The Journal of Educational Research

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/vjer20

Supplemental intervention improves writing of first-grade students: Single case experimental design evaluation

Meaghan McKenna, Howard Goldstein, Xigrid Soto-Boykin, Ke Cheng, Gary A. Troia & John Ferron

To cite this article: Meaghan McKenna, Howard Goldstein, Xigrid Soto-Boykin, Ke Cheng, Gary A. Troia & John Ferron (2021): Supplemental intervention improves writing of first-grade students: Single case experimental design evaluation, The Journal of Educational Research, DOI: 10.1080/00220671.2021.1923450

To link to this article: https://doi.org/10.1080/00220671.2021.1923450

Published online: 22 May 2021.

THE JOURNAL OF EDUCATIONAL RESEARCH
https://doi.org/10.1080/00220671.2021.1923450

Supplemental intervention improves writing of first-grade students: Single case experimental design evaluation

Meaghan McKenna[a], Howard Goldstein[a], Xigrid Soto-Boykin[b], Ke Cheng[c], Gary A. Troia[d], and John Ferron[c]

[a]Department of Communication Sciences and Disorders, University of South Florida, Tampa, Florida, USA; [b]Juniper Gardens Children's Project, University of Kansas, Kansas City, Kanas, USA; [c]Department of Educational Measurement and Research, University of South Florida, Tampa, Florida, USA; [d]Department of Counseling, Educational Psychology, and Special Education, Michigan State University, East Lansing, Michigan, USA

ABSTRACT

The limited research available on writing in Grade 1 led to the development and implementation of an intervention for students who were performing below expectations. Ten students participated in a writing intervention for 11–13 weeks. A multiple baseline design across three units of instruction was focused on (a) paragraph structure, (b) sentence structure and handwriting, and (c) vocabulary and spelling allowed for analysis of the effects of the intervention. Treatment effects were evident from visual analysis, nonoverlap statistics, and multilevel modeling. Descriptive data collected on literacy measures administered before and after the intervention also indicated growth. Educator ratings of student writing and social validity surveys provide further evidence that improvements in student writing were apparent. Students also provided favorable input. These results indicate the malleability of writing behavior in at-risk first-grade students. Although preliminary findings are promising, iterative development would help improve this intervention and determine its efficacy with a broader sample of students.

ARTICLE HISTORY
Received 1 February 2021
Revised 1 April 2021
Accepted 26 April 2021

KEYWORDS
First grade; literacy; writing; intervention

Although the National Commission on Writing (2006) and National Early Literacy Panel (2008) have stressed the need for writing instruction for young children, writing research has attracted little attention compared with research on reading and mathematics (Mo et al., 2014; Puranik & Al Otaiba, 2012). Meta-analyses have synthesized the results from several writing intervention studies, but few studies have enrolled first-grade students (Gillespie & Graham, 2014; Graham et al., 2012; Rogers & Graham, 2008). Moreover, educators are too often unprepared to teach writing during core instruction due to inadequate understanding, skills, or materials (Graham et al., 2013; Graham & Harris, 2013).

Incorporating writing into early childhood curricula is important because it supports both language and literacy, which provide the foundation for academic achievement (Gerde et al., 2012). Although research-based interventions are available to address reading, an evidence-based approach targeting writing in early primary grades is lacking (Kim et al., 2015). The development of a systematic and efficacious intervention targeting early writing is needed to assist students who experience difficulty responding to core writing instruction.

Writing instruction and intervention

Low-level transcription skills (handwriting and spelling) are targeted most frequently in early elementary grades when students experience difficulty writing (Baker et al., 2003; National Commission on Writing, 2006). However, generating ideas, translating ideas into language, and transcribing that language into written text are all necessary components of writing (Kim et al., 2015). The Simple View of Writing (Juel et al., 1986) involves ideation and transcription. Ideation is related to oral language and includes the planning, generation, and organization of a written piece. Transcription denotes handwriting fluency and spelling (Kim et al., 2015). The Simple View of Writing outlines the skills needed for learning and teaching beginning writing (Berninger et al., 2009). Applying the implications from this theory in the development and implementation of writing interventions is important in ensuring children are exposed to every aspect of the writing process, not just to lower-level writing skills such as handwriting and spelling. The Common Core State Standards (CCSS) provide benchmarks for students to meet at each grade level beginning in kindergarten (http://www.corestandards.org/; Graham et al., 2012, 2013). According to the CCSS, writing is a tool for learning, which serves a variety of purposes (Graham et al., 2012, 2015). CCSS address foundations for writing acquisition, handwriting, production of grammatical sentences, conventions, and spelling. Standards for first-grade students include knowledge of text types and purposes (i.e., opinion, informational, and narrative writing), production and

CONTACT Meaghan McKenna meaghanmckenna@ku.edu Juniper Gardens Children's Project, University of Kansas, 444 Minnesota Ave. #300, Kansas City, KS 33612-3807

distribution of writing (i.e., peer collaboration), and research to build and present knowledge (e.g., writing a how-to piece). Unfortunately, the writing benchmarks in CCSS and the prevailing writing theories have largely been ignored (Troia & Olinghouse, 2013). Teachers experience difficulty interpreting the standards and generating associated lessons (Hill, 2001). Few curricula address standards explicitly (Krajcik et al., 2008). Teachers report limited preparation and continuing education opportunities for implementing evidence-based writing practices (Applebee & Langer, 2011; Gilbert & Graham, 2010).

Writing intervention research is needed so that supplemental instruction can be provided to students who are struggling both with writing and meeting CCSS benchmarks (Graham & Harris, 2013). The self-regulated strategy development (SRSD) model is an empirically validated method for teaching writing (Harris & Graham, 1985). When implementing SRSD, students learn how to use a targeted strategy to complete an academic task (Saddler, 2006). Assistance is faded as students' progress to independent use of instructed skills (Troia & Graham, 2003). Troia and Graham (2003) proposed that one should provide explicit instruction (e.g., modeling, guided practice, independent practice) that focuses on all aspects of writing and different writing genres. However, there is limited research on intervention for composing in specific genres (Harris et al., 2015).

Students benefit from opportunities to write for a variety of purposes (Donovan, 2001). Learning about genres is part of literacy development, and the CCSS indicate the value of young children learning to write across several genres (Chapman, 2002). However, the few studies conducted with students in Grades 1 and 2 have addressed narrative writing (e.g., Lane et al., 2011; Saddler, 2006; Spencer & Petersen, 2018; Zumbrunn & Bruning, 2013) because it is deemed most appropriate for young children (Chapman et al., 2007). Although narrative writing is important, the focus on narrative writing and limited exposure to informational writing may be reasons why children experience difficulty with expository writing when they enter upper elementary grades (Duke, 2000; Gee, 2001). Focusing on narrative writing may limit opportunities to address genres students prefer or genres important for future success (Donovan & Smolkin, 2002). A number of investigators identify a need to study nonnarrative writing genres (Chapman, 2002; Donovan & Smolkin, 2011; Duke, 2000; Tower, 2003).

Present study

Differences in student writing ability are evident in early elementary school (Puranik & Lonigan, 2014), and teachers often are unsure of how to deliver appropriate interventions to their students (Taft & Mason, 2011). Most research with first-grade students has addressed handwriting and spelling (e.g., Graham et al., 2018; Wolf et al., 2017). The only two studies with first-grade students to focus on ideation have addressed narrative writing (i.e., Spencer & Petersen, 2018; Zumbrunn & Bruning, 2013). First-grade students also need to learn informational and opinion writing because less attention has been placed on these genres in early elementary grades, yet students are required to compose using them. Thus, a standards-aligned intervention that targets both ideation and transcription writing skills and addresses informational and opinion writing for first-grade students would be useful to teachers. Because of the positive effects of SRSD with students in Grade 2 through high school during small group instruction (Mason et al., 2011), this model was used when designing the intervention.

The aims of the present study include (a) determining if the writing intervention using the SRSD model improved the informational and opinion writing of first-grade students, (b) descriptively analyzing pre- and post- intervention performance on standardized measures of reading and writing measures to determine if student scores increased, (c) identifying if the scoring of writing samples by instructional staff revealed improvements in student writing, and (d) gathering social validity data on teachers' and students' impressions of the writing intervention. Because of the large number of transcription and ideation skills, research also is needed to determine how skills can be grouped together during instruction to aid implementation efficiency. Thus, this exploratory study addressed four research questions:

Research Question 1: To what extent does a writing intervention targeting three units of instruction (Unit 1: organization and topic maintenance; Unit 2: sentence structure, upper/lowercase letter use, punctuation, and handwriting; and Unit 3: spelling and vocabulary) improve the writing performance of first-grade students experiencing delayed writing skills?

Research Question 2: How do students perform on standardized measures of reading and writing administered before and after the intervention?

Research Question 3: Does blind scoring by instructional staff (teachers and a writing coach) reveal differences in student performance in writing samples drawn from baseline, intervention, and maintenance conditions?

Research Question 4: To what extent did teachers and students express satisfaction with and perceive benefits of the writing intervention?

Method

Participants

Ten first-grade students (3 boys and 7 girls) participated in a writing intervention for three and a half months. They attended a Title I public elementary school in Florida in which 98% of the students qualified for free or reduced lunch, 7% were dual language learners, and 19% had Individualized Education Programs. All of the students in this study were eligible for free or reduced lunch. All of the girls were African American, two of the boys were African American, and one of the boys was European American. The students ranged in age from 6 years to 6 years and 11 months. After receiving approval from the Institutional Review Board, classroom teachers and the writing coach

Table 1. Student demographic information.

Student	Age (Years)	Sex	Ethnicity	CELF-5 CL SS $M = 100$ $SD = 15$	i-Ready Overall Score
1	6.4	Female	African American	101	407[a]
2	6.1	Female	African American	81	392[a]
3	6.2	Female	African American	102	413[a]
4	6.11	Male	African American	87	410[a]
5	6.5	Male	European American	101	410[a]
6	6.4	Female	African American	89	428[a]
7	6.9	Female	African American	92	419[a]
8	6	Female	African American	82	360[a]
9	6.1	Female	African American	84	415[a]
10	6.3	Male	African American	104	403[a]

Note. CELF-5 = Clinical Evaluation of Language Fundamentals-Fifth Edition; CL = core language; SS = standard score.
aBelow-grade-level performance on the i-Ready Reading Assessment.

nominated students to participate. Only students with a report card grade of below proficient in writing (minimally, considerably, or substantially) were included. Students were excluded if they unable to recognize and name upper and lowercase letters, unable to write their letters, or unable to identify letter sounds. Parental consent was obtained from all students who met the inclusion criteria prior to beginning the study.

