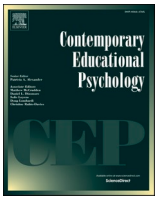
Contemporary Educational Psychology 73 (2023) 102152

Contents lists available at ScienceDirect

Contemporary Educational Psychology

journal homepage: www.elsevier.com/locate/cedpsych



Professional development for evidence-based SRSD writing instruction: Elevating fourth grade outcomes

Debra McKeown [a,*], Kay Wijekumar [a], Julie Owens [b], Karen Harris [c], Steve Graham [c], Puiwa Lei [d], Erin FitzPatrick [e]

[a] Texas A&M University, United States
[b] University of Tennessee-Knoxville, United States
[c] Arizona State University, United States
[d] Pennsylvania State University, United States
[e] University of North Carolina at Charlotte, United States

ARTICLE INFO

Keywords:
Writing
Professional development
Practice-based professional development
Self-regulated strategy development
SRSD

ABSTRACT

Writing is a critical skill for success in all areas of life, but it is one of the least taught skills in school. Teachers consistently report being unprepared to teach writing. In this study, set in a Southern U.S. boomtown, teachers received two days of practice-based professional development for a ten-week implementation of self-regulated strategy development (SRSD), an evidence-based writing intervention, to support student persuasive and informational writing as well as performance on the state standardized writing exam. This multi-site cluster randomized controlled study evaluated the effectiveness of SRSD on student writing outcomes including prompt adherence, elements, and holistic quality. Multilevel modeling analysis was used to evaluate data from 418 fourth -grade students (256 treatment, 162 control) nested across 33 classes (n = 17 treatment taught by 8 departmentalized teachers; 16 control, 9 departmentalized teachers) within 11 schools randomly assigned to condition. Teachers implemented SRSD with high fidelity (M = 92%; range 91–100%). SRSD had a statistically significant and large effect on prompt adherence ($p < .001$; Hedges' $g = 1.87$), elements ($p < .001$; Hedges' $g = 0.84$) and holistic scores ($p < .001$; Hedges' $g = 0.87$), while holding gender and pretest scores constant. Effects of SRSD on all writing measures were not significantly moderated by students' gender, students' pretest scores, or schools' pretest scores. There were complications with teacher observations, especially related to technology. Limitations and future directions are discussed.

1. Introduction

Writing is a critical skill for academic, social, and professional success (Graham, 2006). Poor writing abilities make it difficult for students to use writing effectively as a tool for learning, communication, and self-expression. Unfortunately, high-stakes assessments show a pattern of poor performance on writing for all students at key grade levels in the United States. Over 60% of fourth -grade students in Texas, where this study took place, and 73% of eighth-grade students nationally failed to score proficient or above on the writing test (National Center for Education Statistics [NCES], 2011; *State of Texas Assessments of Academic Readiness* [STAAR], 2017).

Below, we discuss the current state of writing instruction and how it is impacted by state standardized writing assessments. We describe an evidence-based framework for teaching writing, self-regulated strategy development (SRSD). Next, we address the impact of stressful contexts on teachers' acquisition of new skills. Finally, we discuss practice-based professional development (PBPD), a method of supporting teachers to adopt evidence-based practices which is uniquely responsive to the reported lack of pre-service preparation and contextual challenges.

2. Current state of writing instruction

In a content analysis of textbooks and writing curricula used in upper elementary classrooms, Wijekumar et al., (2019) found that most focused on compartmentalized instruction of component skills (e.g., grammar, punctuation, spelling), lacked in-depth instruction on the planning process and ideation, and did little to promote efficacy toward

* Corresponding author.
E-mail address: debramckeown@tamu.edu (D. McKeown).

https://doi.org/10.1016/j.cedpsych.2023.102152

Available online 21 January 2023

writing. The lessons often included worksheets on vocabulary, spelling, and grammar, and teachers reported that most writing instruction occurred a month prior to the state standardized test. This is consistent with prior findings which suggest that when teachers provided writing instruction, it was rarely evidence-based and did not focus on the process of writing (e.g., planning, revising) but instead emphasized brief tasks that did not require collaboration or higher order skills such as synthesis (Troia & Olinghouse, 2013). Graham (2019) refers to this as *writing without composing*.

Generalization. Students also need instruction in each genre in which they are expected to write. Reviews of research has shown that the quality of students' writing vary between genres and writing well in one genre does not transfer to a different genre (Graham, Harris, & Hebert, 2011a). While research is fairly clear that it takes approximately 8–12 weeks to learn a single genre using evidence-based writing instruction (see Harris et al., 2022), we do not know if there are differences for learning subsequent genres.

High-stakes testing. Further complicating matters, high stakes assessments have been associated with a narrowing of the curriculum and a disproportionate focus on tested content areas (National Council of Teachers of English, 2014; Scherff & Piazza, 2005). Yet, despite consistent low writing performance on standardized assessments, writing instruction receives the least attention of any subject in elementary schools; primary grade teachers in the United States reported teaching writing for fewer than 10 min per day and the amount of time decreased with students in upper grades (e.g., Brindle et al., 2016). Nearly 3/4 of Grade 8 students reported spending fewer than 30 min per day writing, but those who reported writing more performed better on the national assessment (NCES, 2011).

2.1. Self-regulated strategy development

Writing is one of the most cognitively demanding tasks asked of students (Alevriadou & Giaouri, 2015). Writing has been explained as the coordinated acts of planning, translating, and reviewing, moderated by long-term memory, working memory, and meta-cognition, all situated within a complex context and community (Flower & Hayes, 1981; Graham, 2018). One evidence-based framework for writing instruction that addresses these multiple demands is self-regulated strategy development (SRSD). SRSD is a criterion-based intervention. It includes explicit, interactive strategies instruction for both general and genre-specific writing. Students learn the writing process (e.g., planning, drafting, evaluating), self-regulation strategies (i.e., goal setting, self-assessment, self-instructions, and self-reinforcement), and the knowledge needed to use them. Students track their improvement and set new, more challenging goals as they continue their development as writers (Harris et al., 2006; 2008). Teachers provide explicit instruction in how to generalize the strategies and understandings to other appropriate writing purposes and settings (Harris et al., 2008). Instruction takes place across six flexible, recursive stages, with gradual release of responsibility to students (Harris et al., 2008; McKeown et al., 2021).

Over 100 experimental studies of SRSD have been conducted to evaluate its impact on writing instruction (Harris & Graham, 2018). Research indicates that SRSD is effective with students who represent the full range of ability in a typical elementary class and that SRSD has the strongest impact of any strategies instruction approach in writing (Graham et al., 2012/2018; Graham et al., 2012). From a meta-analysis of SRSD studies in elementary schools, SRSD's average weighted effect size for writing quality was 1.17 ($n = 14$; Graham et al., 2012).

2.2. Supporting effective writing instruction in stressful contexts

The context for the present study is a boomtown. Teachers in boomtowns are particularly vulnerable as external and internal challenges are magnified in these environments. A boomtown is associated with low unemployment rates and economic stratification. Due to the relatively higher salaries paid in the booming industry, these locations have high housing prices, limited housing options, and difficulty attracting and keeping service workers, including teachers (Jaquet, 2009). Furthermore, local services, including schools, do not grow as rapidly as the number of people in the community, which can result in crowded schools (Jaquet, 2009). School personnel have reported that families responding to the booming industry are often characterized by higher rates of transience and disrupted educational experiences (Schafft et al., 2014). Moreover, Schafft (2014) found that though the local economy was booming, few study participants were realizing any economic benefits within the local schools. This context can keep teachers from coming to or remaining in the area, and those who stay are often subjected to substandard housing, long commutes, economic anxiety, and high levels of stress (Dizon-Ross et al., 2018).

Stress has a negative effect on working memory and increases cognitive load (e.g., Klein & Boals, 2001). Research indicates that when teachers are under stress, new to the profession, and/or exceed their working memory, more recently learned knowledge or strategies are inhibited and reversion to cognitive default occurs (Clark, 2001). That is, older, more reinforced (and presumably less effective) knowledge is more likely to be used as it takes less effort (Feldon, 2007). This is problematic as teachers report a lack of pre-service preparation to teach writing leaving them without adequate schema to draw upon (Brindle et al., 2016; Graham, 2019). Reviews of required literacy coursework within both general and special education teacher preparation programs consistently indicate a lack of writing instruction (Brenner & McQuirk, 2019; Chandler & Sayeski, 2022). However, expertise can be gained, in part, by purposeful practice.

2.3. Practice-based professional development

Practice-based learning has been used in professions such as law, medicine, and education (Billett, 2010). Researchers have used the PBPD model to help teachers implement evidence-based writing practices with fidelity (Harris et al., 2012; Harris et al., 2022). In this study, we incorporated the features of PBPD described by Ball and Cohen in the following ways: (a) *actively engaging with colleagues of similar needs*: teachers from the same grade and school learned, implemented, and planned SRSD instruction together; (b) *contextualizing professional development within the teachers' current classrooms including strengths, needs, and characteristics of present students:* throughout PBPD, student needs were discussed and planned for; teachers were asked to write example essays that included local names, places, and situations; (c) *addressing gaps in content knowledge through assessment and tailored instruction:* informal assessments were used to gauge writing instruction knowledge, beliefs, and practices; content was tailored to address identified needs; (d) *practicing implementation of the lessons and receiving feedback from peers:* teachers taught all lessons with manualized support; (e) *implementing instruction with identical materials that will be used in the classroom:* teachers were provided all materials needed and used those during practice; and (f) *receiving expert feedback on performance and considering how to differentiate instruction for their students while burgeoning skills and understanding can be still be addressed prior to implementation in the classroom:* teachers received feedback on practice and considered opportunities for differentiation for their current students (Ball & Cohen, 1999; Harris et al., 2012; Loewenberg Ball & Forzani, 2009).

Evaluations of effective professional development have indicated that *follow-up support* and *sufficient duration* are also key components (Darling Hamond et al., 2017). PBPD for SRSD has always included follow-up support (Harris et al., 2012) and multiple studies have included ongoing coaching (e.g., McKeown et al., 2016; 2019). Results from>20 studies of SRSD demonstrate that 14–16 h of PBPD is a sufficient duration to result in high implementation fidelity and positive student writing outcomes (Harris et al., 2022).

