TIER 1, TEACHER-IMPLEMENTED SELF-REGULATED STRATEGY DEVELOPMENT FOR STUDENTS WITH AND WITHOUT BEHAVIORAL CHALLENGES

A Randomized Controlled Trial

ABSTRACT

This study took place in the context of schools collaborating with a local university to implement an evidence-based, 3-tiered model of prevention and supports targeting academic, behavioral, and social goals. We examined whether Self-Regulated Strategy Development (SRSD) instruction, delivered by grade 2 and 3 general education teachers to all students (Tier 1), would improve story or opinion essay writing among students with and without behavioral challenges, and whether differential effects existed. SRSD instruction was effective for both groups of students in terms of genre elements and quality. Students without challenging behaviors made greater gains than those with challenging behaviors on some outcome measures. Teachers implemented SRSD with fidelity; SRSD was viewed as socially valid by teachers and students. No results were found for academic engaged time or overall level of behavioral problems. Finally, students' cognitive capabilities were a weak and inconsistent predictor of SRSD writing gains.

Karen R. Harris
ARIZONA STATE
UNIVERSITY

Kathleen Lynne Lane
UNIVERSITY OF KANSAS

Steven A. Driscoll
VANDERBILT UNIVERSITY

Steve Graham
ARIZONA STATE
UNIVERSITY

Kristen Wilson
VANDERBILT UNIVERSITY

Karin Sandmel
JOHNS HOPKINS
UNIVERSITY

Mary Brindle
VANDERBILT UNIVERSITY

Chris Schatschneider
FLORIDA STATE
UNIVERSITY

SCHOOL change and educational reform are critical issues receiving widespread attention across our country. Policy makers, educators, and researchers are invested in improving learning for all students and closing achievement gaps (Buffum, Mattos, & Weber, 2010; Gallimore, Ermeling, Saunders, & Goldenberg, 2009). Two constructs critical to achieving these goals, and that set an important context for much-needed research, are use of evidence-based practices in schools and school-based preventive interventions aimed at academic, behavioral, and social goals (Cook, Smith, & Tankersley, 2012; Lane, Kalberg, & Menzies, 2009). Thus, in this study, we worked in schools implementing evidence-based practices in a comprehensive, integrated, three-tiered model aimed at preventive interventions. Further, theory and research indicate that implementation of evidence-based practices is enhanced through school-university partnerships, such as those in place at the schools we worked with (Buffum et al., 2010; Cook et al., 2012; Gallimore et al., 2009).

Within this context, we focused on general education teacher implementation of an evidence-based intervention in writing. Research indicates that the writing performance of elementary grade students is a national concern, and the majority of elementary grade teachers report inadequate pre- and in-service preparation in writing instruction (Gilbert & Graham, 2010). The teachers and principals in the schools we worked in indicated a need for professional development in writing and were interested in improving their students' writing abilities. Next, we briefly describe what is meant by Tier 1 in a comprehensive, integrated, three-tiered model, and then turn to the theoretical and empirical bases of this study.

School-based preventive interventions are critical components of school change and may be one of the most important vehicles for helping all students achieve (Buffum et al., 2010). While different forms of evidence-based preventive interventions exist, an approach receiving increased research scrutiny and implementation across the country is the three-tiered model for preventing the development of learning and behavior problems and responding effectively to existing concerns (Buffum et al., 2010; Lane, Kalberg, & Menzies, 2009; Sugai & Horner, 2006). In comprehensive, integrated, three-tiered models, Tier 1 (primary level) emphasizes improved instruction and behavioral support in general education, using evidence-based practices to prevent academic and behavioral problems from occurring. Where challenges or problems do exist, interventions that can be implemented school- or class-wide are used. Academic, behavioral, and social progress are monitored frequently, and students who do not respond adequately to Tier 1 interventions receive Tier 2 (secondary level) interventions, typically in small groups. Similarly, students who are unresponsive to Tier 2 supports then receive highly intensive and individualized Tier 3 (tertiary level) interventions. Tier 2 and 3 supports are typically provided in addition to rigorous, effective Tier 1 practices, not in place of them (Buffum et al., 2010). Identification of a disability and provision of special education services typically occur after, or during, Tier 3 intervention.

Theoretical and Empirical Bases of This Study

Theory and research on both challenging behaviors and children's writing provide an important basis for this study and are discussed next. In addition, the theoretical

and empirical bases of the intervention studied here, Self-Regulated Strategy Development (SRSD) for writing, is described.

Challenging Behaviors and Academic Performance

The impact of students' challenging behaviors on their success in school as well as on the success of other students is an important concern (Walker, Ramsey, & Gresham, 2004). Students with behavioral challenges often experience considerable difficulty remaining engaged when working on academic assignments (Sutherland & Wright, in press). While formal diagnosis of emotional/behavioral disturbances typically does not occur until grade 5, it is critical to implement preventive interventions in the early elementary grades (Walker et al., 2004). Theoretically, disengagement or other challenging behaviors (e.g., creating disruptions or manifesting symptoms of anxiety) are likely to reduce the potential effectiveness of Tier 1 interventions (Walker et al., 2004). Students with challenging behaviors often have significant co-occurring academic difficulties in critical areas such as reading, writing, and mathematics (e.g., Nelson, Benner, Lane, & Smith, 2004). Schools and researchers, therefore, have increasingly focused their attention on the achievement gap between students with challenging behaviors and their peers. Writing, although an important predictor of academic success and a basic requirement for participation in civic life (Graham & Perin, 2007), has received little attention in research on Tier 1 interventions for these or other students (Harris, Graham, Brindle, & Sandmel, 2009).

Writing Performance and Challenges

Research indicates that the majority of elementary grade students, not just those with learning or behavioral challenges, demonstrate significant difficulties with narrative, expository, and opinion writing; many students also demonstrate a deteriorating attitude toward writing (Gilbert & Graham, 2010; Harris et al., 2009). This may reflect in part the complex nature of skilled writing. Writing requires extensive self-regulation of a flexible, goal-directed, problem-solving activity. Knowledge about writing, including genre knowledge, and strategies for planning and text production are also critical (Harris & Graham, 2009).

SRSD

SRSD was explicitly designed to address the complex nature of writing and the difficulties most students experience learning to write. Multiple theories and lines of research were, and continue to be, drawn on to develop an intervention responsive to the affective, cognitive, and behavioral demands writing makes on all children, including those with learning and behavioral challenges (Harris & Graham, 2009). Knowledge, strategies, self-regulation, and attitudes about writing are all explicitly targeted. In brief, in SRSD instruction students learn strategies for genre-specific and general writing and the knowledge needed to use these strategies. Next, students are introduced to self-regulation strategies (e.g., goal setting, self-instruction, self-monitoring, self-reinforcement) and supported as they learn to apply the academic strategies in tandem with the self-regulation strategies. Instruction is scaffolded to meet students' needs across six basic stages until students are able to use both writing and self-regulation strategies independently (for a detailed description, see Harris,

Graham, Mason, & Friedlander, 2008). Research indicates that SRSD for writing is effective with students who represent the full range of writing ability in a typical elementary class, and that SRSD has had the strongest impact of any strategies-instruction approach in writing (Harris & Graham, 2009; Harris et al., 2009). In addition, SRSD has been effective with students with learning and behavioral challenges; these children typically have even more significant difficulties with knowledge, strategies, self-regulation, and attitudes about writing than typically achieving children, making the components of SRSD a good match to their needs (Harris et al., 2009).

SRSD Research with Students with Challenging Behaviors

Several research teams have recently focused on investigating SRSD for teaching writing to students with challenging behaviors at Tier 2 or 3. In every study conducted to date, Tier 2 or 3 SRSD instruction improved the writing performance of students with challenging behaviors (Lane et al., 2008, 2010, 2011; Lienemann, Graham, Leader-Janssen, & Reid, 2006; Little et al., 2010; Mason & Shriner, 2008; Reid & Lienemann, 2006). However, research assistants rather than teachers delivered SRSD instruction in all of these studies. Although limited research indicates that upper-grade elementary teachers can successfully implement SRSD (De La Paz & Graham, 2002; Harris et al., 2009), no research was found that focused on the effects of teacher-implemented SRSD at Tier 1 for young students with challenging behaviors or compared effects of SRSD among students at Tier 1 with and without challenging behaviors. Research on differential effects is needed, as limited effects in Tier 1 may indicate the need for Tier 2 intervention with some groups of students.

Finally, while students' scores on measures of intelligence are positively related to their scores on measures of writing achievement (Graham & Harris, 2010), we were unable to locate any studies that examined whether students' cognitive capabilities predicted their ability to profit from SRSD or any other type of writing-strategy instruction. Sandmel et al. (2009), however, reported that they needed to modify SRSD instruction for a student with lower cognitive ability, as he had difficulty with the cognitive and metacognitive demands of strategies instruction, and they suggested that research is needed to investigate cognitive capability as a potential moderator.

The Present Study

In this randomized controlled study, we focused on investigating SRSD at Tier 1 in two text structure genres, story writing and opinion essays; thus, this study addressed both expository and narrative writing (Halliday & Hasan, 1989). We focused only on students receiving Tier 1 support; participating students were neither receiving special education services, nor had they been identified as having emotional and behavioral disorders, as is typical at these grade levels. This study was part of a larger study in a multiyear set of studies (Harris et al., 2012).

Opinion essays and stories were chosen as the genres for intervention because both were included in the local school curriculum and targeted on district and state writing tests, and teachers indicated that the majority or all of their students needed to improve in both genres. Second- and third-grade general education teachers were

randomly assigned to one of two conditions. In one condition, teachers taught the whole class strategies for writing an opinion essay using SRSD. In the other condition, teachers used SRSD to teach the whole class strategies for writing a story.

Each SRSD condition, therefore, served as the control for the other SRSD condition. This design has at least two advantages over the typical experiment, which compares an experimental treatment to no treatment or a business-as-usual control condition (Pressley et al., 1990). First, this design controls for Pygmalion and Hawthorne effects: students in both conditions are exposed to similar novel teaching conditions, teacher expectations should be similar in each condition, and teachers should be equally motivated to prove the worth of each condition. Second, students in each condition receive a meaningful treatment that was shown to be effective with struggling writers in earlier research (Graham, Harris, & Mason, 2005; Harris, Graham, & Mason, 2006). Further, we did not expect that SRSD opinion writing instruction would enhance story writing performance and vice versa, as previous research with young children has demonstrated that instruction in either of these two genres does not affect performance in the other (Graham et al., 2005; Harris et al., 2008). Specific research questions addressed are presented next.