The Clinical Evaluation of Language Fundamentals–Fifth Edition (CELF-5; Wiig et al., 2013) was administered to describe participants' language abilities prior to the study. All 10 students were performing below grade level in reading according to the results of the Fall i-Ready Reading Assessment (2019). Participant demographic information and performance on language and reading assessments administered before the intervention are presented in Table 1. All students remained enrolled for the entire study.

Experimental design

A multiple baseline design across units of instruction was used to evaluate the relation between treatment and the learning of writing skills (Baer et al., 1968). The Simple View of Writing (Juel et al., 1986) along with input from practitioners and policymakers and a pilot study were used to derive the scope and sequence of instruction. Seven students who completed kindergarten and were entering Grade 1 participated in a pilot study during a summer program. They received five weeks of small group writing instruction 2–3 times per week. Instruction addressed eight writing skills: upper/lowercase letter use, punctuation, handwriting, spelling, sentence structure, vocabulary, topic maintenance, and organization. All students made gains in: punctuation, sentence structure, vocabulary, topic maintenance, and organization. Six students made gains in upper/lowercase letter use, five students made gains in spelling, and three students made gains in handwriting. The tools, resources, and materials were further refined based on responses of children.

The eight writing skills were divided into three units of instruction. Unit 1 addressed organization and topic maintenance; Unit 2 addressed sentence structure, letter case, punctuation, and handwriting; and Unit 3 addressed spelling and vocabulary. The multiple baseline design across units of instruction, rather than across participants, was selected because extensive baseline testing placed undue burden and test fatigue on students who were unlikely to acquire writing skills without direct instruction (Gast & Ledford, 2014; Kennedy, 2005). The multiple baseline design across units allowed investigation of how students responded to the grouping of skills addressed in each unit of instruction. This design potentially allowed for 3 within-subject and 10 between-subject replications. Demonstrated improvements associated with the staggered initiation of intervention are used to judge whether replications are sufficient to give one confidence in the internal validity of the findings (Gast et al., 2014).

At least three informational, three opinion, and two narrative writing samples were collected from every student before the study began. The graphed data for all writing samples was used to determine when it was appropriate to begin the intervention. A stable baseline with no upward trends was obtained for all of the students prior to beginning Unit 1. The units were addressed in the same order for all students who received the intervention. Although this is contrary to the ideal multiple baseline design, the findings from the pilot study indicated that instruction be implemented in the order of each of the units. Baseline versus intervention performance was inspected to compare levels of performance, trends, and immediacy of effects. Once intervention commenced, all of the students in the group needed to achieve at least three scores above baseline before the next unit of instruction began. However, after Unit 1 was introduced, ascending baseline trends in subsequent units were present for several students. Before moving on to the next unit of instruction, baseline stability was reestablished with at least three data points.

To address the first research question, multilevel modeling was conducted to summarize results across participants and to augment the visual analysis used to examine individuals' behavior change. To address the second research question, pretest-posttest data on the writing and reading measures were evaluated descriptively and by using nonparametric statistical comparisons. To address the third research question, instructional staff ratings were used to determine if ratings given by individuals blind to the authors and the timing of writing samples indicated differences in student performance across conditions. Teacher and student social validity survey ratings were summarized to address satisfaction with and perceptions of the writing intervention.

Measures

Writing samples

Writing samples were collected across three experimental conditions. The writing samples were untimed; at the beginning of the study students spent about 10 min writing and as the study progressed they wrote for approximately 25 min. Students wrote independently and gave their writing paper to Meaghan McKenna when finished. The prompts administered were counterbalanced across all three conditions. The students wrote a response to each prompt once. When administering each of the writing prompts, Meaghan McKenna said, "Today, you are going to do your best [name of genre] writing. You can pick what you would like to write about: [first writing prompt choice] or [second writing prompt choice]. What do you want to write about today?" Then, the student selected their choice. After that, Meaghan McKenna said, "Please let me know when you are all finished writing."

The initial baseline lasted 12 school days. In the initial baseline, 8–13 writing samples were collected (at least three informational, three opinion, and two narrative samples). Intervention lasted 58–64 school days. During intervention, 11–13 writing samples were collected, which included a rotation between informational and opinion writing (approximately six samples per genre). A writing checklist used in training was made available to students during independent writing activities to prompt self-monitoring while writing. Students also had access to the materials used during intervention sessions (graphic organizer, spacing stick, etc.). After the students completed their independent writing piece, they had the opportunity to review their work and use the check box underneath the symbols on the writing checklist to indicate that evidence of the skill was present in their paper. This also provided students with an opportunity to edit their work. Maintenance lasted for 45–52 school days. Follow-up data collection began one month after the last intervention session. During maintenance, 7–9 samples were collected (at least three informational, three opinion, and one narrative sample). When maintenance samples were collected, students no longer had access to intervention materials.

Researcher-developed writing scoring rubric

An analytic rubric developed by Meaghan McKenna was used to score students' writing samples. Analytic assessment involves analyzing writing by defining a specific set of skills and using criteria to determine student performance (Brookhart, 1999; Coker & Ritchey, 2010; Huot, 1990). Prior to this study, elementary school teachers participated in interviews and practiced using the scoring rubric. These activities provided preliminary content validity evidence to ensure that the skills in the rubric were necessary for assessing early elementary writing.

The rubric contained scoring criteria for the eight writing skills targeted in the instructional units: organization, on topic, sentence structure, letter case, punctuation, handwriting, vocabulary, and spelling. Each skill was scored on a 5-point Likert-type scale ranging from 0 (*never occurring*) to 5 (*occurring almost always*). The entire rubric was used to score all of the writing samples collected, but when graphing each student's score, the maximum score that could be obtained in Units 1 and 3 was 10 points because these units contained two skills that were evaluated for response to treatment. The maximum score obtained in Unit 2 was 20 points because this unit contained four skills.

Kaufman Test of Educational Achievement–Third Edition

The written expression and spelling subtests of the Kaufman Test of Educational Achievement–Third Edition (KTEA-3; Kaufman & Kaufman, 2014) were administered. During the written expression subtest, students wrote sentences from dictation, added capitalization and punctuation, and wrote an essay. During the spelling subtest, students wrote letters that represented sounds and then words from dictation. Two parallel forms were used to evaluate students before and after the intervention. The split-half reliability coefficients of the KTEA-3 written expression subtest for first-grade students is .90 and .96 for spelling. The alternate form reliability for first-grade students is .80 for written expression and .91 for spelling.

i-Ready

The i-Ready Reading (2019) assessment is a universal screener administered three times per year. Data from the initial assessment administered at the end of August (i.e., preintervention) and the second assessment administered in January (i.e., postintervention) were used as descriptive pre and post measures. Overall reliability of i-Ready Reading (2019) subtests ranges from .76 to .80.

Instructional staff ratings

Two first-grade teachers and the writing coach who identified if students met the study's inclusion criteria and taught the students enrolled in the study scored writing samples collected during the baseline, intervention, and maintenance conditions. They were unaware of who produced each sample and when it was collected. A rating rubric developed by Meaghan McKenna allowed the teachers and writing coach to rate the child's writing sample and all eight writing skills targeted in the instructional units using a 4-point scale (1 = poor, 2 = below average, 3 = average, 4 = above average). They were asked to rate writing proficiency based on their expectations for first-grade student writing throughout the first three quarters of the school year.

Social validity measures

Social validity information was elicited from teachers and students. The teachers completed a paper survey and the students answered survey questions presented orally by Meaghan McKenna. Surveys contained questions about perceptions of and satisfaction with the intervention. The teachers also reported if the skills targeted in the small group generalized to classroom writing activities.

Procedure

No writing instruction occurred during baseline except what the teachers typically provided. All of the teachers were providing similar writing instruction based on what was planned during team meetings held throughout the week with the writing and reading coaches. The teachers also had instructional guides that listed specific skills or activities to focus on when teaching writing.

The intervention consisted of three units of instruction. During the first unit students received paragraph structure instruction. They learned how to discriminate between the informational and opinion genres and to use a graphic organizer to plan for writing a single paragraph paper. In the second unit students received grammar, usage, sentence formation, and handwriting instruction. The students practiced revising and editing incomplete sentences and producing sentences that varied in structure and length. Capitalization and punctuation were addressed during the editing activities and when putting together the sentences. Sentence expansion was targeted using question words (e.g., *when*, *where*). Formation of neat letters that were the appropriate size was targeted through repeated practice opportunities to address handwriting. A spacing stick was provided to assist with making adequately sized spaces between words. The students learned to write along the entire line. The third unit addressed vocabulary and spelling. Vocabulary activities included word maps (Westby, 2011) and questions to prompt students to think about attributes for comparing and contrasting words (Montgomery, 2013; Zimmerman, 2007). Students practiced replacing words within written passages to increase the variety of vocabulary and they selected parts of speech (e.g., nouns, verbs, adjectives, etc.) to add to sentences. To address spelling, each student wrote commonly misspelled words on their individual word wall including words from the Basic Spelling Vocabulary list (Graham et al., 1993). Difficult vowel patterns (short and long) were taught using explicit phoneme-grapheme mapping (Berninger et al., 1998; Moats, 2005).

During the intervention condition, the students received instruction two times per week during Units 1 and 2 and three times per week during Unit 3 for approximately 25 min per session. Meaghan McKenna, a licensed and certified speech-language pathologist, implemented the intervention. The 10 students attended school in three different Grade 1 classrooms, naturally forming three intervention groups. There were three students in two groups and four students in one group. The students remained in the same group throughout the intervention. All research activities took place in an empty classroom.

