} \\ &+ \gamma_{101}\left(\text{SWD}_{ijk} * \text{SPD}_k\right) \\ &+ \gamma_{102}\left(\text{SWD}_{ijk} * \text{RPD}_k\right) \\ &+ r_{ijk} + e_{0jk} + u_{00k} \end{aligned} \quad (1)$$

Gates-

$$\begin{aligned} \text{MacGinitie}_{ijk} &= \gamma_{000} + \gamma_{100}\left(\text{Pretest}\right)_{ijk} \\ &+ \gamma_{010}\left(\text{Pretest}\right)_{jk} + \gamma_{001}\left(\text{SPD}\right)_k \\ &+ \gamma_{002}\left(\text{RPD}\right)_k + \gamma_{003}\left(\text{Pretest}\right)_k \\ &+ \gamma_{200}\left(\text{SWD}\right)_{ijk} \\ &+ \gamma_{201}\left(\text{SWD}_{ijk} * \text{SPD}_k\right) \\ &+ \gamma_{202}\left(\text{SWD}_{ijk} * \text{RPD}_k\right) \\ &+ r_{ijk} + e_{0jk} + u_{00kv} \end{aligned} \quad (2)$$

Here, i represents students, j represents teachers, and k represents schools. SPD_k and RPD_k are school-level dummy variables representing, respectively, SPD intervention (coded 1) versus BAU (coded 0) and RPD intervention (coded 1) versus BAU (coded 0). SWD_{ijk} is disability status with the non-SWDs group coded 0 and SWDs coded 1. $\text{SWD}_{ijk} * \text{SPD}_k$ and $\text{SWD}_{ijk} * \text{RPD}_k$ represent the cross-level interactions between disability status (Level 1) and treatment assignment (Level 3). In Equation 2, Pretest_{ijk} represents student-level pretest scores centered around the teacher mean, Pretest_{jk} is the teacher-level pretest aggregate centered around the school, and the Pretest_k is school-level pretest aggregate centered around the grand mean.

All analyses were run with "lme4" package in R (Bates et al., 2015). Two-way interactions were further decomposed using "emmeans" package (Lenth et al., 2020) in R. We used Hedges's (2011) equations for cluster-randomized three-level models to estimate effect sizes. We computed effect sizes at the student level based on t statistics from the three-level models.

Results

Table 5 summarizes observed pretest and posttest means and standard deviations for GMRC and GMV subtests. Posttest means and standard deviations for the content measures are provided in Table 6 for all three conditions and for SWDs and non-SWDs.

GMRC

As shown in Table 7, on the GMRC subtest, students' performance in SPD ($\gamma_{001} = 1.29$, $SE = 3.82$, $p = .74$) and RPD schools ($\gamma_{002} = 3.98$, $SE = 3.93$, $p = .32$) did not differ, on average, from students in BAU schools, and treatment's effect did not differ for SWDs and non-SWDs. The effect sizes for the SPD in the non-SWDs and SWDs groups were 0.04, 95% confidence interval (CI) [−0.17, 0.24], and 0.12, 95% CI [−0.09, 0.32], respectively. For the RPD treatment, the effect size in the non-SWDs group was 0.12, 95% CI [−0.11, 0.34], and for SWDs, it was 0.20, 95% CI [−0.03, 0.44].

GMV

The main effect of SPD on the GMV subtest differed significantly from 0 ($\gamma_{001} = 4.25$, $SE = 1.94$, $p = .03$), meaning that non-SWDs participating in the SPD intervention scored about 4.25 points higher at posttest than non-SWDs assigned to BAU schools. The main effect of disability status was significant as well, indicating that SWDs in BAU schools scored 9.88 points lower than non-SWDs in BAU schools. The two-way interaction of SPD × SWD did not differed statistically

Table 5. Student-Level Descriptive Statistics for Gates-MacGinitie Reading Tests.