Research Questions 1, 2, and 3

The first research question addressed in this study was, Does SRSD instruction in story writing and opinion writing at Tier 1 meaningfully improve the writing of students with behavioral challenges and their matched peers in terms of quality, length, and basic structural elements? We anticipated that SRSD instruction would be profitable for both groups on all writing measures based on the previous research reviewed here. We asked two further questions: Can general education teachers implement SRSD at Tier 1 with integrity? and, Will teachers and students find SRSD to have acceptable social validity? Based on the level of professional development and support provided in this study and on previous research, we expected the answers to both questions to be yes. These are critical questions; without affirmative answers, widespread acceptance of SRSD at Tier 1 is unlikely.

Research Question 4

Is SRSD instruction in these two genres differentially effective at Tier 1 for students with and without challenging behaviors? As noted earlier, it is unlikely that any intervention is equally effective with all types of students, especially when delivered to large groups of students by teachers in the general education classroom (Cook et al., 2012; Harris et al., 2009). We predicted that although students with challenging behavior would make gains across outcome measures, they would make smaller writing gains than students without such behaviors due to the effects of challenging behaviors reviewed earlier. We realized that it was possible, however, that differential effects would not be noted in this study, as SRSD instruction was situated in classrooms where teachers implemented PBIS (i.e., positive behaviors were enhanced, and inappropriate behaviors reduced, via effective behavioral principles).

Research Questions 5 and 6

As noted previously, students with challenging behaviors often have difficulty with academic engagement. Thus, we asked: Would SRSD instruction have differential effects on engagement (as measured by on-task behavior) among students with and without behavioral challenges? We predicted that while all students would show greater engagement, students with challenging behaviors would show greater gains in engagement than their counterparts without behavioral difficulties. Although SRSD instruction in this study did not include explicit instruction in the self-regulation of on-task behavior, we made this prediction based on three elements of SRSD instruction. First, students would be more engaged in these genre tasks because they were taught the strategies and knowledge needed to complete them successfully (Gaskill & Murphy, 2004). Second, setting genre-specific goals for improvement should provide students with an incentive to mobilize effort, and the anticipated satisfaction of achieving these goals should lead them to sustain their effort. Third, self-monitoring of improvement would make students aware that they could successfully carry out specific tasks, increasing the likelihood that they would engage in these tasks (Harris et al., 2009). We anticipated that the impact would be greater for students with behavioral challenges, as these students had more room for improvement.

Based on the same reasoning as presented for question 5, we asked, Would students with challenging behaviors show an overall decrease in problem behaviors (based on teachers' estimates of problem behaviors) after SRSD instruction? This is an important question, but we did not make a prediction, as insufficient research exists related to this question and this measure related to students' behavior across the school day, not just during writing instruction.

Research Question 7

The final question addressed was, Will students' cognitive ability moderate intervention outcomes? We predicted that students with higher cognitive ability would make greater writing gains than those with lower cognitive ability. We based this prediction on two points. One, planning is a cognitively demanding task (Kellogg, 1993), and students who are less cognitively capable may find learning how to plan more challenging than their more cognitively capable peers even given SRSD instruction. Two, SRSD instruction in this study was complex, involving learning not only strategies for planning, but genre-specific knowledge and a variety of self-regulation procedures. As noted earlier, this may overtax the capabilities of students with lower cognitive capabilities (Sandmel et al., 2009).

Method

Participants and Setting

Setting. The study took place in three inclusive elementary schools from one rural district in Tennessee. The district enrolled more than 31,000 students in 40 schools (8 high schools, 8 middle schools, 23 elementary schools, and one K–8 school). In this district, 8.5%, 1.39%, and 9.2% of students received free and reduced-price lunch, were English language learners, and were receiving special education services, respec-

tively. All students receiving special education were served in inclusive schools. Five of the district's elementary schools received Title 1 funding.

The three participating elementary schools each enrolled between 131 and 740 K–5 students who were participating in a larger study of SRSD (for a description and outcomes of the overall study, see Harris et al., 2012). All three schools had a primary prevention program that included academic (Response to Intervention [RTI]), behavioral (positive behavior supports), and social (Character Under Construction) components.

Prior to beginning the study, the necessary approvals were obtained from the university and the district. After consent was gathered from teachers, parental consent was solicited from all students enrolled in participating teachers' classrooms ($N = 443$) for participation in the overall study (see Harris et al. [2012] for a detailed description of consenting procedures). Of the 314 (70.88%) parental consent letters returned, 301 (95.86%) granted consent, and child assent was secured for students.

Teachers. Teacher participants included 20 second- and third-grade teachers from these three schools. Initially, 21 teachers granted consent and were randomly assigned to either the story ($n = 10$) or opinion ($n = 11$) writing intervention using a random-numbers table. However, one teacher assigned to story writing later withdrew, indicating that she did not have adequate time to devote to the intervention. Thus, teacher assignment was as follows: story writing ($n = 9$ [5 second grade, 4 third grade]) or opinion writing ($n = 11$ [5 second grade, 6 third grade]).

Teaching experience ranged from 1 to 21 years, with a mean of 10.57 years ($SD = 5.49$). Eight teachers had attained bachelor's degrees; eight teachers, master's degrees; and one teacher had attained her master's plus 30 units. The remaining teachers did not provide information regarding their educational attainment. All teachers were certified. Six teachers held teaching credentials in early childhood education, nine teachers in elementary education, and one as a reading specialist for K–12. The remaining teachers did not provide specific information regarding their credentialing status.

Students. Student participants were 56 second ($n = 35$) and third graders ($n = 21$), including 38 boys and 18 girls (see Table 1). The majority of students were Caucasian ($n = 54$, 96.43%), ranging in age from 6.93 to 9.61 ($M = 8.01$, $SD = 0.65$) years. None of these students received special education services as per our exclusionary criteria; we focused on students receiving only the primary prevention plan. The short form of the Wechsler Intelligence Scale for Children—Fourth Edition (WISC-IV; Wechsler, 2003) provided an estimate of cognitive ability. This estimate was obtained using two subtests: vocabulary and block design (split-half reliability = .82 for both). We used Sattler's (1991) formula to determine the estimate for conversion to a deviation quotient ($r = 0.91$). In this study, we focused on students with behavioral challenges (BC; $n = 28$) and a matched sample of students (matched on writing performance, description to follow) without such behaviors (typical; $n = 28$) who received instruction in either story or opinion writing groups.

Inclusion Procedures for Students with Behavior Challenges and Typical Students

The Student Risk Screening Scale (SRSS; Drummond, 1994; description follows) was administered as part of school-wide data practices to identify students with

Table 1. Characteristics of Nominated Student Participants

	Intervention Group					
	Story (*n* = 23)		Opinion (*n* = 33)		Total (*n* = 56)	
Variable Level	BC (*n* = 12)	Typical (*n* = 11)	BC (*n* = 16)	Typical (*n* = 17)	BC (*n* = 28)	Typical (*n* = 28)
Screening tools:						
Student Risk Screening Scale *N* (%):						
Low (0–3)	0 (.00)	11 (100.00)	0 (.00)	17 (100.00)	0 (.00)	28 (100.00)
Moderate (4–8)	9 (75.00)	0 (.00)	7 (43.75)	0 (.00)	16 (57.14)	0 (.00)
High (9–21)	3 (25.00)	0 (.00)	9 (56.25)	0 (.00)	12 (42.86)	0 (.00)
District writing assessment (DWA) *M* (*SD*)	3.25 (1.06)	3.55 (1.04)	3.50 (1.46)	3.65 (1.27)	3.39 (1.29)	3.61 (1.17)
Demographics:						
Gender *N* (%):						
Male	10 (83.33)	8 (72.73)	10 (62.50)	10 (58.82)	20 (71.43)	18 (64.29)
Female	2 (16.67)	3 (27.27)	6 (37.50)	7 (41.18)	8 (28.57)	10 (35.71)
Grade *N* (%):						
Second	9 (75.00)	10 (90.91)	8 (50.00)	8 (47.06)	17 (60.71)	18 (64.29)
Third	3 (25.00)	1 (9.09)	8 (50.00)	9 (52.94)	11 (39.29)	10 (35.71)
Ethnicity *N* (%):						
Caucasian	11 (91.67)	11 (100.00)	15 (93.75)	17 (100.00)	26 (92.86)	28 (100.00)
Hispanic	1 (8.33)	0 (.00)	1 (6.25)	0 (.00)	2 (7.14)	0 (.00)
Age *M* (*SD*)	7.86 (.43)	7.51 (.40)	8.29 (.73)	8.15 (.66)	8.11 (.65)	7.90 (.65)
Cognitive ability *M* (*SD*)	90.58 (9.98)	106.06 (10.57)	87.49 (12.11)	100.51 (10.91)	88.81 (11.15)	102.69 (10.93)

Note.—BC = behavior challenge; Student Risk Screening Scale (Drummond, 1994; range 0–21); DWA = most recent District Writing Assessment (DWA scores were not available for all students; story, *n* = 102; opinion, *n* = 132; total, *n* = 234).

behavioral challenges. Students identified as moderate or high risk on the SRSS were identified for the current study. A matched sample was created to identify students without behavioral challenges to serve as a comparison group. Matches were based on the most recent district writing assessment score, the writing strategy the students would receive during the intervention (story or opinion writing), grade level, and gender to the maximum extent possible.

We use the term *behavior challenges* as these students demonstrated initial or soft signs of antisocial behavior as measured by SRSS. However, most students did not exhibit pronounced concerns serious enough to warrant additional supports. Most students formally diagnosed with emotional disturbances requiring special education services are not typically identified for special education under this category until grade 5 (Walker et al., 2004).

Student Risk Screening Scale. The SRSS is a no-cost screening tool used to identify students who are at risk for antisocial behavior. Teachers rate each student in their homeroom class on seven items—(*a*) steal; (*b*) lie, cheat, sneak; (*c*) behavior problems; (*d*) peer rejection; (*e*) low academic achievement; (*f*) negative attitude; and (*g*) aggressive behavior—on a four-point Likert-type scale (never = 0, occasionally = 1, sometimes = 2, frequently = 3). Composite scores range from 0 to 21, with high scores suggesting higher levels of antisocial behavior. Total scores are used to place students into one of three risk categories: low (0–3), moderate (4–8), and high risk (9–21). The SRSS is significantly correlated with the Child Behavior Checklist's aggressive behavior subscale ($r = 0.79$; Walker et al., 2004). More recent validity studies of the SRSS indicate that when comparing students scoring in the low- versus high-risk categories, the SRSS had excellent accuracy for predicting both externalizing (95%) and internalizing (93%) problems on the Systematic Screening for Behavior Disorders (SSBD; Lane, Little et al., 2009).