2.4. PBPD for SRSD to address state writing assessments

To address writing demands meaningful to students and teachers, researchers have tailored SRSD instruction in narrative (Kiuhara et al., 2013; McKeown et al., 2016), persuasive (McKeown et al, 2019), and text-based informational writing (FitzPatrick & McKeown, 2021) to prepare students for state-level standardized assessments. Kiuhara et al. (2013) and McKeown et al. (2016) conducted studies supporting teachers in crafting fantastical narratives that feature the author as narrator – a unique testing genre at the time – and in both studies, student writing outcomes improved and teachers were highly satisfied with the intervention. McKeown and colleagues (2019) conducted a wait-listed quasi-experimental study in which 25 teachers were provided PBPD to implement SRSD writing instruction in the persuasive genre with 685 students. Fidelity in this study was lower than typical in SRSD studies, and teachers attributed that in part to the instruction they were required to do for more prioritized end-of-grade assessments in reading and mathematics. Still, student writing outcomes were enhanced following the intervention, and teachers reported high social validity for both PBPD and SRSD. FitzPatrick and McKeown (2021) conducted a multiple-probe single case experimental design to evaluate PBPD for SRSD writing instruction for the informational genre requiring the citation of text-based evidence with one fifth grade special education teacher's implementation across three small groups of students with and without learning disabilities in an inclusive general education classroom. Again, student writing outcomes uniformly increased and both teachers and students rated the intervention highly on social validity measures.

2.5. Present study

In the present study, we aimed to prepare students for the STAAR writing assessment. We used a cluster randomized controlled trial situated within a large school district in the southern United States, 17 English Language Arts teachers serving 418 students across 33 homerooms were randomly assigned at the school level (n = 11) to either treatment or control.

The STAAR exam emphasizes reading and math skills, and at the time of the study, assessed writing in Grades 4 and 7. To help fourth grade teachers and students prepare for the composition task, we analyzed the past published exams and created responsive writing lessons using the SRSD framework. Students and teachers have found the persuasive genre to be easier to learn than narrative or informational writing (Harris et al., 2012). Therefore, instruction began with persuasive and, once mastered, pivoted to the state required writing task – a prompt identified as expository, and more precisely described as a personal informational essay (e.g., *Why are friends important?*). This writing task aligns well with the persuasive genre as it includes a personal opinion or perspective followed by substantiation, but without an expectation to convince the reader. We believe students can learn this new genre quickly once they have learned the persuasive genre.

Researchers provided PBPD in the persuasive genre as well as how to pivot from persuasive to the expository genre required on the state test. Then eight intervention teachers implemented the SRSD lessons with the full range of students in their inclusive fourth grade classrooms across ten weeks. Researchers observed instruction in person and via video for fidelity of implementation and to provide feedback and support. Student writing outcomes were tested before and after intervention using a prompt similar to the state standardized assessment. Essays were scored for prompt adherence, genre elements, and holistic quality. The prompt and genre were chosen for the strengthened external validity; the schools' primary goal was to improve student performance on the state standardized test.

The SRSD writing intervention for in this study included a new background knowledge component (genre overview, Lesson 0) at the beginning and a transfer task for the state standardized test (Lesson 8) at the end. We added these two lesson components to be more responsive to stated district needs. To avoid earning a score of 0 on the standardized writing assessment, students must adhere to all aspects of the assigned writing task. Therefore, it is important to provide explicit instruction in how to analyze a writing task to detect what is expected, such as the topic, genre (in school, usually persuasive/opinion, narrative, or expository/informational), form (e.g., essay, letter, speech), length, and other expectations (e.g., citing source text). Students need explicit instruction in writing for standardized tests as students do not typically generalize across writing tasks without instruction (Harris et al., 2009; Santangelo & Olinghouse, 2009).

This is the initial study in a series of iterations within a large-scale project called We Write funded by the Institute of Educational Sciences. We Write was designed with four major components: (a) PBPD, (b) teacher-led SRSD writing lessons including two new lessons to prepare for the state writing exam, (c) web-based fidelity of implementation and (d) web-based extension writing lessons for students. The purpose of the present study was to test the effectiveness of the first three components prior to incorporating the web-based lessons.

2.6. Research questions

RQ1: Following PBPD for SRSD, to what degree did teachers implement with fidelity?

Hypothesis 1

We predict teachers will implement with high fidelity based on prior studies (Harris et al., 2022). In PBPD, teachers see others teaching the SRSD lessons multiple times, practice teaching the lessons, and receive feedback. Teachers are provided a manual with complete lesson plans that include meta-scripts to guide them during training and serve as a reference. They are also provided support in terms of materials, collaborating peers, and experts. All of these components are expected to result in high teacher fidelity of implementation which would be consistent with prior studies of PBPD for SRSD (Harris et al., 2022).

RQ2: To what degree does SRSD instruction across persuasive and expository writing improve 4th grade students' performance on a simulation of the Texas state writing task in terms of prompt adherence, genre elements, and holistic quality? Additionally, is the effect of SRSD writing instruction on each of the three outcome measures stronger relative to gender, socioeconomic status, English learner status, disability status and/or pretest performance at the individual, class, and school level?

Hypothesis 2

While past research has shown students to do not generalize from one genre to another, we believe that providing brief instruction on a new, related genre (expository - personal informational essay following persuasive essay instruction) will result in increases in expository writing quality. Following SRSD instruction, we predicted student performance would increase on prompt adherence, included genre elements, and holistic quality. While being on topic has not been separately reported in large comparison studies, single case design studies have shown an increase in on topic writing and prompt adherence following SRSD (FitzPatrick & McKeown, 2020, 2021; McKeown et al., 2016). Prior teacher-implemented SRSD instruction following PBPD has resulted in increased writing performance in genre elements and holistic quality (Festas et al., 2015; Harris et al., 2012; McKeown et al., 2016, 2018, 2019). Research has indicated gender is a predictor of significant variation in writing outcomes with females performing higher than males (Berninger & Fuller, 1992; NCES, 2011). We predicted the same will be the case in this study. Based on individual and group performance in past studies of teacher-implemented SRSD, we predict intervention participants and groups with lower pretest scores will grow more than those with higher pretest scores.

To benefit future studies, an additional purpose was to inform and establish standards and procedures for the use of video devices (SWIVL, iPad, Zoom) to measure implementation fidelity and provide coaching/

modeling.

3. Method

3.1. Setting

This study took place in a Southern US school district situated in a boomtown. The district served students who were primarily Hispanic – over 60%; approximately 25% White, 7% Black, 3% Asian, 2% two or more races, 1% Pacific Islander, <1% Native American. Approximately half were eligible for a free or reduced-price lunch, an indicator of socio-economic status. The fourth grade performance in previous STAAR test administrations at the district level showed approximately 30% of students met or mastered expectations on the state writing assessment.

3.2. Participants

Institutional Review Board approval was obtained for the ethical conduct of human subject research. We received informed teacher consent, parent permission, and student assent from all study participants. Within the district, all fourth grade teachers and their classes were invited to participate in the study. Fourth grade bilingual classroom teachers were invited to attend PBPD and receive support, but were not included in the study. A wait-listed design was used to avoid deprivation of treatment, so all participants had the opportunity to receive the PBPD, but at different points in time. As the intervention is considered an evidence-based practice and was theorized to positively impact student writing scores, we assigned more schools to the intervention group for ethical reasons. The study began with 15 schools randomly assigned to treatment groups (intervention = 9, comparison = 6). In the final analysis, there were 11 schools (intervention = 6, control = 5), 8 intervention teachers (two were teaching pairs) representing 17 homeroom classes, 9 control teachers representing 16 homeroom classes, and 418 students. See Fig. 1 for details regarding attrition.

Of the 43 recruited classes, 6 classes missed pretest and 5 classes missed posttest. Missing pretest at the student level was not significantly associated with treatment condition. The intervention group had a higher percentage of students (28.29%) that missed the posttest than the

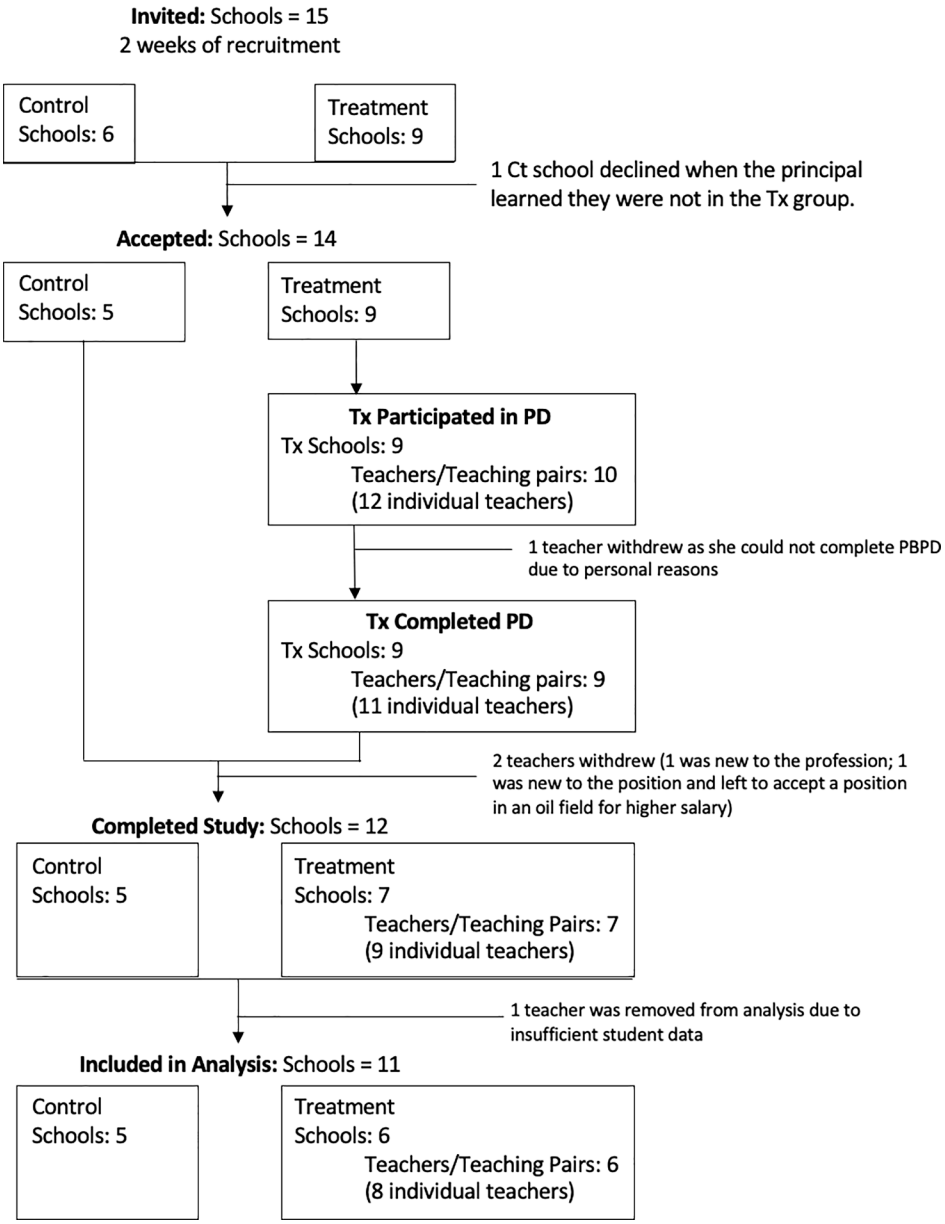


Fig. 1. Attrition of participants across events by condition.

control group (19.17%), although missing posttest was not significantly associated with student pretest performance. Missing pretest or posttest at the class level was not significantly associated with treatment conditions and missing posttest was not associated with class pretest performance.