Variable	Pretest			Posttest		
	n	*M*	*SD*	*n*	*M*	*SD*
Reading comprehension						
Non-SWDs						
SPD	847	470.44	42.69	809	484.56	43.40
RPD	615	473.10	42.72	607	490.81	42.06
BAU	366	471.04	37.21	369	485.42	36.19
SWDs						
SPD	109	452.22	49.67	107	469.98	45.08
RPD	71	470.90	52.59	70	486.60	54.57
BAU	49	466.61	41.38	50	477.12	45.03
Vocabulary						
Non-SWDs						
SPD	1,434	459.73	39.75	1,380	478.66	42.87
RPD	1,280	459.30	40.83	1,220	477.76	42.87
BAU	978	457.41	35.53	929	472.61	39.14
SWDs						
SPD	165	437.47	48.00	164	456.45	51.13
RPD	162	440.75	54.09	161	457.55	54.95
BAU	116	442.22	41.28	106	454.04	46.12

Note. SWDs = students with disabilities; SPD = school-supported professional development; RPD = researcher-supported professional development; BAU = business as usual.

from 0 ($\gamma_{201} = 2.18$, *SE* = 3.29, *p* = .51), indicating that the intervention did not affect SWDs and students without disabilities differentially. The effect size in the non-SWDs group was 0.14, 95% CI [0.04, 0.24]. In the SWDs group, the SPD effect was 0.09, 95% CI [−0.01, 0.19]. Students' performance in RPD schools ($\gamma_{002} = 0.93$, *SE* = 1.97, *p* = .64) did not differ, on average, from students in BAU schools, and treatment's effect did not differ for SWDs and non-SWDs ($\gamma_{202} = 6.21$, *SE* = 3.31, *p* = .06). The effect sizes for the RPD in the non-SWDs and the SWDs groups were 0.02, 95% CI [−0.08, 0.12] and 0.10, 95% CI [0.00, 0.20], respectively.

Unit Tests of Content Knowledge

As shown in Table 8, on the content knowledge measure, students in SPD schools (γ_{001} = 9.16, *SE* = 1.60, *p* = .00) and students in RPD schools (γ_{002} = 8.47, *SE* = 1.61, *p* = .00) outperformed students in the BAU condition. SPD treatment's effect did not differ for

SWDs and non-SWDs, as indicated by the nonsignificant cross-level interaction term ($\gamma_{101} = -1.66$, *SE* = 1.11, *p* = .13). Both groups performed equally well at posttest; the effect sizes for the SPD in the non-SWDs and the SWDs groups, respectively, were 0.57, 95% CI [0.36, 0.79] and 0.40, 95% CI [0.20, 0.61]. The two-way interaction of RPD × SWD was negative and statistically significant ($\gamma_{102} = -2.35$, *SE* = 1.11, *p* = .03), indicating that the intervention did affect SWDs and students without disabilities differentially. Intervention effect was greater for non-SWDs and smaller but still significant for SWDs. The effect size in the non-SWDs group was 0.54, 95% CI [0.32, 0.75]. In the SWDs group, the RPD effect was 0.33, 95% CI [0.13, 0.54].

Content Reading Comprehension

On the content reading comprehension measure, students' performance in SPD schools (γ_{001} = 1.84, *SE* = 0.60, *p* = .00) did differ, on average, from students in BAU schools,

Table 6. Student-Level Descriptive Statistics for Content Measures Test.

Variable	n	M	SD
Unit test of content knowledge			
Non-SWDs			
SPD	1,519	30.02	10.93
RPD	1,347	29.12	10.88
BAU	1,031	20.13	8.22
SWDs			
SPD	181	26.09	11.75
RPD	179	24.08	12.25
BAU	125	17.46	8.56
Content reading comprehension			
Non-SWDs			
SPD	1,519	8.24	4.51
RPD	1,347	7.43	4.53
BAU	1,031	6.28	4.27
SWDs			
SPD	181	7.09	4.08
RPD	179	6.27	4.55
BAU	125	5.47	4.22
Unit test of content vocabulary			
Non-SWDs			
SPD	1,519	39.67	15.86
RPD	1,347	39.44	15.80
BAU	1,031	25.56	13.75
SWDs			
SPD	181	31.90	17.60
RPD	179	29.94	18.65
BAU	125	20.46	14.67