District writing assessment. The district writing assessment was administered during the fall, winter, and spring to all K–5 students in the district. For the spring district writing assessment, students were given a situation and asked to write a story with their self as the main character. Students were given 35 minutes to respond to the writing prompt. Student writing was scored by a teacher other than the students' classroom teacher using a standardized rubric for each grade. Scores for second-grade students ranged from 1 (low) to 6 (high), while scores for third-grade students ranged from 0 to 6. If a student had not received a score for the spring assessment, we used the winter score.

Comparisons. A 2 × 2 ANOVA contrasting group (story vs. opinion) × student status (behavior challenges, BC vs. typical) on age indicated that the interaction term was not significant, $F(1, 52) = 0.44$, $p = .51$. However, there was a significant difference between groups in terms of age, $F(1, 52) = 10.25$, $p = .002$, with students in the opinion group being older (by approximately 3 months) than students in the story writing group, which was expected, as one of the second-grade teachers withdrew from the study. Results of a 2 × 2 ANOVA contrasting group × student status (those with and without behavior challenges) indicated that while the interaction term was not significant, there was a significant difference in the cognitive ability of students with and without behavioral challenges, $F(1, 52) = 22.66$, $p < .0001$. Students with typical behavior patterns had higher estimates of cognitive ability ($M = 102.69$, $SD = 10.93$) relative to students with behavioral challenges ($M = 88.81$, $SD = 11.15$); however, mean scores for all students fell within the average range (scores ranged from 68

to 117 on the cognitive-ability measure). Finally, a 2 × 2 ANOVA contrasting group × student status with respect to initial writing ability on the district test did not reveal any significant differences between subgroups, $F(3, 52) = 0.24$, $p < .87$.

Writing Instruction

As part of the school's primary prevention writing plan, teachers taught the writing process and basic writing skills. To determine how writing was taught in participating teachers' classrooms, these teachers completed a survey adapted from Cutler and Graham (2008). Teachers rated the frequency with which they offered writing activities and practices often recommended for use with young students on a seven-point Likert-type scale ranging from 0 (never) to 7 (several times a day). In addition, research assistants (RAs) conducted two 30-minute observations during teachers' literacy blocks. During these observations, we focused on teachers' use of (a) 31 of the activities included on the Cutler and Graham survey and (b) SRSD practices applied in this study to determine if any contamination occurred.

Survey data indicated that more than 75% of the teachers reported their writing program was best described as combining process writing and basic skills instruction. Two teachers described their program as process writing, and another teacher indicated she focused on traditional skills instruction. Teachers reported utilizing a process approach to writing on a weekly basis that included components such as conferencing, self-selection of writing topics, peer cooperation, planning and revision of papers, and sharing writing with others. Also, they taught spelling, grammar, and usage skills several times a week. Teachers stated that text structure, sentence construction, planning strategies, and revision strategies were taught less often, ranging from once a week to several times a month.

Observational data yielded similar information, with one exception. Teachers did not apply specific writing activities as frequently as expected; this was likely due to the limited number of observations conducted by project staff. For example, grammar and text organization instruction was applied by approximately one-half of the teachers. Spelling instruction was observed in only about one of every six classes. Teachers were observed teaching planning strategies in two out of every five classrooms.

Data from the direct observations of writing practices indicated that teachers did not use the experimental instructional procedures prior to beginning the study, with the following minor exceptions: (a) four teachers were observed teaching students about the parts of stories (two teachers in each condition), (b) one teacher encouraged students to include such parts in their stories, and (c) one teacher taught the parts of an opinion essay. Also, there was no statistical difference between teachers in the two conditions on any of the survey items (all 41 p's $< .11$) or the writing activities observed (all 31 p's $< .07$), suggesting that the writing programs of teachers in the story writing and opinion writing groups were comparable.

Intervention: Self-Regulated Strategy Development for Writing

As explained previously, teachers were randomly assigned to either the story or opinion essay writing conditions. The lesson plans used in this study are available online at http://kc.vanderbilt.edu/projectwrite. In this study, all students learned the

general planning strategy POW (pick my idea, organize my notes, write and say more). In the organize-my-notes step, students made and organized notes in response to the genre-specific opinion essay or story writing strategy they were learning (described next), as well as notes related to general writing strategies such as beginning with a catchy opening, good word choice, and wrapping it up right.

Opinion and story writing strategies. The genre-specific writing strategy for writing opinion essays was the early elementary version of TREE (Harris et al., 2008; *T*, topic sentence—tell what you believe!; *R*, reasons, three or more—Why do I believe this? Will my readers believe this?; *E*, ending—wrap it up right!; *E*, examine—do I have all my parts?). The genre-specific writing strategy for writing stories was WWW, What = 2, How = 2 (*W*, who is the main character or characters?; *W*, When does the story happen?; *W*, Where does the story take place?; *What*, What does the main character do or want to do? What do the other characters do?; *What*, What happens then? What happens with the other characters?; *How*, How does the story end?; *How*, How does the main character feel? How do the other characters feel?). The stages and components of instruction are briefly described next using the opinion essay writing strategy; the same stages and components were implemented in the story writing instruction.

Develop background knowledge. During this first stage of instruction, students acquired knowledge, skills, and vocabulary needed to apply POW and TREE for opinion essay writing. First, POW and its corresponding steps were introduced, and the teacher and students discussed why each step was important. Second, the characteristics of a good opinion essay were discussed. Third, the instructor introduced the appropriate mnemonic *TREE* as a "trick" for remembering the parts to be included in an opinion essay. Vocabulary related to opinion essays (e.g., fact vs. opinion, transition words) was carefully discussed. After discussing examples for each part, students listened as a model opinion essay was read and identified each element. This continued with additional opinion essays until students could accurately identify all of the respective parts. Consistent with the recursive nature of SRSD instruction, students spent a few minutes during each succeeding lesson across the next four stages of instruction reviewing the steps of POW, the appropriate genre-specific mnemonic, and what each stood for.

Discuss it. Students continued to examine opinion essays using POW and TREE, as a class or independently, highlighting the characteristics of an opinion essay (e.g., topic sentence, three or more reasons, transition words, ending) and taking notes on the graphic organizer. Students discussed their current writing and self-regulation abilities, and how the POW and TREE strategies are beneficial for helping them improve their writing performance. Overall goals for writing (e.g., fun to read, fun to write, makes sense, has all its parts, persuades the reader) were introduced to students. For the opinion essay, the goal to include at least five parts (e.g., topic sentence, three reasons, and ending) was also discussed. Students were asked to make a commitment to learning the strategies and working in partnership with the teacher and other students.

Model it. Teachers modeled the writing process using the POW and TREE strategies for a writing prompt (e.g., Should children your age have to do chores at home?) aloud. For example, a teacher modeled *P* (pick my idea; yes, children should do chores). Next, teachers modeled *O* (organize my notes) by using the TREE strategy and writing notes on the graphic organizer; students could help them with ideas for their notes. Finally, teachers modeled *W* (write and say more), by using the notes on the graphic organizer to write the opinion essay; again students assisted. While the

teachers modeled the writing process and the use of the writing and strategies, they also modeled using self-instructions for different purposes including goal setting (e.g., I need to include five parts), problem solving (e.g., What do I have to do now?), self-evaluation (e.g., Do I have all my parts?), self-reinforcement (e.g., This is good. My reader will agree with me), and coping (e.g., I can do this if I try. This isn't so bad).

Teachers then discussed with students the different self-instructions they used. Teachers encouraged students to generate and record their own self-statements to use during the writing process. Students were then introduced to the self-monitoring component. Teachers provided students with a rocket graphing sheet. Each rocket is divided into five parts (one for each of the major elements of TREE) and surrounded by stars (see Harris et al., 2008). Teachers reminded students of the goals for writing opinion essays and the number of parts to include (five). Together, teachers and students examined the opinion essay written and identified the parts. For each part present, the students colored in a box on their rocket. Next, teachers and students identified the transition words used in the opinion essay and colored a star for each transition word included in the essay. The goals for writing were restated, and teachers encouraged students to include all five parts and transition words when they write a opinion essay. Teachers explained that students could "bust" the rocket if they included more than three good reasons, and told them to write their total number of parts on top of the rocket.

Memorize it. The purpose of memorizing the mnemonics (POW and TREE) and what they stand for is to promote automaticity when the students are ready to begin writing on their own. Memorization typically begins in the earlier stages by practicing the mnemonics each day. Teachers may (and in this study often did) develop games (e.g., using flash cards), songs, or hand motions to help students memorize the mnemonics. At this point, teachers ensured that all students had memorized the mnemonics before proceeding to the next stage.

Support it. Teachers and students worked collaboratively to write opinion essays using the POW, TREE, and self-regulation strategies. Teachers gradually faded support as students independently used the writing and self-regulation strategies and met criteria (e.g., including five parts, using transition words, using good reasons). Once students could independently write essays using the graphic organizer, teachers modeled making their own graphic organizer on scratch paper and using this graphic organizer to make notes. Students discussed why it was important to generate their own graphic organizer for writing opinion essays (they won't have the graphic organizer when they write in the future and on tests). Teachers and students identified and discussed other times and reasons the strategies could be used.

Independent performance. Students reached independent performance when they were able to independently write their opinion essays to criteria (e.g., include all five parts, use transition words, etc.) using the self-regulation strategies. Teachers and students continued to discuss other times when the writing strategies can be used and developed plans to maintain use of the writing strategies.

Professional Development

All consented teachers participated in small-group, practice-based professional development sessions at each school site to learn how to implement SRSD instruction in their classrooms (Ball & Forzani, 2009; see Harris et al., 2012, for details on

professional development). Separate sessions occurred for each intervention condition if more than one condition was implemented at that site. Professional development sessions began with a discussion of the students in each teacher's class and their writing strengths and needs. Each group met for 6 hours on 2 days for a total of 12 hours of professional development. One teacher was not available for the initial training and instead completed an individual training lasting 9 hours. Differentiation of SRSD instruction for students with differing levels of writing ability in each teacher's class (Sandmel et al., 2009) was discussed, as was the availability of professional development leaders to answer questions and assist teachers as needed. Fidelity observations to be conducted in each teacher's classroom were also shared and discussed.