Intervention lessons were explicit and systematic. Each lesson contained an outline with general instruction and a script for implementation. The interventionist used the SRSD (Harris & Graham, 1985) model. A writing checklist assisted students with setting goals, self-monitoring, using self-instructions, self-evaluating, and self-reinforcing (Zito et al., 2007). Consistent with SRSD stages, lessons included: (a) reviews of skills and strategies previously introduced, (b) goal setting, (c) instruction targeting the skill or strategy and orientation to the activity, (d) modeling, (e) guided or collaborative practice, (f) independent practice, and (g) review or restatement of the goal addressed at the conclusion of the lesson (Graham & Harris, 1993). The interventionist followed the outline and script to ensure lessons were standardized. However, it was not essential that the exact words were used in every lesson or for every student. Differentiation occurred based on the students' response to instruction. For example, additional modeling, guided practice, or independent practice opportunities were provided based on how successful students were with answering questions and completing the activity addressed in the lesson. During the initial intervention lesson, students were introduced to the genres. Visual symbols represented each genre. For informational writing, an illustration showing a sequence of picture cards for the steps of an activity represented "how to" writing; an illustration showing labels corresponding to pictures represented facts; and an illustration of a teacher with a group of children and lit lightbulbs above the children represented that the genre involves using acquired knowledge. For opinion writing, pictures illustrated children making or thinking about choices and various faces depicting different feelings. The structure for all subsequent lessons included the SRSD stages.

Review of skills and strategies previously introduced

Before beginning a new lesson, students were asked to recall what had previously been addressed. The students would report on the specific skills targeted (e.g., organization, complete sentences, spacing). They also provided a brief summary of the strategies and/or practice activities that had been completed (e.g., "we learned how to use a graphic organizer," "we fixed incomplete sentences," "we used a spacing stick to make spaces in between words"). Then, a goal for the lesson was set.

Setting a goal

A writing checklist was used for setting goals. Visual picture symbols (e.g., index finger pointing up for *spacing*, two children talking about the same item for *on topic*, etc.) represented each of the writing skills addressed. The symbols were introduced as skills were targeted. The checklist and symbols had Velcro backings so symbols could be placed on or taken off the checklist as needed. As students became familiar with the symbols, they would say, "Today we are going to work on organization and on topic." The interventionist would then add a more specific goal aligned with the lesson (e.g., "Today, we will read sentences related to the informational prompt: Write about how to make a sandcastle. You will select the best topic sentence, facts, and closing sentence that are related to [go with] the topic."). Following the goal, instruction targeted the skill or strategy and orientation to the activity.

Instruction targeting the skill or strategy and orientation to the activity

The interventionist and students would discuss the skills to be learned and when and how they would be used. For example, students learning about punctuation heard, "We can use punctuation marks within our sentences or at the end of our sentences." Then, each of the marks was discussed with a visual corresponding to the punctuation marks. "Who can tell me what this is (show visual)?" After students provided an answer, further explanation was given, "A period comes at the end of a sentence if you are done telling something." After each punctuation mark was introduced, the visuals were placed on desks. One student identified a mark and another student picked up the visual corresponding to the mark. After they received instruction for the skill, activities were introduced. Modeling, guided or collaborative practice, and independent practice were provided as needed.

Modeling

When new activities were introduced, a clear example of the expected outcome was modeled. For example, "This sentence says, 'Milk is used to make ice cream'. I am going to turn my sign to the fact side, to show that I know this is stating a fact because this is something I know is true. I could prove it to be true by looking at the ingredients on an ice cream container, reading a book, or researching online." The interventionist also modeled how to use the writing checklist for self-monitoring (e.g., using the symbols as reminders while writing and reviewing work), self-evaluating (e.g., placing a check mark under each symbol when shared evidence of how a strategy or skill was used in their writing), and self-reinforcing (e.g., showing students how to positively praise themselves for successful completion of activities).

Guided or collaborative practice

Guided practice allowed the students and the interventionist to complete the activity together. For example, when using words and picture cards to create sentences that corresponded to a writing prompt, each student received the same bag of pictures. The entire group worked together to unscramble the sentence and add it to their paper. Strategies such as looking for the word(s) with an uppercase letter because a sentence always begins with an uppercase letter or finding the "naming part" and "telling part" of the sentence were given. Students asked and answered questions, stated what would come next, received feedback throughout the activity, along with prompting and scaffolding as needed. Students moved ahead or assisted other group members as they became more independent. Students self-monitored their work by circling the letters they formed that looked the best. Self-evaluation was used to monitor progress while completing activities. Students placed checks on their checklist as they successfully completed activities.

Independent practice

During independent practice, students were asked to complete activities by themselves (e.g., picking up their spacing stick and using it to make appropriate-sized spaces without prompts). They would use self-reinforcement when the expectation for an activity was met (e.g., "I did an amazing job making spaces between the words on the page.").

Review or restatement of the goal

After the lesson concluded, students were asked to review what was addressed during the intervention. When working on handwriting the students said, "Today we practiced forming neat letters. We used the lines on our paper to help with size."

Conferencing

One-on-one conferences were held with students during the intervention phase. After an independent writing sample was collected, students met with the interventionist for approximately five minutes to discuss their self-evaluation and receive feedback. Students used their writing checklist while reviewing their work. Areas of strength were highlighted using a colored marker by the interventionist. Revisions using a colored pencil or maker were made by the students while discussing areas for improvement (e.g., adding a closing sentence).

Intervention implementation fidelity

An observational checklist containing the procedures for the intervention (e.g., setting the goal, completing a practice activity) was used to calculate implementation fidelity. Two research assistants (enrolled in an undergraduate communication sciences program) who were unfamiliar with the intervention but trained to use the observational checklist assessed implementation fidelity of 25% ($n = 23$ of 93) of randomly selected lessons. The research assistants listened to audio recordings of the lessons and used the observational checklist to calculate the percentage of lesson components accurately implemented. Lessons were implemented with 100% fidelity.

Interrater reliability

Xigrid Soto-Boykin, who was unfamiliar with the intervention, used the researcher-developed writing scoring rubric to rescore 33% of the student writing samples to determine interrater reliability. Overall reliability was calculated along with the reliability for all eight of the skills. This was done because it was the first time this rubric was used for research. It also allowed for analysis of the agreement for each skill. Reliability was calculated by dividing the number of agreements by the total number of ratings. Kappa was calculated using SPSS Statistics Software Version 25. An examination of interrater reliability was conducted for 42 of 91 samples in baseline, 45 of 124 samples collected in intervention, and 30 of 81 samples collected in maintenance,

Table 2. Interrater agreement and kappa reliability estimates.

Participant	Overall Reliability	Organization	On Topic	Sentence Structure	Upper/ Lowercase Letters	Punctuation	Handwriting	Vocabulary	Spelling
1	94%	100%	93%	93%	93%	93%	93%	93%	93%
2	85%	90%	70%	100%	70%	80%	90%	90%	80%
3	93%	93%	93%	100%	93%	80%	87%	100%	100%
4	91%	91%	82%	100%	100%	82%	100%	82%	91%
5	95%	91%	82%	100%	100%	100%	91%	100%	100%
6	94%	91%	91%	73%	100%	100%	100%	100%	100%
7	95%	100%	83%	92%	100%	100%	100%	100%	83%
8	90%	91%	91%	91%	100%	82%	91%	100%	73%
9	92%	82%	91%	91%	91%	100%	73%	91%	100%
10	90%	82%	91%	91%	73%	100%	82%	100%	100%
Average	92%	91%	87%	93%	92%	92%	91%	96%	92%
Participant	Overall reliability Kappa	Organization Kappa	On topic Kappa	Sentence structure Kappa	Upper/ Lowercase Kappa	Punctuation Kappa	Handwriting Kappa	Vocabulary Kappa	Spelling Kappa
1	0.92	1	0.90	0.87	0.89	0.91	0.90	0.90	0.87
2	0.81	0.87	0.61	1	0.58	0.75	0.86	0.86	0.76
3	0.92	0.90	0.91	1	0.90	0.71	0.82	1	1
4	0.89	0.87	0.75	1	1	0.76	1	0.76	
5	0.94	0.88	0.76	1	1	1	0.83	1	1
6	0.93	0.86	0.85	0.67	1	1	1	1	1
7	0.94	1	0.75	0.89	1	1	1	1	0.73
8	0.87	0.88	0.88	0.86	1	0.74	0.86	1	0.57
9	0.87	0.74	0.88	0.88	0.87	1	0.64	0.87	1
10	0.88	0.75	0.86	0.87	0.59	1	0.75	1	1
Average	0.90	0.88	0.82	0.90	0.88	0.89	0.87	0.94	0.88

(i.e., 33–48% for each participant). Point-by-point agreement ranged from 85% to 95% and averaged 92%. The mean interrater agreement percentage for each skill was 91% for organization, 87% for on topic, 93% for sentence structure, 92% for upper/lowercase letter use, 92% for punctuation, 91% for handwriting, 96% for vocabulary, and 92% for spelling. Fleiss' kappa coefficient ranged from .81 to .94, indicating very high agreement (Landis & Koch, 1977). The average kappa coefficient for overall writing was .90. The mean kappa for each skill was .88 for organization, .82 for on topic, .90 for sentence structure, .88 for upper/lowercase letter use, .89 for punctuation, .87 for handwriting, .94 for vocabulary, and .88 for spelling. When exact agreement was not present, the discrepancy between the two scorers (Meaghan McKenna and Xigrid Soto-Boykin) was never more than 1 point. Table 2 contains the breakdown of the interrater agreement and kappa coefficients for each student.

Results

Effects of the intervention on first-grade student writing

Data were analyzed visually according to within- and between-phase patterns of responding with respect to level, trend, variability, overlap, and immediacy of effect (Gast & Spriggs, 2014). Visual analysis was conducted to determine the effects of intervention on student writing skills as measured by the writing rubric. Figures 1–3 depict students' performance during baseline, intervention, and maintenance conditions across all three units. The x-axis represents the writing samples collected and the y-axis represents the writing rubric score.

Visual analyses were consistent with a functional relationship between the intervention and the outcome measures (Gast et al., 2014; Kennedy, 2005), as every student made gains in their independent use of writing skills in a staggered fashion. During Unit 1, students made 3–6 point gains and the intervention was effective for all students. The highest score (10) on the rubric was obtained by all of the students at least one time in Unit 1.