Note. SWDs = students with disabilities; SPD = school-supported professional development; RPD = researcher-supported professional development; BAU = business as usual.

but treatment's effect did not differ for SWDs and non-SWDs ($\gamma_{101} = -.50$, $SE = .49$, $p = .32$). The effect sizes for the SPD in the non-SWDs and the SWDs groups were 0.27, 95% CI [0.09, 0.44] and 0.16, 95% CI [−0.02, 0.33], respectively. The main effect for RPD did not differ statistically from 0 ($\gamma_{002} = 1.07$, $SE = 0.60$, $p = .08$) and the cross-level interaction between disability status and RPD treatment was not significant ($\gamma_{101} = -.33$, $SE = .50$, $p = .51$), suggesting that RPD was not effective overall or for SWDs and non-SWDs when evaluated separately. The effect size in the non-SWDs group was 0.16, 95% CI [−0.02, 0.33]. In the SWDs group, the effect was 0.09, 95% CI [−0.09, 0.26].

Unit Test of Content Vocabulary

On the content vocabulary measure, students in SPD schools ($\gamma_{001} = 13.17$, $SE = 2.27$, $p = .00$) and students in RPD schools ($\gamma_{002} = 13.31$, $SE = 2.29$, $p = .00$) outperformed students in the BAU school. The two-way interactions of SPD $\times$ SWDs ($\gamma_{101} = -3.35$, $SE = 1.67$, $p = .05$) and RPD $\times$ SWDs ($\gamma_{102} = -4.22$, $SE = 1.68$, $p = .01$) were negative and statistically significant, indicating that the intervention did affect SWDs and students without disabilities differentially. Intervention effect was greater for non-SWDs and smaller but still significant for SWDs. For the SPD treatment, the effect size in the non-SWDs group was 0.56, 95% CI [0.35, 0.76]. In the SWDs group, the SPD effect was 0.35, 95%

Table 7. Model Estimates for Gates-MacGinitie Reading Tests.

Variable	Reading Comprehension					Vocabulary				
	Estimate	SE	p	Variance	ICC	Estimate	SE	p	Variance	ICC
Fixed effects										
Intercept	484.92	3.11	0.00			471.32	1.43	0.00		
Main effects										
SPD	1.29	3.82	0.74			4.25	1.94	0.03		
RPD	3.98	3.93	0.32			0.93	1.97	0.64		
SWDs	−6.53	4.23	0.12			−9.88	2.56	0.00		
Two-way interactions										
SWDs × SPD	5.29	5.11	0.30			2.18	3.29	0.51		
SWDs × RPD	7.25	5.55	0.19			6.21	3.31	0.06		
Control variables										
Student-level pretest	0.71	0.02	0.00			0.76	0.01	0.00		
Teacher-level pretest	0.59	0.09	0.00			0.95	0.06	0.00		
School-level pretest	0.90	0.08	0.00			1.12	0.04	0.00		
Random effects										
Student				714.95	0.94				553.53	0.91
Teacher									38.02	0.06
School				48.89	0.06				16.38	0.03

Note. ICC = intraclass correlation; SPD = school-supported professional development; RPD = researcher-supported professional development; SWDs = students with disabilities.

CI [0.15, 0.54]. For the RPD treatment, the effect size in the non-SWDs group was 0.57, 95% CI [0.36, 0.78]. In the SWDs group, the RPD effect was 0.32, 95% CI [0.13, 0.52].

Discussion

This study examined the differential effect of STRIVE among students with and without disabilities. SWDs in schools assigned to STRIVE treatment conditions outperformed their peers with disabilities in schools assigned to the BAU condition at a statistically significant level on measures of content knowledge, content vocabulary, and content text-reading comprehension (SPD group only) and a distal measure of vocabulary (SPD group only). These findings align with the pattern of findings in the larger study that included all fourth-grade students (see Swanson et al., 2020). They also align with RCTs examining the effects of a similar intervention implemented in middle-grade general education social studies classrooms on reading outcomes for SWDs (Swanson et al., 2015; Wanzek et al., 2016). In

these studies, SWDs who received the intervention outperformed those who did not on measures of content knowledge and content text-reading comprehension (Swanson et al., 2015; Wanzek et al., 2016). In the current study, teachers across all conditions used the same state standards and sequence of social studies topics covered. They all used a state-adopted textbook. As an extension, students in all conditions had access to the same content and vocabulary, with the difference between treatment and BAU conditions being the set of instructional practices taught during STRIVE. This finding provides evidence that STRIVE can be used to provide access to the general education social studies curriculum for SWDs. In fact, the effect among this population is small to medium in magnitude.