After professional development and preintervention measures were completed, teachers began instruction. Teachers were given a pacing calendar to guide their instruction, although there was flexibility in how students moved through instruction in accordance with SRSD methods. Teachers conducted a maximum of 24 total sessions, with approximately three sessions each week. During the allotted instructional time, no other writing intervention was delivered.

Treatment-Integrity Procedures

Treatment-integrity data were collected for 25% of instructional sessions for each teacher (see Table 2). RAs conducted observations of their assigned teachers using a checklist for each lesson containing key instructional components. When RAs arrived, teachers told the RA which components of what lesson they were going to be teaching that day, and RAs completed the checklist for that lesson by marking off each item they observed. Teachers completed the same lesson-specific checklist as the observer; in addition, teachers completed lesson-specific checklists for all lessons, regardless of whether an observer was present. In the Results section, we report the rate of integrity from (*a*) teacher self-report for each lesson in its entirety, (*b*) RA observations for the 25% of sessions observed, and (*c*) teacher self-reports for the subset of the components observed by the RAs (as lessons often took more than one class session to complete). These data provided a stringent test of whether or not each SRSD treatment was delivered as intended.

Outcome Measures

Before beginning the intervention, teachers administered two writing prompts (story and opinion) on consecutive days to all students in their classrooms; the writing prompt corresponding to their instructional condition was administered first (see Table 3). For example, writing prompts were counterbalanced such that if the teacher was randomly assigned to teach story writing, then that teacher administered the story writing prompt first and then the opinion writing prompt on the next day. It is unlikely that this arrangement influenced students' scores, as there is little relationship between students' performance in the two genres (Graham, Harris, & Hebert, 2010), and it allowed each student to complete their first composition in the instructed genre.

For students in the current study, RAs also collected a measure of academic engagement (Academic Engaged Time) while students completed each of the two writing

Table 2. Social Validity and Treatment Integrity

Variable Level	Intervention Group				Effect Sizes
	Story Instruction M (SD) (n = 9)		Opinion Instruction M (SD) (n = 11)		
Teacher (IRP-15):					
Pre	75.00 (9.51)		73.18 (9.44)		
Post	78.00 (12.00)		78.18 (9.69)		
Δ	3.00 (11.94)		5.00 (11.26)		.17
					.28 [a]
					.52 [b]
	BC (n = 12)	Typical (n = 11)	BC (n = 16)	Typical (n = 17)	
Student (CIRP):					
Pre	35.33 (4.94)	35.55 (3.45)	35.31 (3.14)	32.53 (6.27)	
Post	35.25 (5.15)	35.55 (8.59)	36.88 (5.52)	33.06 (8.93)	
Δ	−.08 (7.57)	.00 (8.68)	1.56 (6.26)	.53 (9.31)	
Treatment integrity:					
Teacher (all sessions)	94.20 (5.01)		97.06 (3.87)		.64
Teacher (observation)	90.04 (11.49)		99.23 (1.72)		1.39
RA (observation)	91.54 (7.27)		88.63 (7.99)		−.38
					−.16 [c]
					2.18 [d]

Note.—IRP-15 = Intervention Rating Profile-15; scores for IRP-15 range 0–90. CIRP = Children's Intervention Rating Profile; scores for CIRP range 0–42. RA = research assistant; AET = academic engaged time.

[a] Compares story pre and story post.

[b] Compares opinion pre and opinion post.

[c] Compares story teacher (observation) and story RA (observation).

[d] Compares opinion teacher (observation) and opinion RA (observation).

prompts (see Table 3). Each RA collected engagement data on one student participant; thus more than one RA was present in each classroom when writing prompts were administered. The same procedures were followed for postassessments.

Story writing prompts. Students were given a line drawing and asked to write a story about the scene presented in the picture. Prior studies have demonstrated that the story writing prompts used in this study are equivalent with respect to how much and how well primary grade students write (Graham et al., 2005).

Students were instructed to write for 20 minutes; however, students were given extra time to finish their writing if necessary. Students were encouraged to plan before they wrote and told they could only ask for assistance in spelling a word. If students asked for other types of assistance, they were told "do the best you can." Each story was scored for number and quality of story elements, overall writing quality, the number of words, and transition words used. Each paper was typed and saved as a Microsoft Word document. Identifying information was removed and spelling, capitalization, and punctuation errors were corrected. Electronic versions were scored rather than the student-written responses, given that appearance of text- or surface-level features (e.g., handwriting legibility, spelling errors) can influence judgments about writing quality and content (Graham et al., 2010).

Each story was scored to determine the number and quality of basic story elements included: main character(s), time, locale, what the main character(s) does or wants

Table 3. Student Participants' Writing Outcomes at the Student Level

| | Intervention Group | | | |
| | Story Instruction ($n = 23$) | | Opinion Instruction ($n = 33$) | |
Writing Style	BC ($n = 12$) M (SD)	Typical ($n = 11$) M (SD)	BC ($n = 16$) M (SD)	Typical ($n = 17$) M (SD)
Story:				
Elements:				
Pre	3.21 (1.05)	3.09 (2.32)	3.13 (2.39)	3.41 (2.24)
Post	4.25 (1.63)	6.59 (2.81)	3.06 (2.46)	.74 (2.68)
Δ	1.04 (2.34)	3.50 (3.31)	−.06 (2.39)	.32 (2.29)
Quality:				
Pre	2.29 (1.08)	2.59 (2.02)	2.38 (2.27)	2.74 (1.97)
Post	2.50 (.85)	4.59 (1.76)	2.28 (1.92)	3.12 (2.08)
Δ	.21 (1.16)	2.00 (2.19)	−.09 (1.63)	.38 (1.74)
Word count:				
Pre	64.42 (21.36)	94.09 (44.15)	91.25 (76.88)	95.12 (60.29)
Post	57.17 (26.76)	113.36 (78.57)	78.81 (66.56)	93.47 (70.98)
Δ	−7.25 (29.21)	19.27 (62.22)	−12.44 (60.22)	−1.65 (28.41)
Transition words:				
Pre	.96 (1.66)	1.68 (2.06)	.81 (1.41)	1.09 (2.18)
Post	.50 (.93)	1.91 (1.66)	1.34 (2.15)	.76 (1.24)
Δ	−.46 (1.89)	.22 (2.53)	.53 (1.92)	−.32 (2.59)
AET:				
Pre	67.17 (19.56)	72.01 (17.73)	62.97 (27.78)	72.11 (19.58)
Post	77.53 (18.34)	76.77 (22.44)	57.64 (26.85)	73.82 (19.41)
Δ	10.36 (24.77)	4.76 (20.45)	−5.33 (34.77)	1.71 (16.85)
Opinion:				
Elements:				
Pre	3.08 (1.92)	6.23 (2.96)	3.13 (4.10)	4.91 (3.12)
Post	2.33 (2.91)	5.27 (4.13)	6.97 (3.50)	7.18 (3.00)
Δ	−.75 (3.43)	−.95 (4.71)	3.84 (3.90)	2.26 (4.28)
Quality:				
Pre	.92 (.47)	1.23 (.34)	.63 (.76)	1.12 (.67)
Post	.75 (.54)	1.64 (1.31)	2.66 (1.54)	3.71 (1.09)
Δ	−.17 (.69)	.41 (1.36)	2.03 (1.28)	2.59 (0.97)
Word count:				
Pre	45.67 (28.82)	92.00 (62.54)	82.31 (68.06)	94.82 (42.78)
Post	38.75 (18.79)	79.82 (64.43)	68.56 (38.58)	73.65 (50.13)
Δ	−6.92 (16.46)	−12.18 (52.78)	−13.75 (48.60)	−21.18 (42.55)
Transition words:				
Pre	.00 (.00)	.36 (.67)	.63 (1.76)	.24 (.47)
Post	.21 (.45)	.45 (1.51)	1.53 (1.73)	3.38 (1.10)
Δ	.21 (.45)	.09 (1.45)	.91 (2.30)	3.15 (1.11)
AET:				
Pre	54.19 (22.31)	61.10 (21.97)	58.03 (31.93)	75.95 (15.68)
Post	57.26 (30.26)	58.48 (34.44)	53.84 (19.23)	67.90 (21.07)
Δ	3.07 (21.52)	−2.62 (35.44)	−4.19 (26.27)	−8.05 (22.52)

Note.—SRSD = Self-Regulated Strategy Development. Effect sizes are report in text. AET = academic engaged time.

to do (goals), what happens next, how the story ends, and how the main character feels. For each element, a score of 0 was assigned if the element was not present, and a score of 1 was awarded if the element was included. Elements that were highly developed received a score of 2 (see Graham, Harris, & Sawyer, 1989, for a fuller

description of this measure). This was the most proximal measure of student growth in the study. All writing prompts were scored by two RAs who had been trained in the scoring method with an interrater reliability (IRR) of .92 during training. For analysis, the mean of the two RA scores was used. IRR between the two RAs' scores was .85.

Number of words was scored by using the word-count feature in Microsoft Word tools, whereas overall quality of stories was scored using a holistic measure. Scorers were told that ideation, organization, sentence structure, word choice, and grammar should all be taken into account in forming a judgment about writing quality, but no single factor should receive undue weight. Papers were scored by RAs using a 0–8-point scale, with higher scores representing higher-quality writing. RAs were provided with a representative (anchor) paper for scores of 1, 3, 5, and 7 that was used to assist them in scoring stories. These anchor points were developed in prior investigations (see Graham et al., 2005; Harris et al., 2006). The scorers received training in how to accurately and reliably use the holistic scoring system (IRR during training was .95). Each rater independently scored all stories produced by the students for quality (IRR was .90).

Opinion essay writing prompts. Students were presented with a question asking their opinion on school or home issues (e.g., Do you think school rules are necessary?). Teachers asked students to respond by writing an essay. The writing prompts were taken from earlier SRSD studies with elementary-age students that have established the equivalency of these prompts with respect to how much and how well the students write (Graham et al., 1989, 2005; Harris et al., 2006).