However, as intervention lessons were implemented for Unit 1, there was a rise in Unit 2 baseline data for Students 1, 4, 5, 6, 7, and 9. Although there were upward trends in Unit 2 baselines for Students 1, 4, and 7, baseline stabilized for at least three sessions prior to beginning intervention on Unit 2 skills. All of the students made gains during the intervention phase of Unit 2. Because there were four skills in this unit, the highest rubric rating score obtained for this unit was 20. Students 1, 2, 5, 6, and 9 all obtained scores of 18 at least once. Student 9 also obtained a score of 19 once. The gains made in Unit 2 were between 5 and 9 points and the intervention was effective for all of the students.

Many of the students achieved high scores in Unit 3 during baseline data collection. As the strategies were introduced during Units 1 and Unit 2, all 10 students demonstrated a slight upward trend in Unit 3. During Unit 3, nine of the ten students all obtained scores of 6 or 7 at baseline leaving little opportunity to show gains after the instruction began in this unit. The overall increase in the scores of each student during the intervention phase for Unit 3 was small. The participants' scores only increased by 1–2 points during this unit. Student 1 and Student 7 were the only two students to obtain a score of 9 of the 10 possible points. The intervention in Unit 3 was ineffective for Student 2.

Follow-up data collection began a month after the last intervention session. Writing samples were collected across a span of three months. All of the students maintained the gains made during the intervention in Units 1 and 2. However, for Unit 1, Students 3 and 7 each had one data

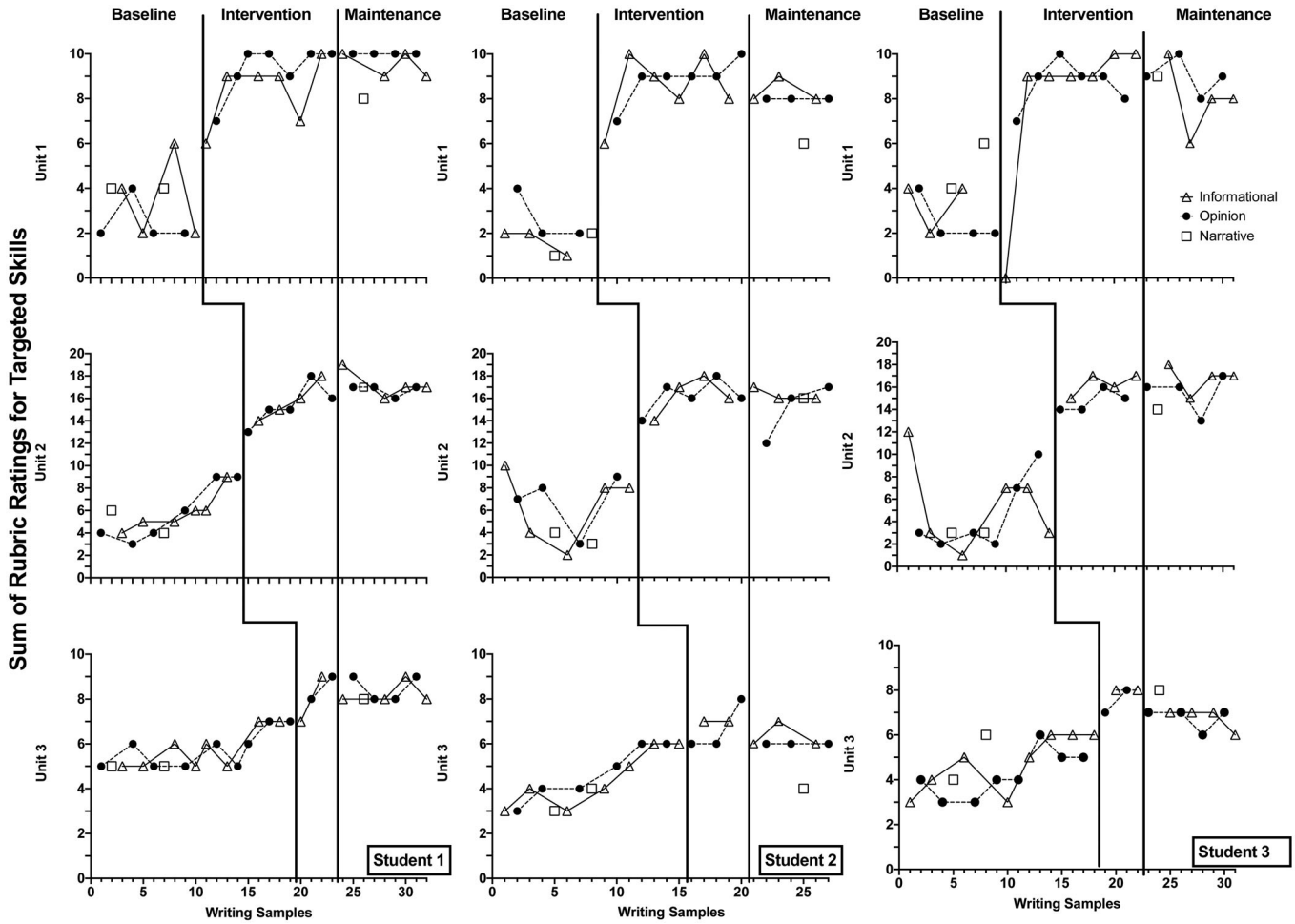

Figure 1. Results from the participants in group 1. Unit 1 is organization and topic maintenance; Unit 2 is sentence structure, letter case, punctuation, and handwriting; and Unit 3 is spelling and vocabulary. A student could receive a score of up to 5 points for each skill.

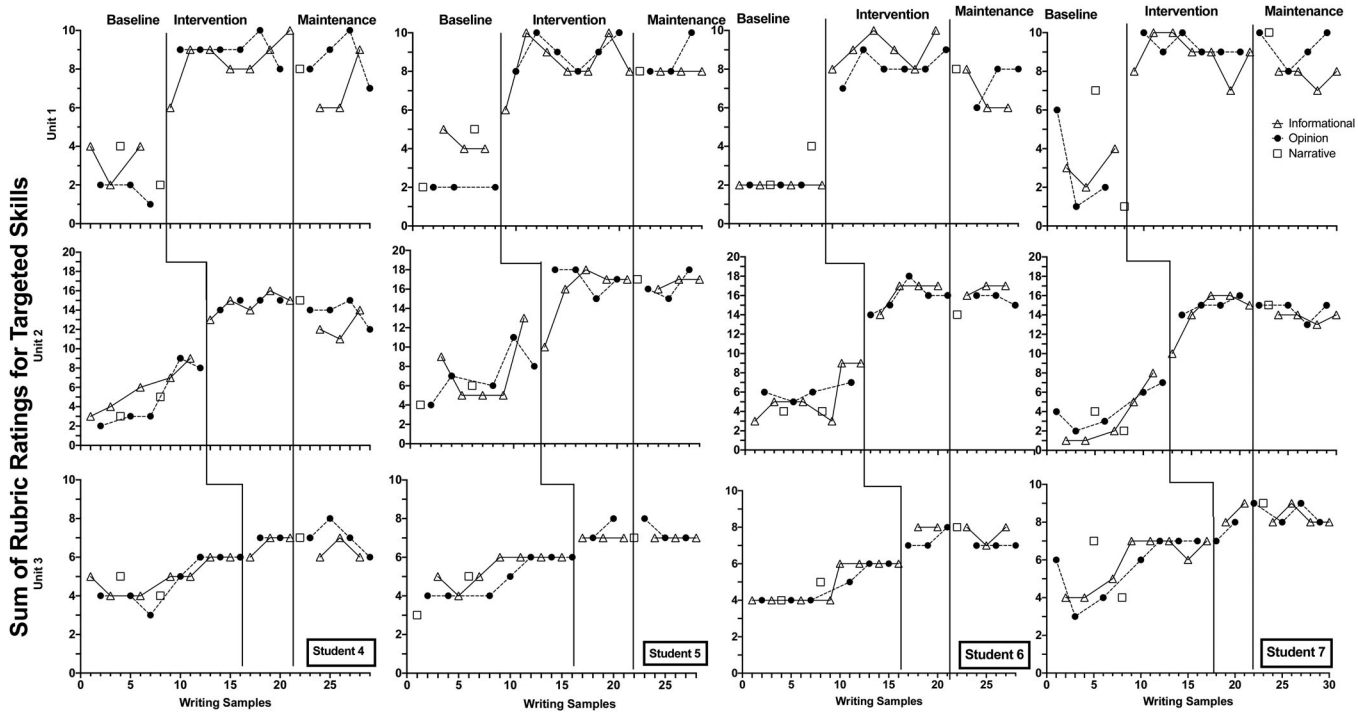

Figure 2. Results from the participants in group 2. Unit 1 is organization and topic maintenance; Unit 2 is sentence structure, letter case, punctuation, and handwriting; and Unit 3 is spelling and vocabulary. A student could receive a score of up to 5 points for each skill.