There were two teaching pairs in the intervention group. These teachers participated in PBPD together, always taught the SRSD lessons together as co-teachers, and their classes were always combined for these lessons. All intervention teachers taught in a departmentalized program and taught writing for all of fourth grade within their schools. Intervention teachers were asked to complete a demographic form during PBPD. All teachers in the final intervention group were female (White = 7, Hispanic = 1) with a Bachelor's degree. Four teachers had one or more Master's degrees, two of which were earned online. Degrees were earned between 1982 and 2017. All teachers were certified to teach fourth grade literacy. They had an average of 16 (range = 7–37) years teaching, 5 years in the current grade (range = 1–10), 2.5 years in the current school and current grade level (range = 1–15). Four teachers reported having leadership experience. Teachers were asked if they volunteered to attend the PBPD. Four reported yes, but upon further discussion, all had been asked by their principal to participate. Teachers in the control condition did not complete a demographic survey.

There were 418 students in the final analysis with 162 in the control group and 256 in the intervention group. Table 1 shows student demographic variables by treatment condition. There were no statistically significant imbalances between treatment and control schools on any of the student level variables.

4. Measures

4.1. Implementation fidelity

Measuring fidelity helps to improve the internal validity of a study by establishing that the intervention is likely responsible for any differences in outcome between groups (Dane & Schneider, 1998). In a recent review of studies of PBPD for SRSD, researchers found that 17 of the 19 studies reported fidelity of implementation and ranged from 74% to 98% (M = 90%; Harris et al., 2022). To measure fidelity of implementation in the present study, trained observers used a checklist containing the core components for each lesson (available from first author). Teachers were evaluated on the completion of each component (0 or 1). Throughout PBPD, teachers were encouraged to use their expertise to adapt the instruction for current students and context, without omitting the core components of the intervention. Thus, a change in order, repeating information, or choosing to skip optional steps was not penalized because SRSD is not scripted.

Fidelity observations were conducted both in person or via video, based on the preference of each teacher. Fidelity was calculated as a percentage (number of steps completed divided by number of steps intended). Seventy observations (approximately 20% of instructional sessions) were conducted on participating teachers across 10 weeks and 20% (n = 14) were conducted by two observers. Agreement was calculated by dividing the total number of agreements by the total number of ratings. Agreement was 100%.

For video observations, we provided teachers with a SWIVL™ device, iPad, and an app to stream and capture video. The iPad was used to capture the video and run the streaming/video capture app (Zoom). The SWIVL device attached to a tripod and held the iPad. The teacher wore a microphone on a lanyard which also served as a tracking device. The SWIVL device rotated to follow the tracking device, thus turning the iPad in direction the teacher was located. Videos were uploaded to secure, encrypted cloud storage. Since an additional purpose of this study was to inform and establish standards and procedures for the use of video devices to measure fidelity, researchers documented issues and solutions with procedures and technology.

Table 1
Descriptive Statistics of Demographic Variables by Condition.

Variables	Intervention		Control	
	Mean (SD)	N	Mean (SD)	N
Total number student participants		256		162
% Female students	48.8	130	48.4	81
% Male students	51.18 (50.08)	124	51.59 (50.13)	76
% Missing gender		2		5
Total number schools		6		5
% approaching reading proficiency	72.83 (7.91)	6	75.00 (8.69)	5
% approaching writing proficiency	61.33 (8.62)	3	67.67 (1.53)	3
Class size	21.53 (1.91)	6	20.90 (3.89)	5
Total enrollment	602 (130.02)	6	570 (136.30)	5
% White	35.46 (9.56)	6	31.28 (12.20)	5
% Hispanic	48.68 (10.33)	6	56.54 (12.18)	5
% African American	6.90 (1.52)	6	8.30 (3.94)	5
% Economic disadvantaged	38.78 (12.90)	6	39.20 (11.80)	5
% English learners	20.05 (28.08)	6	13.44 (13.59)	5
% Disability	36.00 (12.54)	6	41.40 (6.31)	5

Note: % approaching proficiency means those who do not meet proficiency on the state standardized test

4.2. Writing prompts

A team of trained research assistants completed pre- and post-testing procedures with each school, hosting groups of students in a large common area such as the cafeteria or library. The teams used a script to provide instructions to the students on how to complete assessment activities. Students were given 35 min to write in response to a prompt. Later, handwritten responses were typed to control for handwriting bias (Graham, Harris, & Hebert, 2011b; Graham et al., 2011).

Students wrote in response to an expository prompt pulled from released STAAR items. The prompts were chosen because of their parallel form and universal appeal. Since 4th grade was the only elementary school grade in which writing was tested, the students were expected to have no prior exposure to these prompts. Researchers flipped a coin to determine which prompt was given at pretest. In cases where random assignment is used, like in this study, counterbalancing testing probes is not necessary. The pretest and posttest prompts, respectively, for all students were:

READ the information in the box below: In the book *Oh, the Places You'll Go!,* Dr. Seuss writes, "Today is your day! Your mountain is waiting. So … *get on your way!*" **THINK** about all the different places you could visit or things you could do in the future. **WRITE** about something that you look forward to doing. Tell what you want to do and explain why you want to do it.

READ the information in the box below: In the movie *The Wizard of Oz*, Dorothy says, "There's no place like home." **THINK** about a place that is special to you. **WRITE** about your favorite place to spend time. Describe the place and explain what makes it special.

At the end of each prompt, a bulleted list of reminders was included that read "Be sure to clearly state your central idea, organize your writing, develop your writing in details, choose your words carefully, and use correct spelling, capitalization, punctuation, grammar, and sentences."

On Topic. Essays were evaluated for being on topic (0 or 1). Pretests required a reference to something that the student looked forward to doing, and posttests had to include a favorite place to spend time. Essays that scored a 1 for on topic were scored further.

Prompt Adherence. This is a measure of the degree to which

students discern and respond to all requirements of a writing prompt (score = 0–3). Both prompts (pretest and posttest) had two parts (i.e., Tell what you want to do and explain why you want to do it; Describe the place and explain what makes it special). Students earned a point for addressing each of the two parts (0, 1, or 2). They earned another point for responding in the required genre form (for a total of 3 points possible). For example, a narrative with dialogue written in response to the expository prompt would be a 0 on this dimension. On topic and prompt adherence were scored at the same time. During training, inter-rater reliability (IRR) across on topic and prompt adherence was > 0.95 for all scoring pairs; 51% of all essays were scored by two raters. IRR was 0.96.

Genre Elements. All essays were evaluated for the number and quality of basic genre elements of expository essays (0 – 19 possible points). This was the most proximal and sensitive measure as it directly reflects the instruction provided and captures the ability of students to pivot from persuasive (the focus of most of the instruction) to informational. Topic sentence (that adhered to the prompt), up to three ideas (that made sense and supported topic sentence), and conclusion (that made sense) were each scored 0 if not present, 1 if clearly present and adhered to the prompt, and 2 if refined/interesting/elaborated. Up to three supporting details were scored 0 if not present and 1 point each if present, made sense, and supported the ideas or topic sentence. For transition words/phrases, interesting vocabulary, and writer's voice, students were given 1 point for two instances and 2 points for three or more instances. All essays were scored by a pair of trained raters who were provided a code book containing an explanation for each score. During training, IRR was > 0.90 for all scoring pairs. Reliability of scoring was conducted on 51% of the essays. IRR was 0.90. After the pair discussed discrepancies, agreement was 0.93.

Holistic Quality. Holistic evaluation is the most common procedure for scoring writing quality in writing intervention studies (Graham & Perin, 2007). Trained scorers rated each essay on a scale between 0 and 6 where 0 was no response or off topic and 6 represented the highest quality work. Scorers were trained to give equal consideration to organization, development, sentence fluency, word choice, audience awareness, genre parts, flow, and voice. The scoring group identified one or more anchor papers for each score which were used as reference during scoring. Raters were blind to the purpose of the study, condition, and time of test being scored (i.e., pre- or posttest). Each rater independently scored 70% of all assigned essays resulting in 40% of all essays being scored twice for IRR. During training, IRR was 0.90. After all scoring was complete, IRR was 0.94.

4.3. Procedures

The components of the study took place in the following order: First, the research team presented a two-days of PBPD to the intervention teachers. Next, students completed pretests at their school. Then, teachers delivered the SRSD intervention in the treatment schools for approximately 10 weeks, from January to mid-April. At the conclusion of SRSD instruction, students completed a posttest.

4.4. Practice-based professional development

PBPD is important to high fidelity implementation of SRSD and for this study, was delivered face-to-face by the research team. The team followed an agenda and checked off components as they were delivered to maintain fidelity of implementation for professional development. All PBPD steps were completed (100%).

Teachers participated in two consecutive days of PBPD (16 h) at the district office in January, the first week after returning from holiday break. Teachers were well acquainted with their students at this point in the year. Researchers led PBPD. They provided teachers an intervention manual which contained lesson plans, student handouts, logistical information, and other materials such as student writing folders and posters.