Teachers administered the opinion prompts using the same procedures described for administering story writing prompts (e.g., time allotted, planning, and requesting assistance). Scoring procedures for each opinion essay were also consistent with those described for story writing prompt scoring (e.g., typing prompts; removing identifying information and correcting spelling, capitalization, and punctuation errors; scoring; and data entry, including reliability). In terms of scoring basic elements of opinion writing, RAs scored the following: topic sentence (statement indicating what the author believed), reasons (explanation as to why an author believed a particular premise or explanation of why they refuted a counterpremise), conclusion (a closing statement or a statement that brings the author's ideas together), and elaboration (additional information or examples for a premise, reason, or conclusion). For topic sentence and conclusion, a score of 1 was awarded if the element was present and a score of 0 if it was absent. For reasons and elaborations, 1 point was awarded for each separate and unique reason or elaboration included in the paper. Students received additional points based on the quality of the reasons they gave to support their topic sentence (see De La Paz & Graham, 2002, for a fuller description of this measure). Again, number and quality of elements was the most proximal measure of student performance. As with story writing prompts, all opinion writing prompts were scored by two RAs who had been trained in the scoring method with an IRR of .97 during training. For analysis, the mean of the two RA scores was used. IRR was .76.

Number of words and overall quality of opinion papers were scored using the same procedures described for story writing papers. The scorers received extensive training in applying the holistic scoring system (IRR during training was 82%). Each rater independently scored all stories produced by the students. IRR was .82. All papers (opinion text and stories) were scored for number of transition words (e.g.,

first, second, therefore, next, and lastly). All papers were scored by two trained RAs (IRR during training was .98). IRR for all papers was .95.

Academic engaged time. Students' engagement during the teacher-administered writing prompts was assessed using a modified version of the direct-observation procedures delineated in the Systematic Screening for Behavior Disorders (SSBD; Walker & Severson, 1992). Academic engaged time refers to the amount of time a student spends actively engaged attending to and working on relevant academic material, in this case, writing prompts. Students were considered academically engaged if they were (*a*) writing in response to the prompt, (*b*) thinking about their response (5 seconds maximum with eyes gazing away from paper), or (*c*) appropriately asking the teacher a question regarding the prompt. Students were not considered academically engaged if they were (*a*) not writing, (*b*) daydreaming (eyes gazing away from paper for more than 5 seconds), (*c*) out of their seat, or (*d*) completing a follow-up activity (e.g., drawing a picture).

Before conducting direct observations, RAs met a strict criterion for reliability. First, each observer completed the written quiz (Walker & Severson, 1992) at 90% proficiency or better. Second, each observer completed two practice data-collection forms similar to those used during observations with 100% accuracy. RAs then completed observations of student vignettes on video until they met 90% interobserver agreement (IOA) with the project director on three consecutive observations. Finally, RAs completed live student observations at a local school until they met 90% IOA with the project director on three consecutive observations.

Each RA observed one student during pre- and postintervention while he or she completed the teacher-administered writing prompts for story and opinion writing. RAs used a stopwatch to measure the duration of engagement during each 15-minute observation. RAs recorded the time the teacher began the prompt, and then allowed the stopwatch to run while each student wrote. If the student stopped writing, the stopwatch was stopped and then restarted when the student again became engaged in the writing process. At the end of the 15-minute observation, the RA recorded the end time. Total duration was computed by dividing the total number of seconds engaged by 900 and multiplying the quantity by 100 to obtain a percentage. (IOA was obtained for 25% of observations.) When collecting data, observers positioned themselves unobtrusively in the classroom and were instructed to ignore all initiations from the students and to minimize interactions with teachers during data collection.

Each student was observed twice at each time point, once during story writing and once during opinion writing. IOA was collected at both time points for both writing prompts. Reliability was conducted for 25% each of observations during preintervention story writing and opinion writing by having two RAs observe the same student. IOA was computed by dividing the smaller engaged time by the larger engaged time value and multiplying the quantity by 100 to obtain a percentage. Mean IOA for story writing was 96.37% ($SD = 2.94$) and for opinion writing was 95.84% ($SD = 5.05$). Reliability at postassessment was conducted for 25% each of story and opinion observations. Story writing observations had an IOA of 95.91% ($SD = 4.49$), while opinion writing observations had an IOA of 96.76% ($SD = 5.85$).

Social Skills Rating System—Teacher Version (SSRS-T; Gresham & Elliott, 1990). The Social Skills Rating System contains three subscales: social skills, problem behavior, and academic competence. Each subscale has a mean of 100 ($SD = 15$). For this study, teachers completed the problem behavior subscale of the SSRS prior to

Table 4. Student Participants' Behavioral Outcomes

| | Intervention Group | | | |
| | Story Instruction (n = 23) | | Opinion Instruction (n = 33) | |
Variable Level	BC (n = 12) M (SD)	Typical (n = 11) M (SD)	BC (n = 16) M (SD)	Typical (n = 17) M (SD)
Social Skills Rating Scale—Teacher Version:				
Problem behaviors:				
Pre	102.58 (16.83)	89.64 (10.38)	112.25 (17.14)	90.24 (8.61)
Post	100.83 (14.74)	86.73 (5.73)	109.00 (12.69)	91.71 (10.09)
Δ	−1.75 (11.14)	−2.91 (7.19)	−3.25 (13.84)	1.47 (5.10)

Note.—Social Skills Rating System (Gresham & Elliott, 1990; $M = 100$, $SD = 15$).

intervention onset and immediately following the intervention completion to provide their views of students' behavioral performance (see Table 4). The problem behavior scale includes 18 items constituting three domains: internalizing, externalizing, and hyperactivity, which are rated in terms of frequency. The measure has strong internal consistency (ranges .82–.94).

Social Validity

Social validity was assessed from the homeroom teacher's perspective, although in some cases students moved to other teachers for writing instruction (all homeroom teachers also acted as writing instructors), during pre- and postintervention phases, as homeroom teachers spent the most time with their students (see Table 2). Homeroom teachers completed the Intervention Rating Profile (IRP-15; Witt & Elliott, 1985) to obtain their opinions regarding the importance of intervention goals, the acceptability of the procedures, and importance of the intervention outcomes. Students' perspectives on social validity were assessed by having them complete the Children's Intervention Rating Profile (CIRP; Witt & Elliott, 1985).

Intervention Rating Profile. The IRP-15 examines teachers' perceptions of treatment acceptability. Teachers rate 15 statements about procedures and outcomes (e.g., I liked the procedures used in this intervention) on a six-point Likert-type scale (strongly disagree = 1 to strongly agree = 6), with total scores ranging from 15 to 90. High total scores suggest high acceptability. Internal consistency estimates range from .88 to .98. Teachers completed one rating profile for their class as a whole rather than for individual students.

Children's Intervention Rating Profile. The CIRP examines students' perceptions of treatment acceptability. Students rate seven items on a six-point Likert-type scale (I do not agree = 1 to I agree = 6). After reflecting negatively worded items, total scores ranged from 7 to 42. Like the IRP-15, high scores indicated high acceptability. The wording of the CIRP was modified as in previous studies to increase readability for the young participants. Internal consistency ranged from .75 to .89. The CIRP had been administered individually by RAs during preassessment. However, to avoid disrupting the class by again removing students to complete this brief measure, teachers administered the CIRP following intervention completion. After teachers

administered the second of the two writing prompts during the postintervention assessment phase, they administered the CIRP to their students. Teachers provided directions for completing the short survey and read each item aloud. The project director instructed all teachers not to guide the students' answers.

Experimental Design

Data were analyzed using a pre-post experimental design, with teachers randomly assigned to one of two writing conditions as described previously. Data-analytic procedures for each research question are provided in the Results section. It should be noted that during the preintervention assessment procedures, we learned that the third-grade teachers at one school ($n = 6$) had chosen to use "writing groups." Specifically, in one of the schools, homeroom students rotated to other third-grade classrooms to receive writing instruction with other students who shared similar skill sets in terms of writing skills. Consequently, for this subset of students, their homeroom teachers and writing teachers were often not the same teacher. Data completed by the writing instructor were used when analyzing Social Skills Rating Scale data.

Results

In this section, we begin by presenting findings of initial analyses used to examine group differences in preintervention scores. Then, we present findings of the research questions posed in the introduction.

Examining Preintervention Scores

Statistical analysis. A series of two-way analysis of variance (ANOVA) procedures, with two between-subject factors, were conducted to determine the extent to which the groups were equivalent prior to intervention onset. The between-subject factors included treatment group (story writing or opinion essay writing) and student status (behavior challenges or typical behavioral patterns). These analyses were conducted on the following measures: (*a*) story writing measures (number and quality of elements, overall writing quality, word length, transition words, and academic engaged time during writing), (*b*) opinion writing measures (number and quality of elements, overall writing quality, word length, transition words, and academic engaged time during writing), (*c*) overall problem behavior as measured by the Social Skills Rating System, and (*d*) social validity as measured by the IRP-15 and the CIRP. Significant ANOVAs were followed by the Tukey-Kramer modification of the honest significant difference (HSD; $\alpha = 0.05$) simultaneous confidence interval technique. This multiple-comparison technique substitutes the harmonic mean to control for (*a*) unequal group sizes and (*b*) experiment-wise Type I error.

Findings. Results indicated that there were no significant differences on any story writing, opinion writing, or social validity measures. However, the ANOVA examining overall behavior problems was significant, $F(3, 52) = 10.50$, $p < .0001$ ($R^2 = .38$). As expected, there was a significant main effect for students status, $F(1, 52) = 27.49$, $p < .0001$, with students in the behavior-challenges group having significantly higher levels of problem behaviors as rated by the teacher.

Research Questions 1, 4, and 5: Writing Performance, Differential Effects, and Academic Engaged Time

Statistical analysis. Results were analyzed using bivariate correlation and multiple regression procedures. Multiple regression procedures were used to examine how intervention group (INT; story vs. opinion), student status (STUDENT; behavior challenges vs. typical), and the interactions of these variables by group (INT × STUDENT) predicted the following outcome variables: (*a*) story writing measures (number and quality of elements, overall quality, word length, transition words, and academic engaged time during writing; see Table 5), (*b*) opinion writing measures (number and quality of elements, overall quality, word length, transition words, and academic engaged time during writing; see Table 6), and (*c*) problem behavior (see Table 7). *F*-values were examined to determine the significance of the overall model. If the model was significant, then univariate analyses were conducted to determine the unique contribution of each of the three variables in the model. *T*-tests were examined to ascertain the significant contribution of each predictor variable, controlling for the remaining five variables. Beta weights (standardized multiple regression coefficients) and uniqueness indices were examined to determine the relative value of each predictor variable contained in the model. The unique index for a given predictor was the percentage of variance in the criterion variable accounted for by that predictor variable over and above the variance explained by the remaining predictor model. While bivariate correlations examined overall relationships, semipartial correlations delineate the relation between the predictor and criterion variable, controlling for the other variables in the model.

$$Y = \beta_o + \beta_{INT} + \beta_{STUDENT} + \beta_{INT \times STUDENT} + \varepsilon.$$

This model was constructed based on the hypothesis that students with lower levels of problem behavior would be more responsive to intervention efforts. It was also hypothesized that students in either intervention group (story writing or persuasive writing) would experience improvements in the number and quality of elements and the quality of writing for the instructed genre relative to the noninstructed genre. We also hypothesized that students in the opinion essay group would demonstrate a significant increase in the number of transition words used. We did not expect students in the story group to show increases in transition words, as this was not an instructional component of the story writing condition. Effect sizes were computed using the pooled standard deviation in the denominator to determine the magnitude of the differences between intervention groups' change scores (post- minus preintervention scores) and were calculated to examine differences in growth demonstrated by each group and are reported in text. According to Cohen (1988), small, medium, and large effect sizes are 0.2, 0.5, and 0.8, respectively.