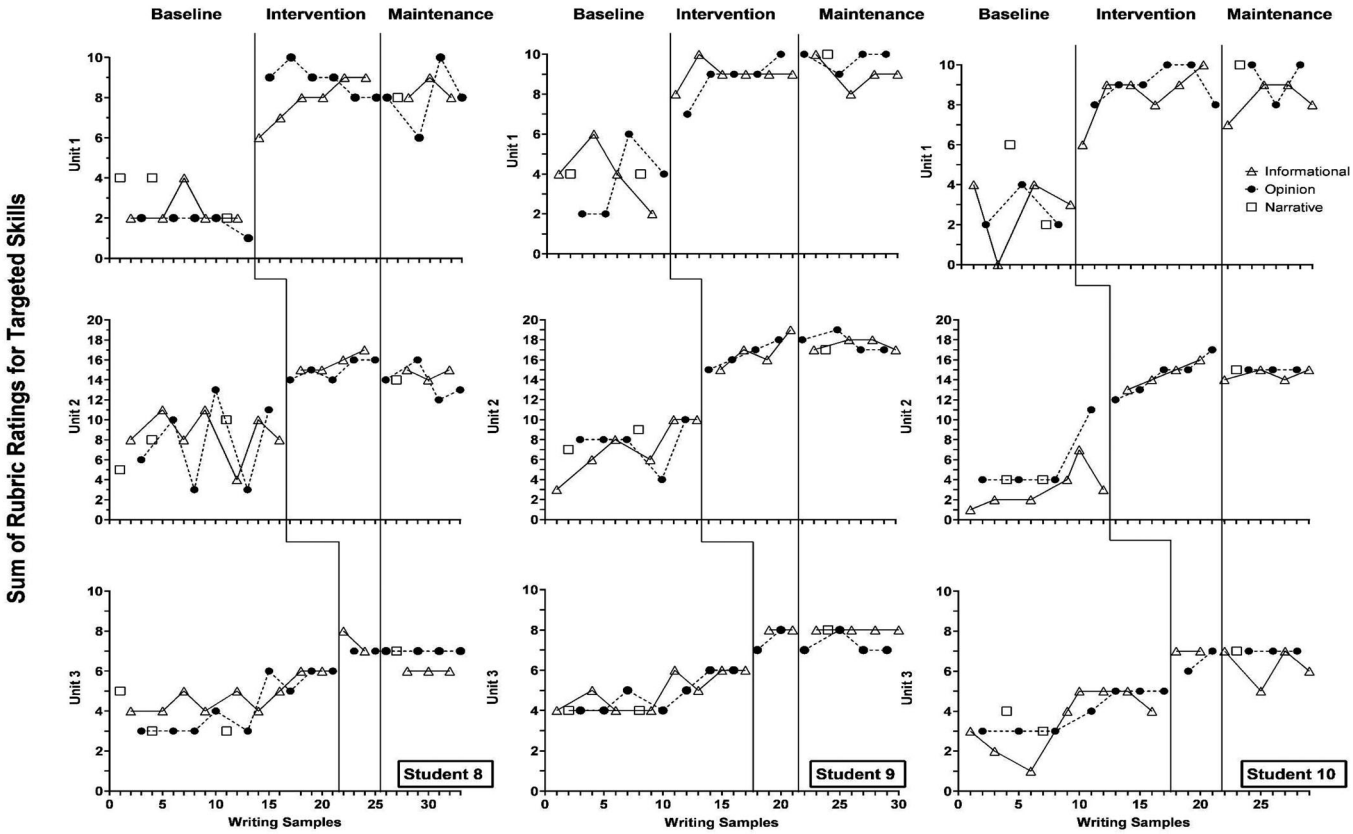

Figure 3. Results from the participants in group 3. Unit 1 is organization and topic maintenance; Unit 2 is sentence structure, letter case, punctuation, and handwriting; and Unit 3 is spelling and vocabulary. A student could receive a score of up to 5 points for each skill.

Table 3. Nonoverlap statistics.

Student	Unit	Tau-U
1	1	0.99
1	2	1
1	3	0.95
2	1	1
2	2	1
2	3	0.89
3	1	0.84
3	2	1
3	3	1
4	1	1
4	2	1
4	3	0.94
5	2	0.96
5	3	1
6	1	1
6	2	1
6	3	1
7	1	0.99
7	2	1
7	3	0.88
8	1	1
8	2	1
8	3	1
9	1	1
9	2	1
9	3	1
10	1	0.99
10	2	1
10	3	1

point that regressed to the baseline range. Students 1, 3, 5, 6, 7, 9, and 10 maintained gains made during Unit 3, but the data points were not always at the same level as they were during the intervention phase. Student 10 had one data point that regressed to the baseline range and Student 3 had two data points that regressed to the baseline range. Students 2, 4, and 8 had three or more maintenance data points that regressed back to their baseline range and were unable to maintain their gains in the Unit 3 skills.

Tau-U, a nonparametric effect size estimate, was calculated, comparing baseline with treatment conditions across the three units of behavior. The benchmarks to interpret the effect sizes calculated using Tau-U are very effective (0.93 or greater), effective (0.66–0.92), and questionable or ineffective (0.65 or less) (Parker et al., 2011). All Tau-U scores fell within the very effective and effective ranges. The Tau-U coefficient shows that the scores for all three units of the intervention increased from the baseline for all students. Students 6, 8, and 9 had a Tau-U of 1.0 for all three units. Nine out of the 10 students had a Tau-U of 1.0 in Unit 2. Six of the 10 students had a Tau-U of 1.0 in Unit 3. Detailed results can be found in Table 3.

Statistical analysis of the multiple baseline design

A two-level model (observations nested within students) was applied to estimate the overall average treatment effect across all 10 students. Because the number of students was relatively small for a multilevel modeling application, we used restricted maximum likelihood estimation with Kenward-Roger adjusted degrees of freedom and standard errors (Kenward & Roger, 1997) as implemented in the mixed procedure of SAS software, version 9.4 of the SAS

Table 4. KTEA-3 assessment performance pre- and postintervention.

Student	Pre-KTEA-3 WE SS	Post-KTEA-3 WE SS	Pre-KTEA-3 Spelling SS	Post-KTEA-3 Spelling SS
1	102	125	107	116
2	82	95	94	94
3	95	121	107	105
4	77	96	91	89
5	90	113	102	99
6	86	109	104	108
7	80	120	107	112
8	82	117	96	102
9	92	116	100	103
10	94	110	98	88
M	88	112	101	102
SD	7.90	10.08	5.78	9.30

Note. KTEA-3 = Kaufman Test of Educational Achievement Third Edition; SS = standard score; WE = written expression.

System for Windows. These small sample adjustments have been shown to produce accurate treatment effect inferences in multiple baseline data with as few as four participants (Ferron et al., 2009). The researchers included a treatment indicator variable to model differences between baseline and treatment phases, and included instructional unit indicator variables to account for variation between the three instructional units. To make the outcome scale comparable across instructional units, the rubric scores from Unit 2 were divided by 2, so that all units had a maximum score of 10. Because the results of the visual analysis indicated trends for some cases, the researchers estimated the treatment effect with a model that adjusted for baseline and treatment phase trends. The researchers centered the time variable such that 0 corresponded with the first treatment observation and thus the initial treatment effect was estimated.

The model for level-1 is:

$$Y_{ij} = \beta_{0j} + \beta_{1j} * Treatment + \beta_{2j} * Time + \beta_{3j} * Time * Treatment + \beta_{4j} * Unit1 + \beta_{5j} * Unit2 + e_{ij} \quad (1)$$

where Y_{ij} is sum of the rubric ratings for the targeted skills for the j^{th} person at time *i*. *Treatment* is coded 0 for baseline phase observations and 1 for treatment phase observations, Unit1 is coded 1 for observations in Instructional Unit 1 and 0 for all other observations, Unit 2 is coded 1 for observations in Instructional Unit 2 and 0 for all other observations, and *Time* is centered such that 0 corresponds to the first observation in the treatment phase. As a consequence β_{0j} is the expected baseline value projected one observation into treatment for student *j* for the reference instructional unit (Unit 3), β_{1j} is the initial treatment effect for student *j*, β_{2j} is the baseline slope for student *j*, and β_{3j} is the change in slope between the baseline and treatment phases for student *j*, β_{4j} is the effect of Instructional Unit 1 relative to Instructional Unit 3, and β_{5j} is the effect of Instructional Unit 2 relative to Instructional Unit 3. The level-2 model is:

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (1.1)$$

$$\beta_{1j} = \gamma_{10} + u_{1j} \quad (1.2)$$

$$\beta_{2j} = \gamma_{20} + u_{2j} \quad (1.3)$$

$$\beta_{3j} = \gamma_{30} + u_{3j} \quad (1.4)$$

$$\beta_{4j} = \gamma_{40} \quad (1.5)$$

$$\beta_{5j} = \gamma_{50} \quad (1.6)$$

where γ_{00}, γ_{10}, γ_{20}, γ_{30}, γ_{40}, and γ_{50} represent the across person average values of β_{0j}, β_{1j}, β_{2j}, β_{3j}, β_{4j} and β_{5j}, and u_{0j}, u_{1j}, u_{2j}, and u_{3j} are the person-level errors, assumed to be distributed multivariate normal.

The expected baseline value projected one observation into treatment is 5.30 ($SE = 0.22$, $p < .0001$) for Instructional Unit 3, and the average initial treatment effect is 3.09 ($SE = 0.21$, $p < .0001$). Using the ideal goal for the outcome of 10 points, this change of 3.09 from the baseline level of 5.30 corresponds to a percent of goal obtained (PoGO) equal to 66%, and thus after adjustment for trend, the initial effect estimate falls in the moderately large range (Ferron et al., 2020). The estimation of the baseline slope is 0.09 ($SE = 0.02$, $p < .0001$), and the estimation of the change in slope between the baseline and treatment phases is 0.08 ($SE = 0.04$, $p < .05$). Because of this increase in slope, the effect estimate increases with time in intervention. For example, four observations into treatment (i.e., centered Time = 3), the baseline is projected to be 5.57 (= 5.30 + 0.09 * 3) for Unit 3 and the treatment effect is estimated to be 3.33 (= 3.09 + 0.08 * 3), which corresponds to a PoGO of 75% at this point in time. There was some variation in the average scores between instructional units, where the scores for Unit 1 were about a point lower than the scores for Unit 3 ($\gamma_{40} = -1.06$, $SE = 0.21$, $p < .0001$), and the scores for Unit 2 were 1.6 points lower than Unit 3 ($\gamma_{50} = -1.60$, $SE = 0.19$, $p < .0001$). The between student variance components were small and not statistically significant, indicating there was little to no reliable variability across students in the baseline levels, immediate treatment effects, or trends within phases. The autoregressive coefficient was estimated to be 0.43 ($SE = 0.04$, $p < .0001$) and the within student residual variance was estimated to be 1.71 ($SE = 0.12$, $p < .0001$).

Standardized assessment performance before and after intervention

To address the second research question, the pre- and post-assessment data for all of the students can be found in Tables 4 and 5. Wilcoxon signed rank tests (Wilcoxon, 1945) were conducted to compare the two sets of scores from each of the students on the KTEA-3 and i-Ready subtests.

Table 5. i-Ready assessment performance pre- and postintervention.