Teachers were taught the foundations of SRSD for writing. Researchers built background knowledge about writing and teaching writing, using academic vocabulary and explaining the underlying theory. We attempted to increase motivation and buy-in by providing empirical evidence from meta-analyses of writing instruction as well as anecdotal evidence from our own classrooms. For each stage, the researcher team alternated modeling lessons by teaching the audience as if they were students using the same materials they would in the classroom. Then, in pairs, teachers practiced teaching the lesson to each other, reflected on their teaching, and refined their skills across the two days. While teachers engaged in practice, the researchers circulated, offered feedback and queried how the lesson could be adapted for specific students. This was repeated with each lesson in the sequence. At the end of Day 1, teachers were asked to create one prompt appropriate to their class and respond using the strategies they had been learning as well as incorporating local context including names, places, or other elements familiar to their students. We shared access to the entire group's essays to use as desired as local exemplars during their instruction. Researchers provided teachers with student writing samples from prior studies. Teachers learned how to assess writing to inform data-based decision-making, group students based on mastery of skills, and promote writing efficacy. Teachers also planned a detailed implementation calendar, learned to use the SWIVL device, and discussed how they would effectively implement the instruction. At the end of PBPD, all teachers opted to begin instruction the very next Monday, a testament to their excitement.

4.5. Teacher-led lessons

This study includes nine teacher-led lessons across the six stages of SRSD. Each teacher-led lesson had an outline of learning objectives, all student resources, and activities to implement for that session. It is important to note that each lesson may take between one to five class sessions depending on how well the students have mastered the knowledge and skills. Table 2 is an overview of each lesson included in the sequence. Below we will describe Lessons 0 and 8 in detail as these were additions to contextualize the intervention to this setting and writing task.

Instruction began with Lesson 0 in which students were given background knowledge about what it means to be a good writer, what good writing is, and the characteristics of the three academic writing genres. Together, the class analyzed writing prompts to identify the purpose. Students needed to understand that each genre has a unique purpose and genre parts that are expected to be included, which helps them to structure essays. To help them remember a basic writing process and the important parts of a genre, mnemonics were introduced. We used POW (Pick my ideas, Organize my notes, and Write and say more) to guide students in picking their ideas for writing. We taught them to "TAP the prompt" (identify Topic, Audience, and Purpose). Teachers learned to introduce the mnemonics and make them memorable with catchy tunes. It is important to develop automaticity with knowledge and skills to lighten the cognitive load throughout the writing process.

The second addition was Lesson 8. This lesson is focused on helping students transfer the persuasive writing strategies to the state standardized writing assessment, a personal expository prompt. Together, the class analyzed writing prompts similar to the state assessment and identified the writing genre (expository in all cases). The class analyzed example essays to identify how authors met the requirements of good writing. Teachers drew attention to the similarities between persuasive genre elements and those expected on the state assessment, noting that the difference was, on the state test they need not convince anyone. In pairs or small groups, students practiced making a writing plan based on the prompt, and finally, wrote a 26-line essay (length limitation of the state test). Finally, students completed the writing task independently under timed conditions, simulating the writing composition portion of

Table 2
Stages of SRSD and Associated Lesson Overviews.

SRSD Stage 1: Develop and Activate Background Knowledge. *Teachers ensure all students have necessary background knowledge to be successful by activating schema related to writing and writing within the instructed genre and providing essential pre-skills instruction (e.g., key vocabulary). The teacher may introduce a memory device or mnemonic to support student knowledge of the writing process, prompt analysis, and genre elements.*	
Lesson 0 Up to 3 sessions	Introduce characteristics of good writing, good writers, common academic writing genres (persuasive, expository, narrative), mnemonics POW (Pick an idea, Organize my notes, Write and say more) for the abbreviated writing process and TAP (Topic, Audience, Purpose) for analyzing the writing task, genre-related purpose words | Lesson wrap-up to review mnemonics and preparation for upcoming assessment
Lesson 1 Up to 3 sessions	Inquire about persuasive genre knowledge, resolve gaps in student schema | Explore exemplar persuasive texts | Introduce TREE mnemonic and instruct parts of persuasive genre (Topic sentence, Reasons, Explain, Ending) | Foster motivation and desire to learn and write in the genre
SRSD Stage 2: Discuss It. *The teacher introduces the meaning and value of good writing and how that extends to the instructed genre as well as what self-regulation is and how it impacts writing. The teacher then deeply explains each genre element, how the elements are related to the previously introduced memory device, and how they can be organized in text. The teacher discusses the value and expected benefits of using a strategy before asking the students for a commitment to use it themselves. Opportunities to generalize the strategy beyond the direct classroom application are discussed.*	
Lesson 2 Up to 3 sessions	Discuss why planning and making quick notes are helpful for students to remember key details and ensure all genre elements are included | Analyze prompts using TAP | Take planning notes from exemplar text finding the parts of TREE and demonstrating notetaking vs. sentence writing
Lesson 3 Up to 2 sessions	Take planning notes from a nonexample text finding the parts of TREE | Discuss that essays must have all the genre parts and make sense | Model revising the notes so the essay makes sense | Discuss using broad vocabulary and transition words
SRSD Stage 3: Model It. *The teacher models analyzing the prompt, planning, and writing from the plan - the entire task that will be asked of students from beginning to end. This includes using the genre-specific strategic approach, the self-regulatory procedures (goal setting, self-assessment, self-instructions, and self-reinforcement), consistent use of the strategy steps, all with an ongoing think aloud to elucidate the internal workings of the writer's mind for students to witness. During this time, the teacher demonstrates motivation and enthusiasm for writing, recognizing successes. Additionally, the teacher will differentiate the model to the classroom context, modeling the struggle of writing in the areas she expects her students may struggle, demonstrating how they might cope with and overcome those challenges. To facilitate engagement, the teacher may offer opportunities for students to collaborate with the process, but being mindful that the teacher controls the model. To close this stage, the teacher will model rereading the essay, evaluating it against the essential genre parts, and setting a writing goal.*	
Lesson 4 Up to 5 sessions Usually repeated	Explicitly model (70% teacher responsibility, 30% student) how to write an essay from the time the prompt is received through the revision process, including using the writing process with POW, analyzing the prompt with TAP, planning using the TREE mnemonic, using the plan to create the essay, reviewing the essay to be sure all essential genre components are included, and graphing the inclusion of genre elements | Discuss and model goal setting based on writing performance | Explicitly model self-statements for problem definition (*What is it I need to do?),* planning (*I need to make a plan.*), strategy use (*I can use TREE to plan!*), self-evaluation (*Did I include all the genre parts?*), coping and self-control (*This is a lot of work, but I have strategy!*), and self-reinforcement (*I really like that academic vocabulary word. Way to go!*) | Discuss and record/write self-statements used and their benefit | Teacher guides all of the process, most of the content, and does all of the writing, releasing responsibility to students to the degree they can be successful.
Lesson 5 Up to 5 sessions Usually repeated	Revisit writing goals set in previous lesson | Conduct a fully collaborative model (30% teacher responsibility, 70% students) with students leading most of the process, most of the content, while the teacher still does all of the writing | Explicitly model self-statements, graphing essay for inclusion of genre parts, discuss use of self-statements and their benefit | Evaluate if goals were met.
SRSD Stage 4: Memorize It. *The fourth stage of SRSD is embedded throughout instruction as teachers consistently reinforce the essential genre elements and meaning of each part of the mnemonic. Students may also be encouraged to memorize common self-regulatory phrases they can use (e.g., I have a strategy to be successful). This stage is critical, particularly for vulnerable learners, as fluency with these basic understandings and procedures reduce the cognitive load allowing students more freedom in applying the strategies independently.*	
Embedded throughout Reinforce writing process, characteristics of good writing and good writers, analyzing prompts, and genre components through the POW, TAP, and TREE mnemonics in every lesson following their initial introduction | Consistently review application of the persuasive genre.	
SRSD Stage 5: Support It. *In a gradual release of responsibility to the student, teachers scaffold student success during this stage as the students begin to work with the tools they have learned thus far – planning, organization, genre elements, self-instruction, self-regulation – with ready access to the teacher's support as needed. Teachers may offer guiding questions about next steps or may work with small groups of students who require additional support at this stage. The teacher deliberately fades coaching and involvement during this stage toward the goal of independent student performance.*	
Lesson 6 Up to 4 sessions Usually repeated	Provide support to small groups or pairs of students as they work with the learned strategy | Remind students to use self-statements and consider their established writing goals as they move through the writing process | Fade scaffolded support as students adopt more responsibility and move toward independent performance
SRSD Stage 6: Independent Performance. *In the final stage of SRSD instruction, students assume full responsibility for the instructed writing task from beginning – deciding what to write about – to the very end – completing the essay, evaluating it compared to expectations, and setting a goal for the next writing task to assure an ongoing trajectory toward improvement. Teachers and students address generalization, maintenance, evaluation, and goal setting.*	
Lesson 7 Up to 4 sessions	Monitor student performance as they independently work with the strategy, writing an essay from start to finish | Support students minimally, as needed | Discuss how students can generalize the strategy to other settings and tasks
Lesson 8 Up to 8 sessions	Transfer writing knowledge and skills to the state test genre (expository) and form (one-page, privileging depth over breadth) | Discuss that the state test is similar to persuasive writing as the prompts ask students to tell about their opinion or perspective, but there is no need to convince anyone.

the state standardized assessment. Teachers taught this lesson across four sessions, on average.

4.6. Implementation, fidelity Observations, performance Feedback, and support

Teachers were asked to teach SRSD for two to three 45-min sessions per week across 10 weeks. A member of the research team visited each intervention classroom during the first two weeks to observe for fidelity, provide support, help setup the SWIVL devices, and train the teachers individually on using the devices. Researchers also made in-person visits during weeks 4 and 6. All other observations were remote, if teachers were willing and able.

The research team employed multiple strategies to prepare teachers to use the video observation devices (i.e., SWIVL, iPad, Zoom). During PBPD, researchers introduced the devices, provided how-to guides, and facilitated practice. There were in-class demonstrations, videos, and

step-by-step guides with pictures. Researchers also trained multiple stakeholders in the school including administration, office staff, and students. Despite extensive efforts, two teachers continued to have issues connecting with SWIVL and these issues continued beyond Week 4.