PD design also likely played a role in group differences. The initial workshop was designed to build teacher knowledge of (a) vocabulary and reading comprehension in general, (b) enacting these skills while using informational text, and (c) learning how to use a specific set of lessons that encompass these

Table 8. Model Estimates for Content Measures.

Variable	Unit test of content knowledge					Content reading comprehension					Unit test of content vocabulary				
	Estimate	SE	p	Variance	ICC	Estimate	SE	p	Variance	ICC	Estimate	SE	p	Variance	ICC
Fixed effects															
Intercept	19.89	1.17	0.00			6.13	0.44	0.00			25.17	1.65	0.00		
Main effects															
SPD	9.16	1.60	0.00			1.84	0.60	0.00			13.17	2.27	0.00		
RPD	8.47	1.61	0.00			1.07	0.60	0.08			13.31	2.29	0.00		
SWDs	−2.93	0.86	0.00			−0.96	0.38	0.01			−5.54	1.29	0.00		
Two-way interactions															
SWDs × SPD	−1.66	1.11	0.13			−0.50	0.49	0.32			−3.35	1.67	0.05		
SWDs × RPD	−2.35	1.11	0.03			−0.33	0.50	0.51			−4.22	1.68	0.01		
Random effects															
Student				77.92	0.70				15.61	0.78				177.72	0.72
Teacher				7.12	0.06				0.80	0.04				16.20	0.07
School				26.48	0.24				3.66	0.18				51.93	0.21

Note. ICC = intraclass correlation; SPD = school-supported professional development; RPD = researcher-supported professional development; SWDs = students with disabilities.

practices. In this way, the PD focused most heavily on what to do in addition to why. The teacher study teams encompassed 4 hr. Although the duration may not be lengthy, the focus of these meetings may actually represent the potent ingredient. Teachers discussed what went well during the prior unit, learned new practices for the upcoming unit, and set goals for the upcoming unit. Because these discussions were situated within the teachers' actual teaching conditions with focus on a specific set of lessons (Gersten et al., 2010; Klingner et al., 2004; Porche et al., 2012), they likely had an influence over classroom practice.

On two outcome measures, SWDs in the SPD group outperformed BAU students on measures of content reading comprehension and general vocabulary. On the same measures, students in the RPD and BAU conditions performed equally well. One explanation for this finding might be the unique features of the SPD condition. During teacher study-team meetings, teachers in both the RPD and SPD groups were led in discussion following the same agenda. We hypothesize that teachers in the SPD condition whose teacher study teams were led by a school coordinator may have felt more comfortable talking about the struggles they faced in the prior unit and may have had more in-depth, open discussions about their students. Social networks, or teachers' interactions with one another, provide ongoing access to knowledge, feedback, and social support that help teachers deepen understanding (e.g., Coburn et al., 2012; Hargreaves & Goodson, 2006; McLaughlin & Mitra, 2001). They also prevent feelings of isolation (McLaughlin & Mitra, 2001) and facilitate the persistence of the practices as they navigate shifting school demands (Hargreaves & Goodson, 2006).