Student	i-Ready Reading Fall Overall Score	i-Ready Reading Winter Overall Score	i-Ready Phonics Fall Score	i-Ready phonics Winter Score	i-Ready HFW Fall Score	i-Ready HFW Winter Score	i-Ready Vocab Fall Score	i-Ready Vocab Winter Score
1	407	463[b]	409	462	411	422	407	480
2	392	434[a]	405	481	386	494	389	409
3	413	462[b]	406	458	375	428	412	427
4	410	406[a]	348	391	374	409	401	424
5	410	464[b]	375	423	441	444	393	417
6	428	448[a]	451	452	464	444	375	447
7	419	459[b]	423	452	522	522	416	509
8	360	441[a]	374	441	434	444	355	401
9	415	471[b]	437	503	428	464	406	447
10	403	417[a]	334	425	408	415	395	420
M	406	447	396	449	424	449	395	438
SD	18.67	21.79	37.80	31.43	45.19	35.87	18.47	33.76

Note. HFW = high-frequency words.
[a]Below-grade-level performance on the i-Ready Reading Assessment.
[b]On-grade-level performance on the i-Ready Reading Assessment.

The scores on the KTEA-3 written expression subtest for all 10 of the participants increased, with improvements ranging from 13 to 40 points. The Wilcoxon signed rank test indicated a statistically significant change ($Z = -2.180$, $p = .005$). Standard scores on the KTEA-3 Spelling subtest did not improve consistently. Five students increased by 3–9 points. One student's score remained the same and scores for four students decreased. The Wilcoxon signed rank test indicated a statistically significant change was not present ($Z = -0.831$, $p = .406$).

The i-Ready Reading overall scores for the winter assessment period along with the phonics, high-frequency words, and vocabulary subtest scores from the fall and winter assessment periods can be found in Table 5. Although all 10 of the students were performing below grade level on the fall assessment according to their overall i-Ready score, 5 of the 10 students were performing on grade level according to the overall score obtained on the Winter i-Ready Reading Assessment. The overall i-Ready Reading score increased for 9 of the 10 students. The Wilcoxon signed rank test indicated a statistically significant change ($Z = -2.703$, $p = .007$). All 10 of the students made gains in their i-Ready vocabulary subtest scores and the results of the Wilcoxon signed rank test indicated a statistically significant change ($Z = -2.803$, $p = .005$). All 10 of the students made gains in their i-Ready phonics subtest scores and the results of the Wilcoxon signed rank test indicated a statistically significant change ($Z = -2.803$, $p = .005$). Eight of the 10 students made gains in their high frequency words subtest scores, one student's score remained the same, and one student's score decreased. The results of the Wilcoxon signed rank test indicated a statistically significant change ($Z = -2.073$, $p = .038$).

Instructional staff ratings

Two of the first-grade teachers and the writing coach rated writing samples collected during baseline, intervention, and maintenance. The raters were unaware of the dates when the samples were collected. They used the following scale: 1 = poor, 2 = below average, 3 = average, 4 = above average. The average ratings for all skills were higher in the intervention and maintenance conditions than the baseline condition. The only exception was the handwriting scores for Student 2, in which an average score of 3 was obtained in the baseline, intervention, and maintenance conditions. Even though the spelling scores for Students 2, 4, 5, 6, 8, and 10 improved when compared with baseline, the scores received on their intervention and maintenance writing samples fell within the poor or below-average range from at least one of the raters. The average ratings for each student across nine areas measured are presented in Table 6.

Social validity

Teacher survey

The three first-grade teachers completed social validity surveys containing questions related to the writing intervention, if the improvements on skills targeted during the intervention were observed during classroom writing activities, and teachers' interest in participating in writing intervention research the following school year. A table containing the average responses for each item can be found in Table 7. The highest possible score was obtained for 9 of the 10 items, indicating that teachers thought that students successfully generalized the skills taught in the small group to their classroom and that they found this intervention to be beneficial. On the short answer questions, all teachers noted the benefits to students receiving small group writing intervention. One teacher stated that the opportunity for more students to receive the intervention would be advantageous. Another teacher indicated that she would like to work with the interventionist on writing in her classroom.

Student survey

All 10 students completed interviews. Table 8 contains their responses to the five yes-and-no questions. All 10 participants indicated that they liked practicing their writing in the small group and that the extra practice made it easier to complete writing activities in class. Nine of the 10 participants indicated that writing lessons were easy to understand; one student said that the lessons were difficult to understand. All of the participants agreed that the pictures helped

Table 6. Instructional staff ratings.

	Average of Baseline Ratings				Average of Intervention Ratings				Average of Maintenance Ratings			
	Poor	BA	AVG	AAVG	Poor	BA	AVG	AAVG	Poor	BA	AVG	AAVG
Overall writing sample	5	5	0	0	0	1	9	0	0	1	8	1
On topic	1	8	1	0	0	0	9	1	0	0	9	1
Well organized	6	4	0	0	0	2	7	1	0	0	9	1
Handwriting	2	5	3	0	0	2	8	0	0	1	9	0
Spelling	5	4	1	0	0	4	6	0	1	5	4	0
Sentence structure/Grammar	5	5	0	0	0	0	10	0	0	1	9	0
Vocabulary	5	4	1	0	0	2	7	1	0	1	9	0
Capitalization	5	5	0	0	0	0	10	0	0	1	9	0
Punctuation	4	5	1	0	0	0	10	0	0	1	8	1

Note. The numbers in this table correspond to the interpretation of the average of all ratings received. AVG = average; AAVG = above average; BA = below average.

Table 7. Teacher social validity data.

Question	Response average
1. I was excited that students in my classroom had the opportunity to participate in the small group writing intervention during the 2018–2019 school year.	5
2. The amount of time required for the students receiving the intervention was reasonable.	5
3. I noticed improvements in classroom writing performance of the students receiving the intervention.	5
4. The students' improved in producing well organized and on topic writing pieces.	5
5. The neatness and readability of the students' writing increased.	5
6. The students' use of complete and grammatically correct sentences while writing improved.	5
7. The students' use of appropriate capitalization and punctuation increased.	4.67
8. The students' use of a variety of vocabulary when writing increased.	5
9. I would be interested in having students in my classroom participate in a writing intervention next school year.	5
10. I would be interested in implementing a small group writing intervention through the use of the lessons and support as needed.	5

Note. 1 = strongly disagree, 2 = disagree, 3= somewhat agree, 4 = agree, 5 = strongly agree.

Table 8. Student social validity data.

Question	Yes	No
1. Did you like practicing your writing in the small group?	10	0
2. Did the extra writing practice make it easier to complete writing activities in the classroom?	10	0
3. Were the writing lessons in the small group easy to understand?	9	1
4. Were the writing lessons in the small group hard to understand?	1	9
5. Did the pictures help you with remembering each of the skills you learned about?	10	0

with remembering each of the skills they learned about. Four of the five students who felt sad, nervous, and frustrated about writing prior to beginning the intervention indicated that after they intervention they felt comfortable, awesome, and good during writing activities. Several students reported that they enjoyed working together with only a few students in a small group on their writing and that they appreciated that there were no distractions. Their favorite activities included the stop light and colors associated with different parts of a writing piece and the activities with dry erase markers (e.g., phoneme grapheme mapping). Many of the students reported that the only thing that they did not like was when they had to complete independent writing activities. They said it was challenging when they had to work on their own and sit by themselves in silence. The students reported that writing in a small group was different than their classroom because the activities felt like warm-ups and they had opportunities to work as a team and learn new things together. During class they felt that the teacher would just tell them what to do instead of working together with them. They said that they would often draw pictures to plan for their writing instead of using the graphic organizer to plan. They reported that the checklist used in small group to talk about their writing was different than what occurred in their classroom.

Discussion

The purpose of this preliminary investigation was to evaluate the effects of a writing intervention using a SRSD model to target informational and opinion writing of first-grade students. In summary, the results indicated that first-grade students experiencing difficulty with writing made significant improvements. Treatment effects were evident from visual analysis, nonoverlap statistics, and multilevel modeling. All 10 students made gains in targeted behaviors as they were introduced to intervention, with effects maintained through the remainder of the school year. Although the data collected on reading and writing measures also indicated growth, our design did not allow us to attribute that to the

intervention. The ratings of instructional staff who were blind to conditions indicated that improvements in writing were readily perceptible. Teachers and students considered the intervention to be acceptable.

The results of this study are consistent Troia and Graham with the findings of previous SRSD writing studies (e.g., Lane et al., 2011; Saddler, 2006; Zumbrunn & Bruning, 2013). Lane et al. (2011) found that second grade students with writing and behavior concerns receiving SRSD instruction made significantly greater gains in the quality of their narrative and opinion writing as compared with the students in the control group. Saddler (2006) found that second-grade students with learning disabilities made improvements in the completeness, length, and quality of the stories they wrote after learning a strategy for planning. Zumbrunn and Bruning (2013) found that the narrative writing skills of first-grade students performing on grade level improved after they received SRSD instruction. The present study demonstrates that SRSD instruction adapted for first-grade students performing below grade level can produce meaningful improvements in informational and opinion writing when strategies are introduced to address the skills targeted in the unit of instruction. Consistent with the findings of Lane et al. (2011), Saddler (2006), and Zumbrunn and Bruning (2013), the quality and length of the writing produced by every student improved. SRSD also increased students' independent use of strategies (e.g., planning for writing using a graphic organizer to include a topic sentence, ideas or facts, and a closing sentence; revising an incomplete sentence by adding words, capitalization and punctuation, etc.). Although narrative writing was not the genre of instructional focus, the scores on the two narrative writing samples collected at baseline were similar to the scores obtained on the informational and opinion genres. There was only one follow-up narrative sample collected after the intervention, and the data from all of the participants showed their scores to be consistent with the scores obtained on the informational and opinion writing samples collected in maintenance. More frequent collection of narrative writing samples should have occurred to better evaluate generalization across genres.

All students received the intervention in the same order, which is contrary to the ideal multiple baseline design. The strongest effects in the present study were present in Unit 1. Unit 1 focused on foundational skills. Not surprisingly, some carryover effects on Units 2 and 3 resulted in higher baseline levels and sometimes upward baseline trends. This indicates that as writing was organized and on topic, students' production of grammatically correct sentences and use of vocabulary related to the writing prompt increased. However, after stabilization of Unit 2 baselines, albeit at higher levels, substantial effects of Unit 2 intervention were evident. All of the students experienced the smallest effects in Unit 3. There was less room for improvement, as baseline scores were at six or above for all but one of the students when Unit 3 instruction began. It also is possible that students may have been experiencing cognitive, linguistic, and physical overload in tracking so many skills (Troia & Graham, 2003; Torrance & Galbraith, 2006). Flower and Hayes (1980) described writing as a dynamic process in which students must fulfill concurrent requirements each time they produce a writing piece. It also is worth noting that six of the students who had the lowest effect scores in Unit 3 struggled with spelling and might benefit from an explicit phonics or spelling intervention. Teachers reported that the students were not currently receiving spelling or phonics instruction in the classroom. When teachers completed the blind ratings of the writing skills of these six students, their spelling scores were in the poor and below average ranges during intervention and maintenance.