After each observation, teachers were provided private performance feedback via email. The feedback followed a standard protocol: Thank you for allowing us to observe, observation data (date, time, lesson, observer), performance feedback (fidelity %), glows (positives from observation), grows (specific, targeted suggestions for improvement), clarification (any observed misunderstandings), logistics (e.g., next visit), appreciation, next scheduled observation, contact information. This means teachers who agreed to more observations also received more feedback.

Teachers were also offered a range of supports from which they could choose. We offered to co-teach or model teaching a lesson with their class. To those who wanted them, we provided ear buds to teachers so we could offer live coaching from afar.

4.7. Assessment procedures

School principals were the point of contact for pre- and post-testing. Testing took place at a date and time and in a location chosen by the principal of each school. At pretest, trained research assistants used a script to conduct the assessment session. The directions were read aloud before students were asked to complete each task. Pencils were provided and research assistants were available to answer questions throughout the process. Students were notified of how much time was remaining. Despite contact and planning with principals, at pretest, teachers at four schools (two intervention, two control) and at posttest, teachers at one school (control) reported not being aware of the testing. At pretest, one principal refused to allow testing to occur after being informed the school was in the control condition, but permitted post-testing.

4.8. Control condition

Teachers in the control condition implemented standard language arts instruction based on the state learning standards. To prepare for the state standardized assessment in writing, they used published workbooks focused on multiple choice questions about revising and editing. To prepare for the composition portion of the state test, these teachers shared published example essays from past tests and asked students to write to the prompts, but did not indicate they gave feedback on the essays. Three teachers in this condition were observed once across the semester. No evidence of SRSD writing instruction was documented.

4.9. Quantitative analysis

We used a *Multi-Site Cluster Randomized Trial (CRT)* design. Schools were randomly assigned to experimental conditions (intervention and control). Because students are nested in homerooms and homerooms are nested within schools, we estimated random intercepts models to account for this nesting data structure. We first ran an unconditional model to estimate variances of each of the outcome measures due to each of the random student, homeroom, and school units (M0). The main effects model (M1) addressed the second research question and examined the effect of SRSD on each of the outcome measures when predictors were included at different levels. At the student level, predictor variables included gender (1 = male, 0 = female) and the corresponding pretest scores (group mean centered). At the homeroom level, class average pretest scores (group mean centered) were included. At the school level, treatment efficacy was tested using dummy codes for experimental conditions (1 = SRSD, 0 = control). Other predictors included at this level were grand mean centered school average pretest scores, percentages of economically disadvantaged students, percentages of English learners, and percentages of students with disabilities, though the latter three predictors were not statistically significant and removed for parsimony. Missing data of the analysis sample were small (1 of 418 cases missed posttest on adherence and elements scores, 7 cases missed gender) and deleted listwise during analysis.

To address Research Question 2, we explored potential interaction effects between treatment and other variables to see whether treatment effect varied by students' gender, students' pretest scores, classes' pretest scores, and schools' pretest scores by adding corresponding product terms to the main effects model (M2). Statistically nonsignificant interactions were dropped for parsimony and statistically significant ones were plotted (based on the parsimonious model estimates) to examine the patterns of interaction. All models were estimated using the mixed procedure in Stata. The 0.05 significance level was used for all statistical tests.

Lastly, we estimated effect size of SRSD as compared to the control based on the main effects model for each outcome measure. Specifically, we computed the effect size as a standardized mean difference by dividing the adjusted (for pretest scores and gender differences) group mean difference by the unadjusted *pooled* within-group standard deviation of the pretest outcome measure (i.e., Hedge's g) because standard deviations were similar across and pretest scores were unaffected by experimental conditions.

5. Results

5.1. Teacher fidelity of implementation

Overall average fidelity of implementation per teacher was 93% (see Table 3) with the lowest *average* fidelity rating observed per teacher being 91% and the highest being 95%. We conducted 6–19 observations of each teacher ($M = 9$). The median and mode raw fidelity scores across all teachers was 100%. The highest *average* fidelity rating *per lesson* was 100% and the lowest was 82%. Lesson 3 was observed only once. The lowest *raw* fidelity score on any lesson was 50% (Lesson 5).

5.2. Student writing outcomes

Descriptive statistics of student outcome measures at pretest and at posttest by treatment condition are reported in Table 4. Pretest variables and student gender were included in the mixed models because they were expected to be related to the outcome measures and would improve statistical power for the test of SRSD effect.

Tables 5-7 present the mixed model estimates for adherence to the prompt, genre elements, and holistic quality scores respectively. Model 1 addressed the second research question about treatment efficacy of the SRSD intervention as compared to the business-as-usual control for each of the writing scores while controlling for students' gender and pretest writing scores. The M1 mixed model results demonstrated that SRSD had a statistically significant and large effect on prompt adherence ($b = 1.67$, $SE = 0.10$, $p < .001$; Hedges' $g = 1.87$), genre elements ($b = 3.08$, $SE = 0.92$, $p < .001$; Hedges' $g = 0.84$), and holistic quality scores ($b = 0.79$, $SE = 0.20$, $p < .001$; Hedges' $g = 0.87$), while holding gender and the corresponding pretest scores constant. That is, after adjusting for gender and pretest differences, students in schools that received SRSD instruction scored 1.87 standard deviation units higher in responding to the prompt, 0.84 standard deviation units higher in genre elements, and 0.87 standard deviation units higher in holistic quality on average at posttest than students in control schools.

The M2 results in Tables 5-7 addressed the moderator analyses and indicated that the effects of SRSD on all three writing measures were not significantly moderated by students' gender, students' pretest scores, or schools' pretest scores. None of these interaction terms were statistically significant at the 0.05 level for any of the writing posttest measures. In other words, the effects of SRSD in improving writing scores (prompt adherence, genre elements, and holistic quality) were consistently large for both male and female students regardless of their pretest performance levels at the individual or school levels. However, the interaction

Table 3
Fidelity score per teacher and per lesson.

			L0	L1	L2	L3	L4	L5	L6	L7	L8
# observations			12	7	6	1	7	13	18	2	4
Teacher	Avg. Fidelity	% observed	Average fidelity score for each lesson								
A	0.94	11	0.85	1.00			1.00	0.94			
B/C	0.91	35	0.80				0.88	0.83	1.00		
D	0.95	14	0.94		1.00		0.92	1.00		1.00	0.92
E/F	0.92	27	0.71		0.78			1.00	0.97		
G	0.93	22	0.95	0.88	0.88		0.96	1.00	0.93		
H	0.92	8	1.00	0.86		0.82	1.00				
Overall Avg.	0.93	0.20	0.90	0.91	0.87	0.82	0.97	0.95	0.96	1.00	0.92

Note: # observations = number of observations conducted; L0 = Lesson 0, L1 = Lesson 1, etc.; % observed = percent of total lessons observed for fidelity of implementation based on an estimated 26 sessions per class; Teachers B/C and E/F were teaching pairs.

Table 4
Descriptive statistics of outcome measures by time and intervention conditions.

	Intervention			Control		
	Pretest	Posttest	N	Pretest	Posttest	N
	Student Level					
Adherence	2.55 (0.86)	2.87 (0.55)	256/255	2.55 (0.95)	1.16 (0.61)	162
Elements	5.85 (3.75)	9.08 (3.75)	256/255	6.36 (3.53)	5.83 (3.21)	162
Holistic	1.81 (0.88)	2.79 (0.97)	256	1.91 (0.96)	2.05 (1.03)	162
	Classroom Level					
Adherence	2.56 (0.23)	2.88 (0.16)	17	2.60 (0.35)	1.19 (0.30)	16
Elements	5.93 (1.79)	9.20 (2.12)	17	6.95 (2.62)	5.94 (1.90)	16
Holistic	1.83 (0.27)	2.81 (0.46)	17	2.00 (0.44)	2.03 (0.55)	16
	School Level					
Adherence	2.56 (0.10)	2.88 (0.18)	6	2.58 (0.22)	1.22 (0.23)	5
Elements	5.96 (1.75)	9.02 (2.05)	6	6.82 (2.45)	6.11 (1.43)	5
Holistic	1.82 (0.19)	2.78 (0.41)	6	2.00 (0.39)	2.09 (0.35)	5

between treatment and homeroom pretest scores was statistically significant ($p < .05$) for prompt adherence and holistic posttest scores but not for genre elements posttest scores. Homeroom classes that had low pretest scores benefited slightly more from participating in the SRSD intervention than classes with higher pretest scores for both prompt adherence and holistic quality adjusted posttest scores. Classes that participated in SRSD consistently had higher adjusted genre elements posttest scores than their control counterparts regardless of class pretest scores.

6. Discussion

In this study, we sought to determine to what degree fourth grade teachers could implement SRSD for writing with fidelity following a two-day PBPD. Consistent with our hypotheses, results indicate that teachers were responsive to PBPD and implemented with high fidelity. However, the number of fidelity observations for some teachers was low and thus may not be representative of their instruction across time. We also wanted to determine to what degree PBPD would impact student writing outcomes and variables that may moderate those outcomes. As hypothesized, results indicate a significant increase in student writing performance, in terms of prompt adherence, genre elements, and holistic quality as a result of the intervention. Fig. 2 contains two examples of student writing progress during the intervention.

Table 5
Mixed model results on prompt adherence posttest scores.

	M0	M1	M2
Fixed Effects			
Intercept	2.13*** (0.25)	1.25*** (0.08)	1.27*** (0.08)
Intervention	–	1.67*** (0.10)	1.64*** (0.11)
Male	–	-0.10 (0.06)	-0.14 (0.09)
Student pretest scores	–	0.06 (0.03)	0.06 (0.05)
Class average pretest scores	–	0.35** (0.13)	0.57** (0.17)
School average pretest scores	–	-0.09 (0.35)	0.15 (0.46)
Intervention × Student gender		–	0.07 (0.11)
Intervention × Student pretest scores		–	-0.01 (0.06)
Intervention × Class pretest scores			-0.56* (0.26)
Intervention × School pretest scores			-0.53 (0.68)
Random Effect Variances			
Schools	0.70 (0.30)	0.02 (0.01)	0.01 (0.01)
Classrooms	0.00 (0.01)	0.00 (0.00)	0.00 (0.00)
Students	0.31 (0.02)	0.31 (0.02)	0.30 (0.02)
Model Fit Statistics			
−2LL	747.84	692.23	686.65
df	4	9	13
AIC	755.84	710.23	712.65
BIC	771.98	746.38	764.86

Note. *p < .05, **p < .01, ***p < .001, $\chi 2$ (df = 5, 0.95) = 11.07; $\chi 2$ (df = 4, 0.95) = 9.49.