Findings related to the question of differential effects between general and special education learners are a bit more mixed. There were differential effects in favor of students without disabilities on measures of content knowledge and content vocabulary. However, the influence of the intervention was not differential between students with and without disabilities on a measure of content reading comprehension and distal measures of reading comprehension and vocabulary. Therefore, it seems that although the intervention is efficacious for both students with and without disabilities when compared with students in the BAU condition, the magnitude of the differences between treatment and BAU is larger for students without disabilities. This provides evidence that despite the fact that SWDs benefit from STRIVE, it does not serve to close the gap in reading outcomes between students with and without disabilities. Though this may be the case, the value of general education classroom instruction should not be minimized. In fact, Stevens et al. (2020) reported that when Tier 1 and Tier 2 content and instructional practices are aligned, struggling readers performed better on measures of content knowledge and vocabulary than when Tier 1 and Tier 2 were not aligned. The current study combined with Stevens et al.'s findings provide support for the additive role Tier 1 instruction can play in outcomes for SWDs.

We think these findings are important for several reasons. First, we know that social studies teachers value content learning and literacy outcomes (Capin et al., 2020; Swanson et al., 2016). Findings from this study revealed that teachers in the treatment conditions provided both literacy learning and content learning—thus not sacrificing content on the path to improved literacy. Second, general education teachers are often asked to differentiate instruction to meet the needs across a range of learners in their classroom but are unclear what and how to provide this differentiation (Firmender et al., 2013; Reis et al., 2004). Teachers may be confused about differentiation and how they are supposed to systematically alter the instruction for multiple learners in their classroom—an overwhelming task. This study combined with prior findings provides evidence that STRIVE is efficacious across a range of learners (Swanson et al., 2021; Vaughn et al., In press). Finally, SWDs are increasingly placed in general education as a means to access the general education curriculum. Yet, physical placement within a classroom does not guarantee access to the

general education curriculum. Mechanisms to support general education teachers in meeting the instructional needs of SWDs while also meeting the needs across a range of learners in their classroom are crucial.

> *Findings from this study revealed that teachers in the treatment conditions provided both literacy learning and content learning— thus not sacrificing content on the path to improved literacy.*

Findings from this study suggest that a distributed PD model that provides teachers with practices for enhancing literacy and content learning can yield positive effects for students with and without disabilities in the classroom. For SWDs, this may be especially important given the need for multiple opportunities to practice reading text throughout the school day rather than during isolated opportunities in the intervention setting. High-quality general education instruction that weaves vocabulary and reading comprehension practices into social studies provides this opportunity.

> *A distributed PD model that provides teachers with practices for enhancing literacy and content learning can yield positive effects for students with and without disabilities in the classroom.*

Limitations

We recognize that there are several limitations to this research. First, SWDs included in this study represented a range of disability categories (e.g., learning disabilities, autism spectrum disorder). As a result, we do not know the relative effects of the treatment for each specific category of special education. Perhaps even more relevant than the category of special education may be the extent to which students' initial reading (i.e., word-reading scores) predicted their success in benefiting from STRIVE. These more nuanced questions about which SWDs benefited from STRIVE would be valuable to address in future research.

Practical Implications

The STRIVE practices provide teachers with a set of supports for promoting access to content vocabulary, text comprehension, and content knowledge that are associated with improved outcomes for a range of learners. The high scores for participating teachers on the fidelity measure suggests that STRIVE practices are feasible for implementation. The PD was distributed over time but was not excessive. The total amount of teacher time engaged in PD was approximately 12 hr—well within an acceptable amount of PD typically provided. Findings also provide support for a PD model that combines expert-led initial PD followed by school-directed teacher study teams (STRIVE SPD condition). This model advantages both expert knowledge and local context to support implementation of instructional practices with high levels of fidelity.

Future Research

We appreciate that there are many unanswered questions worthy of future research. As mentioned previously, we think it would be valuable to know even more about the differential effects of STRIVE for various learners with disabilities. As this study was limited to fourth grade, we think it is worthwhile to better understand the extent to which these findings might generalize to other grade groupings. Also, the content area selected for this study was social studies, and future research questions examining effects within other content areas, such as science, are worth addressing.