Initially, the treatment effects of two models were estimated when statistically analyzing the multiple baseline design. One model assumed no trends and one model assumed linear trends in baseline and treatment phases because the results of the visual analysis indicated some ambiguity about the presence of trends. Although the average baseline slope in the model with trends is small, it is statistically significant, thus the findings from the model with trends are reported. The positive effect for the change in slope between baseline and treatment indicates that the effect gets larger with more time in treatment. Despite a large effect, students did not fully master all the writing skills targeted. Nevertheless, the overall PoGO indicates a clinically significant improvement. There was little to no variability across students in baseline and treatment conditions in either of the models. The multilevel model results can be generalized to other students when this research is replicated.

All students maintained the skills learned in Units 1 and 2 after the intervention was over. However, three students did not maintain their gains in Unit 3. All students were successful without access to the materials (e.g., word wall) that had been available when independent writing samples were collected during the intervention phase (which they often noted). Students found innovative ways to recreate materials (e.g., using an extra pencil as a spacing stick).

The ratings of students' samples by instructional staff who were unaware of the condition were consistent with study findings. Thus, even a more global rating indicated that the improvements in students' writing were readily perceptible to relevant stakeholders. Teachers also noted improved writing in the classroom. Notably, four of the five students with negative feelings toward writing before intervention reported positive feelings after intervention.

While enrolled in this study, the students received Tier 1 writing instruction in their classroom. There were no formal observations of what this instruction looked like. However, during the social validity survey, teachers reported that they often taught a mini lesson for approximately 15 min to the whole group and then provided students with independent writing time. Even though writing instruction was occurring in the classroom, effects were not clearly evident. If effective, the researchers might have seen improvements in baseline writing performance. With the initiation of intervention, especially for the first two units, immediate and rather large improvements in writing were demonstrated by all students.

Although the researchers do not know which components of this intervention are most powerful, the student responses on the social validity survey regarding the visual and tactile cues were especially positive. Visuals are powerful tools for instruction that provide concrete and memorable prompts (Barton & Sawyer, 2003). Cues were generated to represent the genres of writing instruction, to symbolize skills addressed in units of instruction, and to assist students as they learned new strategies and participated in activities. The tactile cues (e.g., spacing stick) allowed students to have a tool to assist them with a skill as they worked toward independence.

Limitations and future research

Because the study applies a single subject experimental design, the researchers are only able to generalize these findings to students with characteristics similar to the students in this study. The effects of future iterations of this intervention with more diverse samples will allow further investigations of efficacy and broaden claims of generalizability. Future researchers could rule out a possible order effect by counterbalancing the order of units. Experiments in which the units are reorganized to address different skills should be conducted. The lessons addressing spelling and vocabulary could be strengthened. Spelling lessons should be aligned to the phonics instruction in the classroom and incorporating texts may allow for increasing background knowledge prior to writing and use of vocabulary from the text in the writing.

Further research is necessary to fully develop a tiered approach to effective writing instruction in early elementary grades (Saddler & Asaro-Saddler, 2013). During this study, the intervention was implemented as supplemental, small group, Tier 2 instruction. Additional work is needed to better understand the instruction occurring during general classroom instruction (Tier 1). Presently, inconsistencies in writing instruction across classrooms and teachers are present, and writing is taught for less than 30 min each day in Grade 1 (Coker et al., 2016). The limited writing instruction occurring in classrooms may exacerbate the need for supplemental (Tier 2) writing instruction (Johnson et al., 2013). However, interventions need to be identified and validated for all tiers (Saddler & Asaro-Saddler, 2013) to provide educators with resources for implementation.

In this preliminary study, Meaghan McKenna conducted the intervention sessions and collected the writing samples. Teachers need to implement the intervention in the future to determine if these results are replicable. Implementation of this intervention by educators will require training and ongoing job embedded professional development. A clear scope and sequence that allows for differentiation based on the areas of mastery of students is necessary (Reigeluth, 1999). Explanations of when and how to use the materials and provide feedback also should be included (Scheeler et al., 2004).

Educators also will need to be able to score writing samples with a valid and reliable measure to assess student progress (Lembke et al., 2010). In this study a researcher-developed writing scoring rubric was used to measure students' responses to intervention. During baseline, students experienced difficulty producing writing that showed evidence of the skills assessed and the majority of rubric scores fell within the *never*, *rarely*, and *sometimes* ranges. After receiving the intervention, scores fell within the *sometimes*, *often*, and *almost always* ranges. However, this rubric only had content validity. Measurement studies need to be conducted to gather multiple sources of validity and reliability evidence.

Conclusion

This study highlights the positive response of first-grade students to an intervention targeting informational and opinion writing. These preliminary findings are promising and may be improved through iterative development of this intervention. Providing our youngest learners with a strong writing foundation will prepare them for success in school and life.

Acknowledgments

The authors are extremely grateful to the principal, assistant principal, writing coach, and first-grade teachers who supported this study. The leadership team members truly believe that success is the only option and they are champions of research to improve educational practice. Thank you all so much! This project would have never been possible without your partnership.

Disclosure statement

There are no relevant conflicts of interest.

Funding

Grant R324B180004 from the Institute of Education Sciences, U.S. Department of Education, awarded to the University of Kansas supported the time spent making minor revisions to this article. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

ORCID

Meaghan McKenna http://orcid.org/0000-0001-9430-415X
Howard Goldstein http://orcid.org/0000-0003-3270-2739
Xigrid Soto-Boykin http://orcid.org/0000-0002-5135-9078
Ke Cheng http://orcid.org/0000-0001-8429-0567
Gary A. Troia http://orcid.org/0000-0002-4827-5674
John Ferron http://orcid.org/0000-0002-6263-5912

References

Applebee, A., & Langer, J. (2011). The national study of writing instruction: Methods and procedures. *Center on English Learning & Achievement*, 1–122. Retrieved from https://www.yumpu.com/en/document/read/7740163/the-national-study-of-writing-instruction university-at-albany.

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis 1. *Journal of Applied Behavior Analysis*, *1*(1), 91–97.

Baker, S., Gersten, R., & Graham, S. (2003). Teaching expressive writing to students with learning disabilities: Research-based applications and examples. *Journal of Learning Disabilities*, *36*(2), 109–123. https://doi.org/10.1177/002221940303600204

Barton, J., & Sawyer, D. M. (2003). Our students are ready for this: Comprehension instruction in the elementary school. *The Reading Teacher*, *57*(4), 334–347.

Berninger, V. W., Garcia, N. P., & Abbott, R. D. (2009). Multiple processes that matter in writing instruction and assessment. In G. A. Troia (Ed.), *Challenges in language and literacy. Instruction and assessment for struggling writers: Evidence-based practices* (pp. 15–50). Guilford Press.

Berninger, V. W., Vaughan, K., Abbott, R. D., Brooks, A., Abbott, S. P., Rogan, L., Reed, E., & Graham, S. (1998). Early intervention for spelling problems: Teaching functional spelling units of varying size with a multiple-connections framework. *Journal of Educational Psychology*, *90*(4), 587–605. https://doi.org/10.1037/0022-0663.90.4.587

Brookhart, S. M. (1999). *The art and science of classroom assessment: The missing part of pedagogy*. ASHE-ERIC Higher Education Report, Volume 27, Number 1. ERIC Clearinghouse on Higher Education.

Chapman, M. (2002). A longitudinal case study of curriculum genres, K-3. *Canadian Journal of Education / Revue Canadienne de L'éducation*, *27*(1), 21–44. https://doi.org/10.2307/1602186

Chapman, M., Filipenko, M., McTavish, M., & Shapiro, J. (2007). First graders' preferences for narrative and/or information books and perceptions of other boys' and girls' book preferences. *Canadian Journal of Education / Revue Canadienne de L'éducation*, *30*(2), 531–553. https://doi.org/10.2307/20466649

Coker, D. L., Farley-Ripple, E., Jackson, A. F., Wen, H., MacArthur, C. A., & Jennings, A. S. (2016). Writing instruction in first grade: An observational study. *Reading and Writing*, *29*(5), 793–832. https://doi.org/10.1007/s11145-015-9596-6

Coker, D. L., & Ritchey, K. D. (2010). Curriculum-based measurement of writing. In *Kindergarten and first grade: An investigation of production and qualitative score*.

Common Core State Standards: National Governors Association and Council of Chief School Officers. (2010). Common Core State Standards Initiative Preparing America's Students For College and Career Readiness. http://www.corestandards.org/

Donovan, C. A. (2001). Children's development and control of written story and informational genres: "Insights from one elementary school". *Research in the Teaching of English*, *35*, 394–447.

Donovan, C. A., & Smolkin, L. B. (2002). Considering genre, content, and visual features in the selection of trade books for science instruction. *The Reading Teacher*, *55*(6), 502–520.

Donovan, C. A., & Smolkin, L. B. (2011). Supporting informational writing in the elementary grades. *The Reading Teacher*, *64*(6), 406–416. https://doi.org/10.1598/RT.64.6.2

Duke, N. K. (2000). 3.6 minutes per day: The scarcity of informational texts in first grade. *Reading Research Quarterly*, *35*(2), 202–224. https://doi.org/10.1598/RRQ.35.2.1

Ferron, J. M., Bell, B. A., Hess, M. R., Rendina-Gobioff, G., & Hibbard, S. T. (2009). Making treatment effect inferences from multiple-baseline data: The utility of multilevel modeling approaches. *Behavior Research Methods*, *41*, 372–384.