6.1. Fidelity of implementation

Average fidelity of implementation was high which aligns with prior research as fidelity reported in most SRSD studies is approximately 90% (Harris et al., 2022). In the current study, Lesson 3 had the lowest average fidelity (82%), though this was based on a single observation and is not likely representative. In Lesson 3, teachers help students identify the parts of a silly essay that does not make sense, but has all genre parts.

Teachers modeled taking notes on the genre parts included in the example essay and then discussed how a good essay is more than having all the parts – the parts must also make sense. The most frequent error in this lesson involved taking notes. For example, not taking notes at all (e.g., discussing only, highlighting instead of writing notes), using full sentences (not notes), or not explicitly linking the notes to the essay. SRSD instruction includes explicit modeling of planning. Teachers demonstrate how using phrases for notetaking can help students organize their ideas and avoid fatigue. Another lesson component involves demonstrating how to change the notes to make sense, systematically returning to the plan, writing a sentence or more in response to a note, and then crossing off the note on the plan while working toward a complete essay. This process must be repeatedly modeled for students to

Table 6
Mixed model results on elements posttest scores.

	M0	M1	M2
Fixed Effects			
Intercept	7.67*** (0.64)	6.27*** (0.71)	6.30*** (0.75)
Intervention	–	3.08*** (0.92)	3.04** (0.98)
Male	–	-0.70* (0.31)	-0.78 (0.52)
Student pretest scores	–	0.26*** (0.05)	0.23* (0.09)
Class average pretest scores	–	0.53* (0.24)	0.56* (0.28)
School average pretest scores	–	0.27 (0.24)	0.27 (0.33)
Intervention × Student gender		–	0.11 (0.65)
Intervention × Student pretest scores		–	0.05 (0.11)
Intervention × Class pretest scores			-0.16 (0.55)
Intervention × School pretest scores			-0.01 (0.49)
Random Effects (variance)			
Schools	4.07 (1.91)	1.77 (0.95)	1.76 (0.95)
Classrooms	0.40 (0.41)	0.36 (0.39)	0.37 (0.39)
Students	10.37 (0.75)	9.44 (0.69)	9.43 (0.69)
Model Fit Statistics			
−2LL	2197.00	2114.75	2114.43
df	4	9	13
AIC	2205.00	2132.75	2140.43
BIC	2221.13	2168.90	2192.64

Note. *p < .05, **p < .01, ***p < .001, $\chi 2$ (df = 5, 0.95) = 11.07; $\chi 2$ (df = 4, 0.95) = 9.49.

Table 7
Mixed model results on holistic posttest scores.

	M0	M1	M2
Fixed Effects			
Intercept	2.47*** (0.15)	2.16*** (0.16)	2.17*** (0.17)
Intervention	–	0.79*** (0.20)	0.81*** (0.22)
Male	–	-0.30*** (0.09)	-0.27 (0.14)
Student pretest scores	–	0.33*** (0.05)	0.36*** (0.08)
Class average pretest scores	–	0.56 (0.34)	1.27** (0.49)
School average pretest scores	–	0.51 (0.35)	0.40 (0.40)
Intervention × Student gender		–	-0.05 (0.18)
Intervention × Student pretest scores		–	-0.05 (0.10)
Intervention × Class pretest scores			−1.30* (0.64)
Intervention × School pretest scores			0.44 (0.81)
Random Effects (variance)			
Schools	0.20 (0.10)	0.05 (0.05)	0.06 (0.05)
Classrooms	0.04 (0.04)	0.07 (0.05)	0.05 (0.04)
Students	0.86 (0.06)	0.73 (0.05)	0.73 (0.05)
Model Fit Statistics			
−2LL	1158.71	1069.72	1065.19
df	4	9	13
AIC	1166.71	1087.72	1091.46
BIC	1182.85	1123.88	1135.67

Note. *p < .05, **p < .01, ***p < .001, $\chi 2$ (df = 5, 0.95) = 11.07; $\chi 2$ (df = 4, 0.95) = 9.49.

internalize the value of planning and using the plan to write. Planning is associated with higher writing outcomes (Limpo & Alves, 2013). In past studies, we have found that students struggle with taking notes and using their notes to write an essay; thus, teacher modeling of taking notes for planning is key to unlocking this skill for students (McKeown et al., 2016; 2019). Future instances of PBPD may need to include increased emphasis on note-taking to improve fidelity of these skills.

Lesson 5 had the lowest raw fidelity score (50%), but it was a single instance. The teaching pair who earned this fidelity score did so while teaching a collaborative modeling session where they focused on analyzing the prompt and planning/organizing ideas for the essay. They engaged the students in small group discussions of ideas for each genre element, giving time for the small groups to talk and debate before presenting the best ideas to the whole group. For each element, the whole group debated amongst the ideas from the small groups. There was rich discussion about ideation and creative thinking, with each teacher circulating to monitor and stimulate further thought, ensuring each student had a chance to share. Due to the time spent on these activities, the steps for the closing review were not included. This adaptation reflects prioritization of higher-level thinking skills. Though it is preferred to end the lesson by summarizing what they learned, we do not believe learning was harmed on this day.

Previous studies have shown that teachers struggle most with the Model it stage of SRSD (Lessons 4 and 5 in this study; see FitzPatrick & McKeown, 2020; McKeown et al., 2019). Teachers have reported being uncomfortable when modeling (McKeown et al., 2014; McKeown at al., 2018). Yet, in the present study, the modeling lessons were among the highest average fidelity scores (97%, 95% respectively). They were also among the most frequently observed lessons in the present study. We prioritized observing during those complex lessons so we were able to provide individualized feedback and help to ensure students received the highest quality instruction possible. There is not enough variability to discern whether being observed or receiving feedback more frequently impacted implementation, but future researchers may consider these potential interactions.

6.2. Instruction observations

All teachers were asked to record or stream a writing lesson (using the video devices) at least once a week; two teachers achieved this goal. One of these teachers had a SWIVL device already and had experience using it, so the practice was not new. She made it part of her daily routine. The other teacher said she was eager to receive feedback on her teaching, so always recorded her lessons. Four teachers never used the video device for fidelity observations and were only observed in person. Nearly all teachers had delays in video observations due to difficulties with the SWIVL devices and other technical issues which impacted the number of observations that could be conducted. There were also issues with internet access required to stream or upload video files. For internet issues, principals filed technical support requests, but in one case, it was never resolved; that is to say, this teacher had no wireless access within her classroom for the duration of the study. Other technology issues included failing to set up or charge the equipment properly. Technology infrastructure in schools needs to be assessed to ensure immediate support. Since teachers received expert feedback each time they were observed, the teachers who made the video observations part of their daily writing routine (a key practice for successful implementation; Outhwaite et al., 2019) received more feedback than those who did not.

One teacher represented a range of difficulties with video observations. She was supported throughout the intervention, but was not included in the final analysis due to missing student tests. This teacher may have been resistant to being observed, both in person and via video. Twice, when we arrived at the scheduled observation time, she chose to teach grammar, which she told us was content she was comfortable delivering (thus, perhaps grammar was her cognitive default). She never turned on the video devices and reported a variety of issues including needing to teach the basics of grammar before implementing SRSD (not a requirement), having behavior issues, and having technology issues. To address the technology issues, we trained a student as well as the Assistant Principal to use the equipment, but the video tools were never used successfully. After five in-person visits, we had only one in-person observation of her providing SRSD instruction. Nonetheless, visits to her classroom revealed she was implementing SRSD as we saw evidence of

Sample 1 Pretest

"I wish to visit"

Sample 1 Progress

"All children should learn to use a computer. To do homework, like do image math and iStation. To look online for recipes they probably don't know how to cook. To play video games…this is maybe why children should learn to use the computer."

Sample 2 Pretest

"Do you have a place to go in the future. Well I do it is Disney Land. Disney Land is in California before you go to Disneyland you need to know the weather so you can know what kind of clothing to wear. I want to go to Disneyland because I can see all your favorite Disney characters and buy some toys. If you go to there they will have sleeping sweets and food. And more I hope you visit there.

Sample 2 Progress

"All children should be able to learn how to use a computer. First of all, kids need to learn how to use a computer for school projects. For example, when you are stuck on facts about your project, you can search them up on the internet. Another reason is that you can play games on the internet and search other things. For example, when you are outside then you go inside and you get bored you can play Fortnight or other games like Protegy and imagine math. These games are fun to play on the computer whenever you want. Finally, this is way all kids all over the world should learn how to use a computer for school projects, video games, and other things."

Fig. 2. Writing samples from participating students.

student writing samples in response to prompts we had provided, use of the mnemonics in their planning, student knowledge of the genre parts, and a vibrant, hand-drawn poster illustrating the strategy.

Leinhardt & Greeno (1986) found that when being observed by a principal or supervisor, teachers were more likely to make errors. They theorized that being observed increased the cognitive load and diminished the available working memory capacity required to manage the cognitive processing necessary to provide instruction. She was the most experienced teacher in our group, but experience alone is a poor predictor of expertise (Ericsson, 2008). What we observed of this teacher is consistent with other work that suggests stress has a negative effect on working memory and increases cognitive load (e.g., Klein & Boals, 2001). It is natural to default to a less effortful practice when cognitive resources are diminished (Feldon, 2007). This may explain, in part, what we encountered - resistance to being observed and/or recorded while teaching a new and complex intervention in writing, an area most teachers admit they are uncomfortable (Brindle et al., 2016). Collecting qualitative feedback from teachers in future studies will be useful in understanding their perspectives and designing observation techniques that balance teacher needs with the requirements of science.

6.3. Student outcomes

As hypothesized, the student writing outcomes were significant for the intervention group across all measures and each had a large effect size, which is consistent with prior SRSD research (Harris & Graham, 2016; 2018).