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 48. https://doi.org/10.18637/jss.v067.i01

Capin, P., Stevens, E. A., Stewart, A., Swanson, E., & Vaughn, S. (2020). Examining vocabulary,

comprehension, and content knowledge instruction during fourth grade social studies teaching. *Reading and Writing: An Interdisciplinary Journal*. Advance online publication. https://doi.org/10.1007/s11145-020-10106-5

Coburn, C. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher*, *32*(6), 3–12. https://doi.org/10.3102/0013189X032006003

Coburn, C. E., Russell, J. L., Kaufman, J. H., & Stein, M. K. (2012). Supporting sustainability: Teachers advice networks and ambitious instructional reform. *American Journal of Education*, *119*(1), 137–182. http://doi.org/10.1086/667699

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute.

Espin, C. A., Busch, T. W., Shin, J., & Kruschwitz, R. (2001). Curriculum-based measurement in the content areas: Validity of vocabulary-matching as an indicator of performance in social studies. *Learning Disabilities Research & Practice*, *16*(3), 142–151. https://doi.org/10.1111/0938-8982.00015

Firmender, J. M., Reis, S. M., & Sweeny, S. M. (2013). Reading comprehension and fluency levels ranges across diverse classrooms: The need for differentiated reading instruction and content. *Gifted Child Quarterly*, *57*(1), 3–14. https://doi.org/10.1177/0016986212460084

Gagné, F. (2005). From noncompetence to exceptional talent: Exploring the range of academic achievement within and between grade levels. *Gifted Child Quarterly*, *49*, 139–153. https://doi.org/10.1177/001698620504900204

Gajria, M., Jitendra, A. K., Sood, S., & Sacks, G. (2007). Improving comprehension of expository text in students with LD: A research synthesis. *Journal of Learning Disabilities*, *40*(3), 210–225. https://doi.org/10.1177/00222194070400030301

Garet, M. S., Cronen, S., Eaton, M., Kurki, A., Ludwig, M., Jones, W., Uekawa, K., Falk, A., Bloom, H. S., Doolittle, F., Zhu, P., Sztejnberg, L., & Silverberg, M. (2008). *The impact of two professional development interventions on early reading instruction and achievement* (NCEE 2008-4030). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Garet, M. S., Porter, A., Desimone, L., Birman, B., & Yoon, K. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, *38*(4), 915–945. https://doi.org/10.3102/00028312038004915

Gersten, R., Baker, S. K., Shanahan, T., Linan-Thompson, S., Collins, P., & Scarcella, R. (2007). *Effective literacy and English language instruction for English learners in the elementary grades: IES practice guide* (NCEE 2007-4011). What Works Clearinghouse.

Gersten, R., Dimino, J., Jayanthi, M., Kim, J. S., & Santoro, L. E. (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*(3), 694–739. https://doi.org/10.3102/0002831209361208

Gilmour, A. F., Fuchs, D., & Wehby, J. H. (2019). Are students with disabilities accessing the curriculum? A meta-analysis of the reading achievement gap between students with and without disabilities. *Exceptional Children*, *85*(3), 329–346. https://doi.org/10.1177/0014402918795830

Hargreaves, A., & Goodson, I. (2006). Educational change over time? The sustainability and non-sustainability of three decades of secondary school change and continuity. *Educational Administration Quarterly*, *42*(1), 3–41. https://doi.org/10.1177/0013161X05277975

Hedges, L. V. (2011). Effect sizes in three-level cluster-randomized experiments. *Journal of Educational and Behavioral Statistics*, *36*(3), 346–380. https://doi.org/10.3102/1076998610376617

Hoffman, L. (2015). *Longitudinal analysis: Modeling within-person fluctuation and change*. Routledge. https://doi.org/10.4324/9781315744094

Klingner, J. K., Vaughn, S., Arguelles, M. E., Tejero Hughes, M., & Ahwee Leftwich, S. (2004). Collaborative strategic reading: "Real-world" lessons from classroom teachers. *Remedial and Special Education*, *25*(5), 291–302. https://doi.org/10.1177/07419325040250050301

Klingner, J. K., Vaughn, S., Boardman, A., & Swanson, E. (2012). *Now we get it! Boosting comprehension with collaborative strategic reading*. Jossey-Bass.

Lee, C. D., & Spratley, A. (2010). *Reading in the disciplines: The challenges of adolescent literacy*. Carnegie Corporation of New York.