Story writing outcomes. Story writing variables included number and quality of elements, overall writing quality, word count, transition words, and academic engaged time during writing. In terms of number and quality of story elements, the three-variable model accounted for 23% of the variance in students' improvement, $R^2 = 0.22$, $F(3, 52) = 4.84$, $p = .0048$. Inspection of semipartial correlations indicated that the interaction term was not significant. Only one variable, intervention group (INT), was significant in predicting number and quality of story-elements scores, $t = -3.21$, $p = .0022$, accounting for 15.53% of the variance in story elements after con-

Table 5. Results of Multiple Regression Analyses on Story Writing Measures

	Variable									
	Elements		Quality		Word Count		Transition Words		AET	
Predictor Variable	Beta Weight Value (t)	Unique Indices Value (F)	Beta Weight Value (t)	Unique Indices Value (F)	Beta Weight Value (t)	Unique Indices Value (F)	Beta Weight Value (t)	Unique Indices Value (F)	Beta Weight Value (t)	Unique Indices Value (F)
INT	−.56* (−3.21)	.1553** (10.33)	−.44* (−2.46)	.0960* (6.04)	−.22 (−1.15)	.0239 (1.32)	−.12 (−.63)	.0292 (1.57)	−.06 (−.31)	.0153 (.84)
STUDENT	−.81 (−1.96)	.0578 (3.84)	−.86* (−2.02)	.0648* (4.08)	−.45 (−.99)	.0178 (.98)	−.50 (−1.09)	.0220 (1.18)	.36 (.79)	.0115 (.63)
INT × STUDENT	.64 (1.49)	.0335 (2.23)	.63 (1.42)	.0322 (2.03)	.29 (.61)	.0069 (.38)	.60 (1.25)	.0073 (.39)	−.43 (−.91)	.0018 (.10)
Overall model	$F(3, 52) = 4.84$, $p = .0048$, $R^2 = .22$		$F(3, 52) = 3.64$, $p = .0188$, $R^2 = .17$		$F(3, 52) = 1.06$, $p = .3745$, $R^2 = .06$		$F(3, 52) = .61$, $p = .6100$, $R^2 = .03$		$F(3, 52) = .92$, $p = .44$, $R^2 = .05$	

Note.—INT = intervention group; AET = academic engaged time. Standardized beta weights and unique indices shown reflect all variables in each model. For t-tests that tested the significance of the beta weights, $df = 52$. For F-tests that tested the significance of the uniqueness indices, $df = 1, 52$. R^2 refers to the proportion of explained variance of the overall model.

*$p < .05$.

**$p < .01$.

Table 6. Results of Multiple Regression Analyses on Opinion Writing Measures

	Outcome Variables									
	Elements		Quality		Word Count		Transition Words		AET	
Predictor Variable	Beta Weight Value (*t*)	Unique Indices Value (*F*)	Beta Weight Value (*t*)	Unique Indices Value (*F*)	Beta Weight Value (*t*)	Unique Indices Value (*F*)	Beta Weight Value (*t*)	Unique Indices Value (*F*)	Beta Weight Value (*t*)	Unique Indices Value (*F*)
INT	.36* (2.03)	.0631* (4.13)	.69**** (5.10)	.2365**** (25.99)	−.11 (−.54)	.0056 (.30)	.77**** (5.14)	.2922**** (26.45)	−.10 (−.53)	.0053 (.28)
STUDENT	−.13 (−.32)	.0059 (.39)	−.19 (−.60)	.0032 (.35)	.04 (.08)	.0000 (.01)	.63 (1.78)	.0881 (7.97)	.15 (.32)	.0018 (.10)
INT × STUDENT	.27 (.62)	.0016 (.10)	.01 (.03)	.0000 (0.00)	.04 (.09)	.0000 (.01)	−1.04** (−2.82)	.0351 (3.18)	−.06 (−.13)	.0003 (.02)
Overall model	$F(3, 52) = 4.46$, $p = .0073$, $R^2 = .20$		$F(3, 52) = 19.30$, $p < .0001$, $R^2 = .53$		$F(3, 52) = .28$, $p = .8430$, $R^2 = .02$		$F(3, 52) = 12.84$, $p < .0001$ $R^2 = .43$		$F(3, 52) = .43$, $p = .74$, $R^2 = .02$	

Note.—INT = intervention group; AET = academic engaged time. Standardized beta weights and unique indices shown reflect all variables in each model. For *t*-tests that tested the significance of the beta weights, $df = 52$. For *F*-tests that tested the significance of the uniqueness indices, $df = 1, 52$. R^2 refers to the proportion of explained variance of the overall model.

*$p < .05$.

**$p < .01$.

***$p < .001$.

****$p < .0001$.

Table 7. Results of Multiple Regression Analyses on Overall Problem Behavior and Social Validity Measures

| | Measure | | | |
| | Problem Behavior | | CIRP | |
Predictor Variable	Beta Weight Value (t)	Unique Indices Value (F)	Beta Weight Value (t)	Unique Indices Value (F)
INT	.22	.0237	.03	.0012
	(1.13)	(1.29)	(.17)	(.06)
STUDENT	.36	.0217	−.08	.0005
	(.78)	(1.18)	(−.17)	(.03)
INT × STUDENT	−.52	.0112	.12	.0005
	(−1.08)	(.61)	(.26)	(.03)
Overall model	$F(3, 52) = 0.74$, $p = .5321, R^2 = 0.04$		$F(3, 52) = 0.13$, $p < .94, R^2 = 0.01$	

Note.—INT = intervention group. Standardized beta weights and unique indices shown reflect all variables in each model. For t-tests that tested the significance of the beta weights, $df = 52$. For F-tests that tested the significance of the uniqueness indices, $df = 1$, 52. R^2 refers to the proportion of explained variance of the overall model.

trolling for the other two variables in the model. Students in the story writing group ($M = 2.22$, $SD = 3.05$) had large improvements in number and quality of story elements at the end of the intervention relative to students in the opinion writing conditions ($M = 0.14$, $SD = 2.31$, effect size = 0.78). The main effect of student was not significant, suggesting that students with and without behavioral challenges show comparable levels of improvement in terms of elements (see Table 5).

In predicting overall story quality, the three-variable model accounted for 17% of the variance in students' improvement in quality, $R^2 = 0.17$, $F(3, 52) = 3.64$, $p = .0188$. Inspection of semipartial correlations indicated that the interaction term was not significant. However, both intervention group, INT ($t = -2.46$, $p = .0174$), and students status, STUDENT ($t = -2.02$, $p = .0487$), were significant in predicting changes in story quality. Group accounted for 9.60% of the variance in improvement in story quality after controlling for the other two variables in the model, and student status (STUDENT) accounted for 6.48% of the unique variance. Students in the story writing group ($M = 1.07$, $SD = 1.92$) had greater improvements in story quality at the end of the intervention relative to students in the opinion writing conditions ($M = 0.15$, $SD = 1.67$, effect size = .51). In addition, students without behavioral challenges ($M = 1.02$, $SD = 2.05$) had greater improvements in story quality relative to students with behavioral challenges ($M = 0.04$, $SD = 1.43$, effect size = .56).

Contrary to our original hypotheses, the models predicting changes in word length as well as academic engaged time during writing were not statistically significant (see Table 5). As expected, however, the model predicting changes in transition words was statistically significant (see Table 5).

Opinion writing outcomes. Opinion writing variables also included number and quality of elements, overall writing quality, word count, transitions, and academic engaged time during writing. In terms of opinion elements, the three-variable model accounted for 20% of the variance in students' improvement, $R^2 = 0.20$, $F(3, 52) = 4.46$, $p = .0073$. Inspection of semipartial correlations indicated that the interaction term was not significant. Only one variable, intervention group (INT), was significant in predict-

ing number and quality of opinion-elements scores, $t = 2.03$, $p = .0474$, accounting for 6.31% of the variance in opinion elements after controlling for the other two variables in the model. Students in the opinion writing group ($M = 3.03$, $SD = 4.12$) had great improvements in opinion elements at the end of the intervention relative to students in the story writing conditions ($M = 0.85$, $SD = 4.00$, effect size = .54). The main effect for student was not significant. This suggests that students with ($M = 1.88$, $SD = 4.31$) and without behavioral challenges ($M = 1.00$, $SD = 4.65$) show comparable levels of improvement over SRSD-story-instructed students in terms of number and quality of elements (effect size = .20; see Table 6).

In terms of overall quality of the opinion writing essays, the three-variable model accounted for 53% of the variance in students' improvement, $R^2 = 0.53$, $F(3, 52) = 19.30$, $p < .0001$. Inspection of semipartial correlations indicated the interaction term was not significant. Intervention group (INT) was significant in predicting opinion-elements scores, $t = 5.10$, $p = < .0001$, accounting for 23.65% of the variance in story elements after controlling for the other two variables in the model. Students in the opinion writing group ($M = 2.32$, $SD = 1.15$) had great improvements in the quality of their opinion essays at the end of the intervention relative to students in the story writing conditions ($M = 0.11$, $SD = 1.08$, effect size = 1.98). The main effect for student was not significant, suggesting that students with ($M = 1.09$, $SD = 1.53$) and without behavioral challenges ($M = 1.72$, $SD = 1.55$) show comparable levels of improvement in terms of elements (effect size = .42).