6.4. Prompt adherence

Prompt adherence had the strongest effect size and, to our knowledge, has not been reported in prior studies of SRSD, though it has been incorporated into other scoring procedures (e.g., McKeown et al., 2019). Released prompts from past state writing exams were used as outcome measures, but some students at a school in the control condition reported they had practiced writing to the posttest prompt. This may have inflated control student outcomes at posttest. Despite this potential contamination and inflation, effect sizes for the treatment group were both large and significant. This gives weight to the value of SRSD writing instruction over business-as-usual instruction even when it is used, quite literally, to teach the test.

The group level data for the control schools on the prompt adherence measure showed scores decreasing at posttest. These students still had similar on topic scores to students in the treatment condition, however they failed to have high prompt adherence scores indicating they were not responsive to *all* aspects of the prompt. This may indicate that the chosen writing curriculum in the control schools did not include explicit instruction in responding to all aspects of the writing prompts as required on the state level assessment.

6.5. Genre elements

The number of genre elements significantly increased for students in treatment over those in the control condition (ES = 0.84). Students receiving SRSD instruction included, on average, just over three additional elements compared to control students at posttest and compared to their own performance at pretest. Genre elements are the expected near transfer skill, so it was hypothesized that students would grow; this is consistent with prior work (2.83–3.02; Harris et al., 2012; McKeown et al., 2019). Yet, in this study, there was a single lesson focused on the tested genre which followed eight lessons on persuasive writing. This demonstrates that after learning the persuasive genre, it is possible for

students to learn a new, related genre quickly.

6.6. Holistic quality

The most common method to score writing is holistic quality as it reflects global growth valuing characteristics such as coherence, organization, ideation, flow, and clarity. Holistic growth is challenging to attain in interventions that are relatively short in duration. In this study, holistic scores significantly increased for students in treatment over those in control (ES = 0.87); this result is consistent with prior *meta*-analytic work that suggests SRSD produces positive holistic scores (Graham et al., 2012). Students who received SRSD instruction increased nearly an entire point in holistic scoring from pretest to posttest.

6.7. Moderating variables

Overall main effects had both significant and practical improvement across all three writing measures. There was no differential performance for prompt adherence, genre elements, and holistic quality scores based on percentages of English Learners, students with disabilities, economically disadvantaged students, nor gender. This indicates that all students across these categories benefited similarly from SRSD writing instruction. This contrasts with other SRSD studies which reported higher overall scores for females (e.g., Mason et al., 2017).

Additionally, treatment classes with low pretest scores garnered slightly more benefit from the intervention than those with higher pretest scores for prompt adherence and holistic scoring. While we did not make a prediction regarding an interaction between intervention and class, SRSD classes starting the study with the lowest prompt adherence and holistic quality scores made greater gains on these measures than SRSD classes starting the study with higher scores. This interaction was more pronounced for holistic quality than prompt adherence. A previous SRSD study by De La Paz and Graham (1997) with fourth and sixth grade students with learning disabilities also reported differential effects depending upon students' entry-level writing skills. Somewhat similar to the current study, SRSD students whose papers evidenced the lowest cohesiveness at the start of the study made greater gains on a maintenance writing probe than SRSD students who produced more cohesive written compositions when the experiment began. While these findings suggest that students with lower writing performance may benefit most from SRSD instruction, such a conclusion must be viewed as tentative at best. In the De La Paz and Graham (1997) study such interactions were not observed for other writing variables. Additionally, when SRSD and control students were asked to dictate compositions in this earlier study, SRSD students whose papers were longer and of higher quality at pretest made the greatest gains at posttest and maintenance, respectively. Even so, researchers need to pay greater attention to possible differential effects of SRSD with students who differ in entry-level writing skills. The writer(s)-within-community model (Graham, 2018) makes it clear that what and how students write as well as writing development are impacted by the class (i.e., writing community) where instruction takes place and individual differences among students in these classes.

6.8. Stressors

Boomtowns present a unique context, particularly as they impact services like schools. It was not rare within the district for teachers to leave the profession to pursue opportunities in the higher-paying booming industry. When that happened, students were redistributed among other teachers in the school since hiring new teachers or substitutes was difficult, a finding consistent with prior research (Jaquet, 2009). This increased the workload and stress on teachers who could receive up to ten new students on any given day. This uncertainty also created anxiety in the students as their educational experiences were disrupted (Schafft et al., 2014). Numerous teachers expressed their struggle to cope with the unpredictable nature of the job in this context, being able to make ends meet financially, and find affordable housing, consistent with findings identified by Dizon-Ross (2018). At the end of the school year, four of the participating teachers left the area citing their inability to afford living there.

Teacher shortages, instability, increased workloads, and stress are characteristics of a boomtown, but they can also be found in schools affected by the COVID-19 pandemic (see Jotkoff, 2022 for similarities). Thus, there are lessons from this study that may be generalized. In terms of teacher shortages, the competitive pay from industry was enticing; states and school systems must find compensation packages that are desirable to maintain qualified teachers. While researchers cannot increase teacher pay, there may be value to increasing teacher competencies to lighten the cognitive load of their daily work. By using PBPD where teachers learned and practiced implementation to achieve deep levels of comfort with the instruction, the evidence-based practice may slowly become the teachers' cognitive default. This may also free cognitive resources to devote to managing these others stressors present during the study and are even more pronounced in the post-pandemic context. Teachers also reported, anecdotally, that they were thankful for the time given during PBPD to plan the implementation pacing calendar with colleagues. They also were pleased they were able to practice the lessons in advance as it helped them to feel more confident and eager to implement the new writing instruction.

While PBPD does not address root causes of the stress – lack of competitive pay and overall instability – it does address the resources teachers have to be responsive to the stress that will likely be an ongoing part of their work. PBPD was effective at helping teachers in this stressful context implement evidence-based writing instruction with high fidelity which resulted in impressive student gains in a relatively short period of time. Retrospectively viewing the boomtown teachers' needs through the lens of a global pandemic, it is clear that we need high quality PBPD that can be provided on demand to meet the needs of a workforce impacted by instability and unpredictability.

That said, PBPD, though effective for the great majority of the teachers in this study, was inadequate to address the compounded needs of new teachers in this unique context. Two of the three teachers who withdrew from the study were new. One teacher had just graduated and was new to teaching. Another was new to the grade and school. Both were only days into their positions and were asked to participate in PBPD before their classrooms were even set up. The brand-new teacher withdrew from the study after four weeks because she felt overwhelmed with her job. The second teacher left her position without notice after three weeks as she was offered a higher salary in industry. A case might be made for wanting new teachers to begin with knowledge of effective writing instruction, and Graham (2019) has indicated special attention must to be paid to the needs of writing teachers who are new or new to the school system. Yet, both of these teachers were clearly overwhelmed and intensive professional development and implementation added to, rather than alleviated, their stress.

6.9. Limitations, Implications, and future directions

This study resulted in strong effect sizes on all student outcomes, but there are several limitations to be recognized. While researchers requested weekly reports on the amount of time spent teaching SRSD, only one teacher consistently reported this information. Teachers committed to teaching SRSD at least 45-min per session 2–3 times per week, but we do not know with certainty if that commitment was adhered to for all teachers despite consistent attempts to collect this information. In future studies, it would be useful to more accurately measure dosage by lesson to determine if there are correlations between instructional time and outcomes. In this study, we did not collect social validity measures from teachers or students, but this should be included in future studies.

There were many barriers to fully utilizing video conferencing for fidelity observations and coaching, both technical and human. Future research should more systematically identify these barriers so we can work toward solutions to improve the possibility of researchers partnering with more and more remote schools.

Principals and administrative staff were eager to discuss the teachers' fidelity of implementation with researchers, yet they did not have a clear understanding of the intervention or its components. The majority of the principals appeared at the PBPD session, but none stayed past initial introductions. Future research should address implementing professional development to help leadership improve support of evidence-based practices.

This study demonstrates that students can be taught to write persuasive essays using SRSD and then taught to transfer those skills to a personal informational essay. In the future, PBPD should be adapted to make it available online to increase accessibility in the post-pandemic world. A variety of issues can impede daily in-person classroom instruction such as parent employment, global illness, and academic challenges. A responsive intelligent tutoring system that assesses and responds to students' particular understanding would be helpful in supporting teachers to differentiate to meet the needs of all students. As school districts look for effective approaches to improve writing outcomes, including for standardized writing assessments, PBPD is an approach that has consistently resulted in high fidelity implementation of complex writing interventions resulting in positive impacts on student writing outcomes.

Funding

This work was funded by Institute of Education Sciences (CFDA No: 84.305A; Award No: R305A180212).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Alevriadou, A., & Giaouri, S. (2015). The impact of executive functions in the written language process: Some evidence from children with writing disabilities. *Journal of Psychologists and Counselors in Schools, 25*, 24–37. https://doi.org/10.1017/jgc.2015.3

Ball, D., & Cohen, D. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession* (pp. 3-32). Jossey-Bass.

Berninger, V., & Fuller, F. (1992). Gender differences in orthographic, verbal, and compositional fluency: Implications for assessing writing disabilities in primary grades. *Journal of School Psychology, 30*, 363–382. https://doi.org/10.1016/0022-4405(92)90004-O

Billett, S. (2010). *Learning through practice* (Vol. 5). Springer.

Brenner, D., & McQuirk, A. (2019). A snapshot of writing in elementary teacher preparation programs. *The New Educator, 15*(1), 18–29. https://doi.org/10.1080/1547688X.2018.142729

Brindle, M., Graham, S., Harris, K. R., & Hebert, M. (2016). Third and fourth grade teacher's classroom practices in writing: A national survey. *Reading and Writing, 29*, 929–954. https://doi.org/10.1007/s11145-015-9604-x

Chandler, B., & Sayeski, K. (2022, February). *Writing instruction within teacher preparation: An exploratory study*. [Poster Session]. Pacific Coast Research Conference, Coronado Bay, CA.

Clark, R. E. (2001). New directions: Cognitive and motivational research issues. In R. E. Clark (Ed.), *Learning from media: Arguments, analysis, and evidence.* (pp.263–298). Information Age Publishing.

Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review, 18* (1), 23–45. https://doi.org/10.1016/s0272-7358(97)00043-3

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

De La Paz, S., & Graham, S. (1997). The effects of dictation and advanced planning instruction on the composing of students with writing and learning problems. *Journal of Educational Psychology, 89*, 203–222.

Dizon-Ross, E., Loeb, S., Penner, E., & Rochmes, J. (2018). *Stress in boom times: Understanding teachers' economic anxiety in a high-cost urban district* (CEPA Working Paper No.18-09). Retrieved from Stanford Center for Education Policy Analysis: http://cepa.stanford.edu/wp18-09.