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2020). emmeans: Estimated

marginal means. R package Version 1.4. 4 [Computer software].

MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L, G. (2000). *Gates-MacGinitie Reading Tests* (4th ed.). Riverside.

Marinak, B., & Gambrell, L. B. (2008). Elementary informational text instruction: A research review. *International Journal of Learning*, *15*(9). https://doi.org/10.18848/1447-9494/cgp/v15i09/45930

McFarland, J., Hussar, B., Zhang, J., Wang, X., Wang, K., Hein, S., Diliberti, M., Cataldi, E. F., Mann, F. B., Barmer, A., Nachazel, T., Barnett, M., & Purcell, S. (2019). *The condition of education 2019* (NCES 2019-144). National Center for Education Statistics.

McLaughlin, M., & Mitra, D. (2001). Theory-based change and change-based theory: Going deeper, going broader. *Journal of Educational Change*, *2*(4), 301–323. https://doi.org/10.1023/A:1014616908334

National Council for the Social Studies. (2013). *College, career, and civic life (C3) framework for social studies state standards*.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards*.

Pashler, H., Bain, P. M., Bottge, B. A., Graesser, A., Koedinger, K., McDaniel, M., & Metcalfe, J. (2007). *Organizing instruction and study to improve student learning: IES practice guide* (NCER 2007-2004). National Center for Education Research.

Payne, G., & Payne, J. (2004). Secondary analysis. In G. Payne & J. Payne, *Sage key concepts:Key concepts in social research* (pp. 214–218). Sage. https://doi.org/10.4135/9781849209397

Porche, M., Pallante, D., & Snow, C. (2012). Professional development for reading achievement. Results from the Collaborative Language and Literacy Instruction Project (CLLIP). *Elementary School Journal*, *112*(4), 649–671. https://doi.org/10.1086/665008

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*(3), 879–891. https://doi.org/10.3758/brm.40.3.879

Reis, S. M., Gubbins, E. J., Briggs, C. J., Schreiber, F. J., Richards, S., Jacobs, J. K., Eckert, R. D., & Renzulli, J. S. (2004). Reading instruction for talented readers: Case studies documenting few opportunities for continuous progress.

Gifted Child Quarterly, *48*(4), 315–338. https://doi.org/10.1177/001698620404800406

Richardson, V. (2003). The dilemmas of professional development. *Phi Delta Kappan*, *84*(5), 401–406. https://doi.org/10.1177/003172170308400515

Scruggs, T. E. (2012). Differential facilitation of learning outcomes: What does it tell us about learning disabilities and instructional programming? *International Journal for Research in Learning Disabilities*, *1*, 4–20.

Scruggs, T. E., Mastropieri, M. A., Berkeley, S., & Graetz, J. (2010). Do special education interventions improve learning of secondary content? A meta-analysis. *Remedial and Special Education*, *36*, 437–449. https://doi.org/10.1177/0741932508327465

Shanahan, T., Callison, K., Carriere, C., Duke, N. K., Pearson, P. D., Schatschneider, C., & Torgesen, J. (2010). *Improving reading comprehension in kindergarten through 3rd grade: IES practice guide* (NCEE 2010-4038). What Works Clearinghouse.

Simmons, D., Hairrell, A., Edmonds, M., Vaughn, S., Larsen, R., Willson, V., Rupley, W., & 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*(2), 121–156. https://doi.org/10.1080/19345741003596890

Stevens, E. A., Vaughn, S., Swanson, E., & Scammacca, N. (2020). Examining the effects of a Tier 2 reading comprehension intervention aligned to Tier 1 instruction for fourth-grade struggling readers. *Exceptional Children*, *86*(4), 430–448. https://doi.org/10.1177/0014402919893710

Swanson, E., Hairrell, A., Kent, S., Ciullo, S., Wanzek, J. A., & Vaughn, S. (2014). A synthesis and meta-analysis of reading interventions using social studies content for students with learning disabilities. *Journal of Learning Disabilities*, *47*(2), 178–195. https://doi.org/10.1177/0022219412451131