In predicting transition words used in the opinion writing essays, the three-variable model accounted for 43% of the variance in students' improvement in transition words, $R^2 = 0.43$, $F(3, 52) = 12.84$, $p < .0001$. Inspection of semipartial correlations indicated that the interaction term was significant ($t = 2.82$, $p = .0067$). Specifically, students in the story writing conditions with and without behavioral challenges demonstrated comparable responses. However, students in the opinion writing group with typical behavior patterns ($M = 3.15$, $SD = 1.11$) showed greater improvement in the number of transition words used compared to students with behavioral challenges ($M = 0.91$, $SD = 2.30$) in the same group (effect size = 1.31). Contrary to our original predictions, as with story writing, the models predicting changes in word length and AET were not significant.

Research Question 2: Treatment Integrity

Statistical analysis. We report the mean session integrity scores for three aspects of treatment integrity: (*a*) the teacher self-report of each whole lesson, (*b*) the portions of sessions observed by the RA, and (*c*) the portions of sessions from teacher self-reports corresponding to the portions observed by RAs. These data were analyzed using descriptive procedures (e.g., means and standard deviations).

Findings. Overall, treatment-integrity scores were high (> 85%) from all three perspectives for both intervention groups (see Table 2). Consequently, teachers administered the two SRSD treatments as intended. Teacher self-report scores for the entire lessons were comparable for the story ($M = 94.20$, $SD = 5.01$) and opinion writing groups ($M = 97.06$, $SD = 3.87$). Effect sizes indicated medium magnitude differences (0.64), with integrity higher for the opinion writing condition.

In terms of direct observation sessions, RA observations of treatment integrity were (*a*) slightly higher than teacher ratings of the same components in the story

writing condition and (b) lower than teacher ratings of the same components in the opinion writing condition. Teachers' self-reported integrity for the full lessons and for the direct observation sessions both indicated that (according to the teachers' perspective) treatment integrity was higher for the opinion writing group relative to the story writing group. Effect sizes suggest medium- to high-magnitude difference for the full session (Cohen's $d = 0.64$) and partial session (Cohen's $d = 1.39$) scores.

Research Question 3: Social Validity

Statistical analysis. Social validity data from the teacher perspective were analyzed using descriptive procedures, with mean IRP-15 preintervention, postintervention, and change scores (post- minus prescores) examined using descriptive procedures. Effect sizes were computed to determine the magnitude of differences in teacher ratings between the story and opinion writing groups prior to intervention onset and following intervention completion. Student social validity data were analyzed in a series of 2 (group) $\times$ 2 (student status) ANOVAs comparing difference scores as well as pre- and postintervention scores.

Findings. Prior to intervention onset, teachers rated the intervention relatively favorably for students in the story writing ($M = 75.00$, $SD = 9.51$) and opinion writing ($M = 73.18$, $SD = 9.44$) conditions. After intervention completion, mean teacher ratings of students in both the story and opinion writing conditions increased to 78.00 ($SD = 12.00$) and 78.18 ($SD = 9.69$), respectively, suggesting that the intervention slightly exceeded their expectations as evidenced by low-to-medium magnitude in improvement (Cohen's d story = .28, opinion = .52).

In terms of students' perceptions, results of the three analyses revealed no significant differences on CIRP preintervention, postintervention, or changes scores, suggesting that students' perceptions were highly comparable at intervention onset and again following intervention completion, with little change in students' perceptions of social validity following intervention completion. Thus, students with and without behavioral challenges, in both conditions, rated the intervention favorably (see Table 2).

Research Question 6: Overall Problem Behavior

The three-variable model (intervention group, student status, and the interaction of these variables) was not significant in predicting changes in overall problem behavior. Consequently, the unique indices of these variables were not examined (see Table 7).

Research Question 7: Cognitive Ability as a Moderating Variable

Statistical analysis. A series of regression analyses were conducted to determine whether cognitive ability moderated any improvements in story writing measures, opinion writing measures, or overall problem behavior. To address this question, four sets of analyses were conducted to examine cognitive ability for (a) students with behavioral challenges who received story instruction, (b) students without behavioral challenges who received story instruction, (c) students with behavioral challenges who received opinion writing instruction, and (d) students without behavioral challenges who received opinion writing instruction. In each set of analyses, cognitive ability served as the predictor variable.

Findings. For students with behavioral challenges in the story writing condition, cognitive ability moderated changes in number of words written, $F(1, 10) = 9.70$, $p = .0110$, accounting for 49% (R^2) in the model ($\beta = 0.70$). Students with higher cognitive ability wrote longer stories. Cognitive ability was not a significant moderator for change on the other measures for this group. In contrast, for students without behavioral challenges who were in the story writing condition, cognitive ability did not moderate changes on any of the measures.

For students with behavioral challenges in the opinion writing condition, cognitive ability did not moderate changes in any of the measures. However, for students without behavioral challenges, cognitive ability did moderate changes in the number of elements in opinion essays, $F(1, 15) = 7.34$, $p = .0162$, accounting for 33% (R^2) in the model ($\beta = 0.57$). Again, students with higher cognitive ability wrote opinion essays with more elements. Cognitive ability did not moderate outcomes for the other measures for this subsample of students.

Discussion

In the present study, we investigated SRSD instruction for story or opinion essay writing among students with and without behavioral challenges, and whether differential effects were found for these groups of students, when delivered by the classroom teacher to all students as part of the regular class primary prevention plan within the context of a comprehensive, integrated, three-tiered model of prevention. We also addressed integrity of teacher implementation, social validity, impact on engagement among second- and third-grade students with and without challenging behaviors in these classrooms, whether or not teachers' estimates of overall behavior problems improved after SRSD instruction, and whether cognitive capabilities predicted writing gains made by students.

The Impact of SRSD Instruction at Tier 1: Research Question 1

As predicted, teacher-implemented, whole-class SRSD instruction at Tier 1 enhanced the writing performance of students with challenging behaviors and of a matched group of students without challenging behaviors. Stories written by participating students whose class received SRSD story writing instruction evidenced greater improvements in number and quality of story elements as well as story writing quality when compared to students receiving SRSD opinion writing instruction. Similarly, students who received whole-class SRSD opinion writing instruction produced arguments with more transition words, more and better opinion elements, and greater overall quality than SRSD-story-instructed students. Thus, providing SRSD instruction as part of the regular class primary prevention plan was an effective treatment for improving the writing performance of student participants.

Effect sizes for SRSD instruction ranged from .51 to 1.15 for story and opinion writing quality, respectively, and .78 to .54 for number and quality of elements for each genre. The effects for story quality obtained in this study exceeded those obtained by Tracey, Reid, and Graham (2009) in a quasi-experimental study (effect size = .35) in which the performance of all third graders in SRSD classrooms was compared to controls receiving traditional skill-based writing instruction. However, larger effect sizes for both story and opinion writing quality were obtained in an

experimental study (Lane, Little, et al., 2009) in which second- and third-grade writers with challenging behaviors received SRSD instruction individually (effect sizes for overall quality ranged from 1.14 to 1.23). Based on these statistics, it may be tempting to conclude that SRSD instruction is more effective for students with behavioral challenges when it is delivered as either a secondary or tertiary support versus a primary support for the whole class. Such a claim is premature, as this study did not make such comparisons. Additional research directly comparing whole-class to small-group and individualized SRSD instruction is needed.

Contrary to predictions, neither SRSD intervention resulted in a reliable increase in the number of words written by students with and without challenging behaviors. There was considerable variability amongst students in terms of how much they wrote at each assessment point. While the theoretical rationale for increased length of compositions is logical (i.e., word length should increase as SRSD-instructed students learn how to generate relevant ideas for a specific genre), the findings from the existing SRSD experimental studies with primary-grade students have been mixed. For example, one study found that students who struggled with writing produced longer stories and opinion papers following SRSD instruction (Graham et al., 2005), another reported an increase in words for opinion papers only (Lane et al., 2010), whereas a third investigation indicated that neither stories nor opinion papers were longer following instruction (Harris et al., 2006). In this study and these previous studies, however, SRSD instruction resulted in more complete and qualitatively better papers. Students in this and previous studies appeared to include less inappropriate text and more appropriate text after SRSD instruction. Further, increasing length was not a goal for students in these studies; future research should examine adding goals for increasing length when appropriate.

Treatment Integrity and Social Validity: Research Questions 2 and 3

Results of this study were positive in terms of both treatment integrity and social validity. Treatment integrity was high across three aspects: (*a*) teacher self-report of each whole lesson, (*b*) the portions of sessions observed by the RA, and (*c*) the portions of sessions from teacher self-reports corresponding to the portions observed by RAs. Both teachers and students believed SRSD instruction was an effective intervention; teachers rated SRSD instruction even more highly after they provided such instruction to their class. These findings provide critical impetus for widespread adoption of SRSD instruction at Tier 1.

When interpreting the social validity data, however, it is important to note that the CIRP was administered differently at pre- and postintervention assessment. It was individually administered by RAs following assenting and then by the classroom teacher within a whole-class context after the intervention concluded. This difference in procedures must be considered when interpreting outcomes, although we have no reason to expect that either method of administration would skew the results in a given direction.

The Differential Impact of SRSD Instruction: Research Question 4

We predicted that students with challenging behaviors would make smaller gains than students without such behaviors on writing measures. This prediction was

based on the assumption that students with challenging behaviors would engage in behaviors (e.g., disruptive actions) that would make SRSD less effective for them (Lane, Kalberg et al., 2009; Sutherland & Wright, in press; Walker et al., 2004). This prediction was only partially supported. For students who received SRSD story writing instruction, students with challenging behaviors made smaller gains in story quality than students without challenging behaviors. A similar result for transition words was obtained for students who received SRSD opinion writing instruction, with students with challenging behaviors using fewer of these words in their opinion text than their counterparts without challenging behaviors.

Finally, in this study (as reported in other SRSD studies, see Lane et al., 2011), teachers and students reported that opinion essay writing was easier than story writing, even after SRSD instruction. Story writing and the associated strategies are more complex, require greater creativity, and may take more effort for many students. Future research should investigate such genre effects.

Research Question 5: Academic Engaged Time

We predicted that SRSD instruction would increase engagement (defined as on-task behavior); we further predicted that SRSD-instructed students with challenging behavior would become more engaged when writing than SRSD-instructed students without challenging behaviors. As discussed, SRSD instruction includes multiple mechanisms for promoting engagement. We reasoned that students with challenging behavior would make the greatest growth as they experience considerable problems remaining on task during academic work (Sutherland & Wright, in press), providing them with greater opportunities for improvement. These predictions were not supported. Neither of the SRSD interventions significantly influenced on-task behavior, and on-task behavior was not significantly related to student type.