Ericsson, A. K. (2008). Deliberate practice and acquisition of expert performance: A general overview. *Academic Emergency Medicine, 15*, 988–994. https://doi.org/10.1111/j.1553-2712.2008.00227.x

Feldon, D. F. (2007). Cognitive load and classroom teaching: The double-edged sword of automaticity. *Educational Psychologist, 42*(3), 123–137. https://doi.org/10.1080/00461520701416173

Festas, I., Oliveira, A. L., Rebelo, J. A., Damião, M. H., Harris, K. R., & Graham, S. (2015). Professional development in self-regulated strategy development: Effects on the writing performance of eighth grade Portuguese students. *Contemporary Educational Psychology, 40*, 17–27. https://doi.org/10.1016/j.cedpsych.2014.05.004

FitzPatrick, E., & McKeown, D. (2020). Meeting the needs of middle school writers in a special education classroom: SRSD for the informational genre citing text-based evidence. *Education and Treatment of Children, 43*, 71–84. https://doi.org/10.1007/s43494-020-00006-2

FitzPatrick, E., & McKeown, D. (2021). Writing from multiple source texts: SRSD for fifth grade learners in inclusive settings. *Learning Disabilities Research and Practice*. https://doi.org/10.1111/ldrp.12257

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication, 32*(4), 365–387. https://doi.org/10.2307/356600

Graham, S. (2006). Writing. In P. Alexander & P. Winne (Eds.), *Handbook of educational psychology* (pp. 457–477). Erlbaum.

Graham, S. (2018). The writer(s)-within-community model of writing. *Educational Psychologist, 53,* 258-279.

Graham, S. (2019). Changing how writing Is taught. *Review of Research in Education, 43* (1), 277–303. https://doi.org/10.3102/0091732X18821125

Graham, S., Bollinger, A., Booth Olson, C., D'Aoust, C., MacArthur, C., McCutchen, D., & Olinghouse, N. (2012/2018). *Teaching elementary school students to be effective writers: A practice guide* (NCEE 2012- 4058). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/publications_reviews.aspx#pubsearch.

Graham, S, Harris, K. R., & Hebert, M. (2011). It is more than just the message: Presentation effects in scoring writing. *Focus on Exceptional Children, 44*(4). 10.17161/fec.v44i4.6687.

Graham, S., Harris, K., & Hebert, M. A. (2011a). *Informing writing: The benefits of formative assessment. A Carnegie Corporation Time to Act report.* Washington, DC: Alliance for Excellent Education.

Graham, S., McKeown, D., Kiuhara, S., & Harris, K. R. (2012). A meta-analysis of writing instruction for students in the elementary grades. *Journal of Educational Psychology, 104*(4), 879–896. https://doi.org/10.1037/a0029185.supp

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

Harris, K. R., & Graham, S. (2016). Self-regulated strategy development in writing: Policy implications of an evidence-based practice. *Policy Insights from Behavioral and Brain Sciences, 3,* 77-84. https://journals.sagepub.com/doi/abs/10.1177/2372732215624216?journalCode=bbsa.

Harris, K. R., & Graham, S. (2018). Self-regulated strategy development: Theoretical bases, critical instructional elements, and future research. In R. Fidalgo, K.R. Harris, & Braaksma, M. (Eds.), *Design principles for teaching effective writing: Theoretical and empirical grounded principles* (pp. 119–151). Leiden, The Netherlands: Brill.

Harris, K. R., Graham, S., Brindle, M., & Sandmel, K. (2009). Metacognition and children's writing. In Hacker, D., Dunlosky, J., Graesser, A. (Eds.), *Handbook of metacognition in education* (pp. 131-153). Mahwah, NJ: Lawrence Erlbaum.

Harris, K. R., Graham, S., & Mason, L. (2006). Improving the writing, knowledge, and motivation of struggling young writers: Effects of self-regulated strategy development with and without peer support. *American Educational Research Journal, 43*(2), 295–340.

Harris, K., Graham, S., Mason, L., & Friedlander, B. (2008). *Powerful writing strategies for all students*. Baltimore, MD: Brookes.

Harris, K. R., Lane, K. L., Graham, S., Driscoll, S. A., Sandmel, K., Brindle, M., & Schatschneider, C. (2012). Practice-based professional development for self-regulated strategies development in writing: A randomized controlled study. *Journal of Teacher Education, 63*, 103–119. https://doi.org/10.1177/0022487111429005

Harris, K., Camping, A , & McKeown, D., &. (2022). A review of research on professional development for multicomponent strategy-focused writing instruction: Knowledge gained and challenges remaining. In F. De Smedt, R. Bouwer, T. Limpo, & S. Graham (Eds). *Brill studies in writing: Methodological issues in high-quality intervention research in writing.* Manuscript under review.

Jaquet, J. B. (2009). *Energy boom towns and natural gas: Implications for Marcellus Shale local governments and rural communities.* NERCRD Rural Development Paper No. 43. The Northeast Regional Center for Rural Development. The Pennsylvania State University.

Jotkoff, E. (2022, February 1). NEA survey: Massive staff shortages in schools leading to educator burnout; alarming number of educators indicating they plan to leave profession. [Press release]. https://www.nea.org/about-nea/media-center/press-releases/nea-survey-massive-staff-shortages-schools-leading-educator.

Kiuhara, S., Harris, K., Graham, S., Brindle, M., McKeown, D., & Gilbert, J. (2013, February 7-10). *The effectiveness of practice-based professional development at tier 1: SRSD with an on-demand writing task* [Panel presentation]. Pacific Coast Research Conference, Coronado, CA, United States.

Klein, K., & Boals, A. (2001). The relationship of life events, stress, and working memory capacity. *Applied Cognitive Psychology, 15*, 565–579.

Leinhardt, G., & Greeno, J. G. (1986). The Cognitive Skill of Teaching. *Journal of Educational Psychology, 78*(2), 75–95. https://doi.org/10.1037/0022-0663.78.2.75.

Limpo, T., & Alves, R. A. (2013). Teaching planning or sentence-combining strategies: Effective SRSD interventions at different levels of written composition. *Contemporary Educational Psychology, 38*(4), 328–341. https://doi.org/10.1016/j.cedpsych.2013.07.004

Loewenberg Ball, D., & Forzani, F. M. (2009). The work of teaching and the challenge for teacher education. *Journal of Teacher Education, 60*(5), 497–511. https://doi.org/10.1177/0022487109348479

Mason, L. H., Cramer, A. M., Garwood, J. D., Varghese, C., Hamm, J., & Murray, A. (2017). Efficacy of self-regulated strategy development instruction for developing writers with and without disabilities in rural schools: A randomized controlled trial. *Rural Special Education Quarterly, 36*(4), 168–179. https://doi.org/10.1177/8756870517723617

McKeown, D., FitzPatrick, E., & Sandmel, K. (2014). SRSD in practice: Creating a professional development experience for teachers to meet the writing needs of students with EBD. *Behavioral Disorders, 40*(1), 15–25. https://doi.org/10.17988/0198-7429-40.1.15

McKeown, D., Brindle, M., Harris, K. R., Graham, S., & Collins, A. A. (2016). Illuminating growth and struggles using mixed methods: Practice-based professional development and coaching for differentiating SRSD instruction in writing. *Reading and Writing: An Interdisciplinary Journal, 29*, 1101–1140. https://doi.org/10.1007/s11145-016-9627-y

McKeown, D., Brindle, M., Harris, K. R., Sandmel, K., Steinbrecher, T., Graham, S., Lane, K., & Oakes, W. (2018). Teachers' voices: Understanding effective practice-based professional development for elementary teachers on SRSD in writing. *American Educational Research Journal, 56*, 753-791. doi.org/10.3102/0002831218804146.

McKeown, D., FitzPatrick, E., Brown, M., Brindle, M., Owens, J., & Hendrick, R. (2019). Urban teachers' implementation of SRSD for persuasive writing following practice-based professional development: Positive effects mediated by compromised fidelity. *Reading and Writing, 32*(6), 1483–1506. https://doi.org/10.1007/s11145-018-9864-3

McKeown, D., & FitzPatrick, E., Ennis, R., & Sanders, S. (Eds.) (2021). Self-regulated strategy development: A framework for effective instruction across the content areas. *Learning Disabilities Research and Practice, 36(3), 184-187.* Doi: 10.1111/ldrp.12256.

National Center for Education Statistics [NCES] (2012). The Nation's Report Card: Writing 2011 (NCES 2012–470). Institute of Education Sciences, U.S. Department of Education, Washington, D.C.

National Council of Teachers of English [NCTE]. (2014). How standardized tests shape – and limit – student learning: A policy research brief.

Outhwaite, L. A., Gulliford, A., & Pitchford, N. J. (2019). A new methodological approach for evaluating the impact of educational intervention implementation on learning outcomes. *International Journal of Research & Method in Education, 43*(3), 225–242. https://doi.org/10.1080/1743727X.2019.1657081

Santangelo, T., & Olinghouse, N. G. (2009). Effective writing instruction for students who experience writing difficulties. *Focus on Exceptional Children, 43*(4), 97–119.

Schafft, K. A., Glenna, L. L., Green, B., & Borlu, Y. (2014). Local impacts of unconventional gas development within Pennsylvania's Marcellus Shale region: Gauging boomtown development through the perspectives of educational administrators. *Society & Natural Resources, 27*(4), 389–404. https://doi.org/10.1080/08941920.2013.861561

Scherff, L., & Piazza, C. (2005). The more things change, the more they stay the same: A survey of high school students' writing experiences. *Research in the Teaching of English, 39*(3), 271-304. http://www.jstor.org/stable/40171667.

State of Texas Assessments of Academic Readiness: STAAR 2017.

Troia, G. A., & Olinghouse, N. G. (2013). The common core state standards and evidence-based educational practices: The case of writing. *School Psychology Review, 42*(3), 343–357. https://doi.org/10.1080/02796015.2013.12087478

Wijekumar, K., Beerwinkle, A. L., Harris, K. R., & Graham, S. (2019). Etiology of teacher knowledge and instructional skills for literacy at the upper elementary grades. *Annals of Dyslexia, 69*, 5–20. https://doi.org/10.1007/s11881-018-00170-6