Swanson, E., Solis, M., Ciullo, S., & McKenna, J. W. (2012). Special education teachers' perceptions and instructional practices in response to intervention implementation. *Learning Disability Quarterly*, *35*(2), 115–126. https://doi.org/10.1177/0731948711432510

Swanson, E., Stevens, E., A., Scammacca, N., Capin, P., Stewart, A., & Austin, C. (2017). The impact of Tier-1 reading instruction on reading

outcomes for students in Grades 4–12: A meta-analysis. *Reading and Writing*, *30*(8), 1639–1665. https://doi.org/10.1007/s11145-017-9743-3

Swanson, E., Stewart, A., Stevens, E. A., Scammacca, N., Capin, P., Hamilton, B. J., Roberts, G., & Vaughn, S. (2021). *The efficacy of two models of professional development mediated by fidelity on fourth grade student reading outcomes*. [Manuscript submitted for publication]. Meadows Center for Preventing Educational Risk, The University of Texas at Austin.

Swanson, E., Wanzek, J., McCulley, L., Stillman-Spisak, S., Vaughn, S., Simmons, D., Fogarty, M., & Hairrell, A. (2016). Literacy and text reading in middle and high school social studies and English language arts classrooms. *Reading & Writing Quarterly*, *32*(3), 199–222. https://doi.org/10.1080/10573569.2014.910718

Swanson, E., Wanzek, J., Vaughn, S., Roberts, G., & Fall, A. M. (2015). Improving reading comprehension and social studies knowledge among middle school students with disabilities. *Exceptional Children*, *81*(4), 426–442. https://doi.org/10.1177/0014402914563704

Swanson, E., Vaughn, S., & Roberts, G. (2018). *Examining the efficacy of differential levels of professional development for teaching content area reading strategies*. Grant awarded by the U.S. Department of Education, Institute of Education Sciences to The University of Texas at Austin.

Tudge, J., & Scrimsher, S. (2003). Lev S. Vygotsky on education: A cultural-historical, interpersonal, and individual approach to development. In B. J. Zimmerman & D. H. Schunk (Eds.), *Educational psychology: A century of contributions* (pp. 207–228). Lawrence Erlbaum.

Vaughn, S., Fall, A.-M., Roberts, G., Wanzek, J., Swanson, E., & Martinez, L. R. (2019). Class percentage of students with reading difficulties on content knowledge and comprehension. *Journal of Learning Disabilities*, *52*(2), 120–134. https://doi.org/10.1177/0022219418775117

Vaughn, S., Roberts, G., Swanson, E., Fall, A-M., & Stillman-Spisak, S. (2014). Improving middle school students' knowledge and comprehension in social studies: A replication. *Educational Psychology Review*, *27*(1), 31–50. https://doi.org/10.1007/s10648-014-9274-2

Vaughn, S., Swanson, E., Fall, A-M, Roberts, G., Capin, P., Stevens, E. A., & Stewart, A. A. (in press). The efficacy of school and researcher provided professional development on English learners' literacy. *Journal of Educational Psychology*.

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

Wanzek, J., Roberts, G., Vaughn, S., Swanson, E., & Sargent, K. (2019). Examining the role of pre-instruction academic performance within a text-based approach to improving student content knowledge and understanding. *Exceptional Children*, *85*(2), 212–228. https://doi.org/10.1177/0014402918783187

Wanzek, J., Swanson, E., Vaughn, S., Roberts, G., & Fall, A. M. (2016). English learner and non-English learner students with disabilities: Content acquisition and comprehension. *Exceptional Children*, *82*(4), 428–442. https://doi.org/10.1177/0014402915619419

What Works Clearinghouse. (2020). Standards handbook: Version 4.1. *Institute of Education Sciences*. Retrieved from https://ies.ed.gov/ncee/wwc/Handbooks

Authors' Note

ORCID iDs

Elizabeth Swanson [iD] https://orcid.org/0000-0002-2716-4078

Elizabeth A. Stevens [iD] https://orcid.org/0000-0002-8412-1111

Supplemental Material

Supplemental material for this article is available online.