Nevertheless, we believe that additional research examining the effects of SRSD on students' engagement is needed. Lane et al. (2011) found that SRSD-instructed primary-grade students with challenging behaviors did make greater on-task gains when writing opinion text than control students receiving process writing and skills instruction (similar results were not found for story writing). In Lane et al.'s study, students were off task more frequently at the start of the study than students in the current study. Improved on-task behavior when writing may be dependent on initial rates of off-task behaviors. As noted previously, the behavioral supports in place in the schools in our study may have resulted in less off-task behavior before the study began. Finally, measuring off-task behavior in writing is complex. It is difficult to distinguish, for example, thinking about a composition from daydreaming. We constructed an operational definition that attempted to address this complexity. Future studies should employ a greater range of measures, including total time students spend composing as well as secondary reaction time tasks designed to measure cognitive effort (see Kellogg, 1993, for examples).

Level of Overall Problem Behaviors: Research Question 6

We investigated whether students with challenging behaviors showed an overall decrease in problem behaviors (based on teachers' estimates) after SRSD instruction. We did not make a prediction due to insufficient research specifically related to this

question in the context of schools implementing three-tiered models, and because this measure related to students' behavior across the school day, not just during writing instruction. Teachers' estimates of problem behavior among students with challenging behaviors did not decrease after SRSD instruction in either story or opinion essay writing. This finding indicates that further interventions at Tier 1, both academic and behavioral, should be explored in future research.

The Relationship between Cognitive Ability and Improved Writing: Research Question 7

We predicted that student gains in writing would be related to their overall cognitive ability. To test this prediction, we conducted separate analyses for each student type in each SRSD treatment. This prediction received limited support. Cognitive ability accounted for a statistically significant amount of variance in two instances. For SRSD-story-instructed students, more cognitively capable students with challenging behaviors wrote longer stories than those who were less cognitively capable. In addition, for SRSD-opinion-essay-instructed students, more cognitively capable students without challenging behaviors produced opinion essays with higher scores on the element measure than their less cognitively capable counterparts.

Cognitive capabilities may better predict an aspect of writing not assessed in this study (e.g., revising). Additional research is needed to replicate the current findings; explore other genres, tasks, and cognitive measures; and look at students with lower cognitive ability. Findings from this study are promising: they suggest that students in the normal range of cognitive capabilities profited from SRSD instruction.

Conclusion

This study took place in the context of rural schools implementing an evidence-based three-tiered model that supported academic, behavioral, and social development. In this district, 8.5% of students received free and reduced-price lunch, 1.39% were English learners, and 9.2% were receiving special education services in inclusive schools. Further, all teachers were credentialed, the schools were working collaboratively with a local university, and principals and teachers chose to focus on writing intervention. The positive results found in this study may be related to many of these contextual factors. Such results speak well to the impact of evidence-based practices and preventive, school-based, three-tiered approaches. Research is needed to determine generalizability to, and different outcomes in, different school contexts.

Limitations and additional directions for future research were identified throughout this discussion; a great deal more work remains to be done. We found that Tier 1 teacher-implemented SRSD instruction was effective for students with challenging behaviors and matched students without such behaviors for both narrative (stories) and expository (opinion) writing. While both groups of students profited from instruction, students without challenging behaviors made greater gains than those with challenging behaviors on some outcome measures. General education teachers implemented SRSD with fidelity; SRSD was viewed as socially valid by teachers and students. SRSD-instructed gains in writing were only minimally related to participants' cognitive capabilities. These findings are important, as previous SRSD studies with young students with challenging behaviors have involved Tier 2 or Tier 3 inter-

ventions where students were instructed in small groups or individually by research assistants. This study demonstrates that when SRSD is included as part of regular classroom instruction, it benefits students who are vulnerable due to behavioral challenges as well as those who are not.

Note

References

→ Ball, D. L., & Forzani, F. M. (2009). The work of teaching and the challenge for teacher education. *Journal of Teacher Education*, **60**(5), 497–511.

Buffum, A., Mattos, M., & Weber, C. (2010). The why behind RTI. *Educational Leadership*, **68**(2), 10–16.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cook, B. G., Smith, G. J., & Tankersley, M. (2012). Evidence-based practices in education. In K. R. Harris, T. Urdan, & S. Graham (Eds.), *American Psychological Association Educational Psychology Handbook* (Vol. **1**, pp. 495–527). Washington, DC: American Psychological Association.

→ Cutler, L., & Graham, S. (2008). Primary grade writing instruction: A national survey. *Journal of Educational Psychology*, **100**, 907–919.

→ De La Paz, S., & Graham, S. (2002). Explicitly teaching strategies, skills, and knowledge: Writing instruction in middle school classrooms. *Journal of Educational Psychology*, **94**, 687–698.

Drummond, T. (1994). *The Student Risk Screening Scale (SRSS)*. Grants Pass, OR: Josephine County Mental Health Program.

→ Gallimore, R., Ermeling, B., Saunders, W., & Goldenberg, C. (2009). Moving the learning of teaching closer to practice: Teacher education implications of school-based inquiry teams. *Elementary School Journal*, **109**, 537–553.

→ Gaskill, P. J., & Murphy, P. K. (2004). Effects of a memory strategy on second-graders' performance and self-efficacy. *Contemporary Educational Psychology*, **29**(1), 27–49.

→ Gilbert, J., & Graham, S. (2010). Teaching writing to elementary students in grades 4 to 6: A national survey. *Elementary School Journal*, **110**, 494–518.

Graham, S., & Harris, K. R. (2010). Writing. In R. Allington & A. McGill-Franzen (Eds.), *Handbook of reading disabilities research* (pp. 232–241). Mahwah, NJ: Erlbaum.

Graham, S., Harris, K. R., & Hebert, M. (2010). *Assessing writing*. New York: Carnegie Corp.

→ Graham, S., Harris, K. R., & Mason, L. H. (2005). Improving the writing performance, knowledge, and self-efficacy of struggling young writers: The effects of self-regulated strategy development. *Contemporary Educational Psychology*, **30**, 207–241.

→ Graham, S., Harris, K. R., & Sawyer, R. (1989). A components analysis of cognitive strategy instruction: Effects on learning disabled students' compositions and self-efficacy. *Journal of Educational Psychology*, **81**, 353–361.

Graham, S., & Perin, D. (2007). *Writing next: Effective strategies to improve writing of adolescents in middle and high schools.* (Report to the Carnegie Corporation). New York: Carnegie Corp.

Gresham, F. M., & Elliott, S. N. (1990). *Social Skills Rating System.* Circle Pines, MN: American Guidance Service.

Halliday, M., & Hasan, R. (1989). *Language, context, and text: Aspects of language in a social semiotic perspective.* Oxford: Oxford University Press.

Harris, K. R., & Graham, S. (2009). Self-regulated strategy development in writing: Premises, evolution, and the future. *British Journal of Educational Psychology* (monograph series), no. 6, 113–135.

Harris, K. R., Graham, S., Brindle, M., & Sandmel, K. (2009). Metacognition and children's writing. In D. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Handbook of metacognition in education* (pp. 131–153). Mahwah, NJ: 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**, 295–340.

Harris, K. R., 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., 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**(2), 103–119.

Kellogg, R. T. (1993). Observations on the psychology and thinking of writing. *Composition Studies*, **21**, 3–41.

→ Lane, K. L., Graham, S., Harris, K. R., Little, M. A., Sandmel, K., & Brindle, M. (2010). Story writing: The effects of self-regulated strategy development for second grade students with writing and behavioral difficulties. *Journal of Special Education*, **44**, 107–128.

→ Lane, K. L., Harris, K., Graham, S., Driscoll, S. A., Sandmel, K., Morphy, P., . . . Schatschneider, C. (2011). Self-regulated strategy development at tier 2 for second-grade students with writing and behavioral difficulties: A randomized control trial. *Journal of Research on Educational Effectiveness*, **4**, 322–353.

→ Lane, K., Harris, K. R., Graham, S., Weisenbach, J., Brindle, M., & Morphy, P. (2008). The effects of self-regulated strategy development on the writing performance of second grade students with behavioral and writing difficulties. *Journal of Special Education*, **41**, 234–253.

Lane, K. L., Kalberg, J. R., & Menzies, H. M. (2009). *Developing schoolwide programs to prevent and manage problem behaviors: A step-by-step approach.* New York: Guilford.

→ Lane, K. L., Little, M. A., Casey, A. M., Lambert, W., Wehby, J. H., Weisenbach, J. L., & Phillips, A. (2009). A comparison of systematic screening tools for emotional and behavioral disorders: How do they compare? *Journal of Emotional and Behavioral Disorders*, **17**, 93–105.

→ Lienemann, T., Graham, S., Leader-Janssen, B., & Reid, R. (2006). Improving the writing performance of struggling writers in second grade. *Journal of Special Education*, **40**, 66–78

Little, M. A., Lane, K. L., Harris, K., Graham, S., Brindle, M., & Sandmel, K. (2010). Self-regulated strategies development for persuasive writing in tandem with schoolwide positive behavioral support: Effects for second grade students with behavioral and writing difficulties. *Behavior Disorders*, **35**, 157–179.

→ Mason, L. H., & Shriner, J. G. (2008). Self-regulated strategy development instruction for writing an opinion essay: Effects for six students with emotional/behavior disorders. *Reading & Writing: An Interdisciplinary Journal*, **21**, 71–93.

Nelson, R., Benner, G., Lane, K., & Smith, B. (2004). Academic achievement of K–12 students with emotional and behavioral disorders. *Exceptional Students*, **7**, 159–173.

→ Pressley, M., Woloshyn, V., Lysynchuck, L., Martin, V., Wood, E., & Willoughby, T. (1990). A primer of research on cognitive strategy instructional: The important issues and how to address them. *Educational Psychology Review*, **2**, 1–58.

Reid, R., & Lienemann, T. (2006). Self-regulated strategy development for written expression with students with attention deficit/hyperactivity disorder. *Exceptional Children*, **73**, 53–68.

Sandmel, K., Brindle, M., Harris, K. R., Lane, K., Graham, S., Little, A., . . . Mathias, R. (2009). Making it work: Differentiating tier two writing instruction with self-regulated strategies development in tandem with schoolwide positive behavioral support for second graders. *Teaching Exceptional Children*, **42**, 22–33.

Sattler, J. (1991). *Assessment of children* (3rd ed.). San Diego, CA: Sattler.

Sugai, G., & Horner, R. R. (2006). A promising approach for expanding and sustaining school-wide positive behavior support. *School Psychology Review*, **35**, 245–259.