Ferron, J., Goldstein, H., Olszewski, A., & Rohrer, L. (2020). Indexing effects in single-case experimental designs by estimating the percent of goal obtained. *Evidence-Based Communication Assessment and Intervention*, *14*(1–2), 6–27. https://doi.org/10.1080/17489539.2020.1732024

Flower, L., & Hayes, J. R. (1980). The cognition of discovery: Defining a rhetorical problem. *College Composition and Communication*, *31*(1), 21–32. https://doi.org/10.2307/356630

Gast, D. L., & Ledford, J. R. (2014). *Single case research methodology: Applications in special education and behavioral sciences* (2nd ed.). Routledge.

Gast, D. L., Lloyd, B. P., & Ledford, J. R. (2014). Multiple baseline and multiple probe designs. In D. L. Gast, (Ed.), *Single Case research methodology* (pp. 251–296). Routledge.

Gast, D. L., & Spriggs, A. D. (2014). Visual analysis of graphic data. In D. L. Gast, & J. R. Ledford, (Eds.), *Single case research methodology* (pp. 176–210). Routledge.

Gee, J. (2001). Reading, language abilities, and semiotic resources: Beyond limited perspectives on reading. In J. Larson (Ed.), *Literacy as snake oil: Beyond the quick fix* (pp. 7–26). Peter Lang.

Gerde, H. K., Bingham, G. E., & Wasik, B. A. (2012). Writing in early childhood classrooms: Guidelines for best practice. *Early Childhood Education Journal*, *40*(6), 351–359. https://doi.org/10.1007/s10643-012-0531-z

Gilbert, J., & Graham, S. (2010). Teaching writing to elementary students in grades 4–6: A national survey. *The Elementary School Journal*, *110*(4), 494–518. https://doi.org/10.1086/651193

Gillespie, A., & Graham, S. (2014). A meta-analysis of writing interventions for students with learning disabilities. *Exceptional Children*, *80*(4), 454–473. https://doi.org/10.1177/0014402914527238

Graham, S., Gillespie, A., & McKeown, D. (2013). Writing: Importance, development, and instruction. *Reading and Writing*, *26*(1), 1–15. https://doi.org/10.1007/s11145-012-9395-2

Graham, S., & Harris, K. R. (1993). Self-regulated strategy development: Helping students with learning problems develop as writers. *The Elementary School Journal*, *94*(2), 169–181. https://doi.org/10.1086/461758

Graham, S., & Harris, K. R. (2013). Common core state standards, writing, and students with SLD: Recommendations. *Learning Disabilities Research & Practice*, *28*(1), 28–37. https://doi.org/10.1111/ldrp.12004

Graham, S., Harris, K. R., & Adkins, M. (2018). The impact of supplemental handwriting and spelling instruction with first grade students who do not acquire transcription skills as rapidly as peers: A randomized control trial. *Reading and Writing*, *31*(6), 1273–1294. https://doi.org/10.1007/s11145-018-9822-0

Graham, S., Harris, K. R., & Loynachan, C. (1993). The basic spelling vocabulary list. *The Journal of Educational Research*, *86*(6), 363–368. https://doi.org/10.1080/00220671.1993.9941230

Graham, S., Hebert, M., & Harris, K. R. (2015). Formative assessment and writing: A meta-analysis. *The Elementary School Journal*, *115*(4), 523–547. https://doi.org/10.1086/681947

Graham, S., McKeown, D., Kiuhara, S., & Harris, K. R. (2012). A meta-analysis of writing instruction for students in the elementary grades. *Journal of Educational Psychology*, *104*, 879–896.

Harris, K. R., & Graham, S. (1985). Improving learning disabled students' composition skills: Self-control strategy training. *Learning Disability Quarterly*, *8*(1), 27–36. https://doi.org/10.2307/1510905

Harris, K. R., Graham, S., & Adkins, M. (2015). Practice-based professional development and self-regulated strategy development for Tier 2, at-risk writers in second grade. *Contemporary Educational Psychology*, *40*, 5–16. https://doi.org/10.1016/j.cedpsych.2014.02.003

Hill, H. C. (2001). Policy is not enough: Language and the interpretation of state standards. *American Educational Research Journal*, *38*(2), 289–318. https://doi.org/10.3102/00028312038002289

Huo, B. (1990). The literature of direct writing assessment: Major concerns and prevailing trends. *Review of Educational Research*, *60*(2), 237–263. https://doi.org/10.3102/00346543060002237

i-Ready. (2019). North Billerica, MA: Curriculum Associates.

Johnson, E. S., Hancock, C., Carter, D. R., & Pool, J. L. (2013). Self-regulated strategy development as a tier 2 writing intervention. *Intervention in School and Clinic*, *48*(4), 218–222. https://doi.org/10.1177/1053451212462880

Juel, C., Griffith, P. L., & Gough, P. B. (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. *Journal of Educational Psychology*, *78*(4), 243–255. https://doi.org/10.1037/0022-0663.78.4.243

Kaufman, A., & Kaufman, N. (2014). *Kaufman test of educational achievement third edition (KTEA-3)*. NCS Pearson.

Kennedy, C. H. (2005). *Single-case designs for educational research*. Prentice Hall.

Kenward, M. G., & Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, *53*(3), 983–997.

Kim, Y. S., Al Otaiba, S., & Wanzek, J. (2015). Kindergarten predictors of third grade writing. *Learning and Individual Differences*, *37*, 27–37. https://doi.org/10.1016/j.lindif.2014.11.009

Krajcik, J., McNeill, K. L., & Reiser, B. J. (2008). Learning-goals-driven design model: Developing curriculum materials that align with national standards and incorporate project-based pedagogy. *Science Education*, *92*(1), 1–32. https://doi.org/10.1002/sce.20240

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Lane, K. L., Harris, K., Graham, S., Driscoll, S., Sandmel, K., Morphy, P., Hebert, M., House, E., & Schatschneider, C. (2011). Self-regulated strategy development at tier 2 for second-grade students with writing and behavioral difficulties: A randomized controlled trial. *Journal of Research on Educational Effectiveness*, *4*(4), 322–353. https://doi.org/10.1080/19345747.2011.558987

Lembke, E. S., Garman, C., Deno, S. L., & Stecker, P. M. (2010). One elementary school's implementation of response to intervention (RTI). *Reading & Writing Quarterly*, *26*(4), 361–373. https://doi.org/10.1080/10573569.2010.500266

Mason, L. H., Harris, K. R., & Graham, S. (2011). Self-regulated strategy development for students with writing difficulties. *Theory into Practice*, *50*(1), 20–27. https://doi.org/10.1080/00405841.2011.534922

Mo, Y., Kopke, R. A., Hawkins, L. K., Troia, G. A., & Olinghouse, N. G. (2014). The neglected "R" in a time of Common Core. *The Reading Teacher*, *67*(6), 445–453. https://doi.org/10.1002/trtr.1227

Moats, L. C. (2005). How spelling supports reading. *American Educator*, *6*(12-22), 42–43.

Montgomery, J. K. (2013). Teaching words to adolescents with language disabilities. *Perspectives on Language Learning and Education*, *20*(2), 67–74. https://doi.org/10.1044/lle20.2.67

National Commission on Writing. (2006). *Writing and school reform*. http://www.collegeboard.com

National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel*. National Institute for Literacy.

Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining nonoverlap and trend for single-case research: Tau-U. *Behavior Therapy*, *42*, 284–299. https://doi.org/10.1016/j.beth.2010.08.006

Puranik, C. S., & Al Otaiba, S. (2012). Examining the contribution of handwriting and spelling to written expression in kindergarten children. *Reading and Writing*, *25*(7), 1523–1546. https://doi.org/10.1007/s11145-011-9331-x

Puranik, C. S., & Lonigan, C. J. (2014). Emergent writing in preschoolers: Preliminary evidence for a theoretical framework. *Reading Research Quarterly*, *49*(4), 453–467. https://doi.org/10.1002/rrq.79

Reigeluth, C. M. (1999). What is instructional-design theory and how is it changing. *Instructional-Design Theories and Models: A New Paradigm of Instructional Theory*, *2*, 5–29.

Rogers, L., & Graham, S. (2008). A meta-analysis of single subject design writing intervention research. *Journal of Educational Psychology*, *100*(4), 879–906. https://doi.org/10.1037/0022-0663.100.4.879

Saddler, B. (2006). Increasing story-writing ability through self-regulated strategy development: Effects on young writers with learning disabilities. *Learning Disability Quarterly*, *29*(4), 291–305.

Saddler, B., & Asaro-Saddler, K. (2013). Response to intervention in writing: A suggested framework for screening, intervention, and progress monitoring. *Reading & Writing Quarterly*, *29*(1), 20–43. https://doi.org/10.1080/10573569.2013.741945

Scheeler, M. C., Ruhl, K. L., & McAfee, J. K. (2004). Providing performance feedback to teachers: A review. *Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children*, *27*(4), 396–407. https://doi.org/10.1177/088840640402700407

Spencer, T. D., & Petersen, D. B. (2018). Bridging oral and written language: An oral narrative language intervention study with writing outcomes. *Language, Speech, and Hearing Services in Schools*, *49*(3), 569–581. https://doi.org/10.1044/2018_LSHSS-17-0030

Taft, R. J., & Mason, L. H. (2011). Examining effects of writing interventions: Highlighting results for students with primary disabilities other than learning disabilities. *Remedial and Special Education*, *32*(5), 359–370. https://doi.org/10.1177/0741932510362242

Torrance, M., & Galbraith, D. (2006). The processing demands of writing. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 67–80). Guilford.

Tower, C. (2003). Genre development and elementary students' informational writing: A review of the literature. *Literacy Research and Instruction*, *42*(4), 14–39.

Troia, G. A., & Graham, S. (2003). The consultants corner: Effective writing instruction across the grades: What every educational consultant should know. *Journal of Educational and Psychological Consultation*, *14*(1), 75–89.

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

Westby, C. (2011). Web support for vocabulary building. *Word of Mouth*, *23*(1), 13–15. https://doi.org/10.1177/10483950114150 23d

Wiig, E. H., Semel, E., & Secord, W. A. (2013). *Clinical evaluation of language fundamentals-fifth edition (CELF-5)*. NCS Pearson.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometries*, *1*, 80–83.

Wolf, B., Abbott, R. D., & Berninger, V. W. (2017). Effective beginning handwriting instruction: Multi-modal, consistent format for 2 years, and linked to spelling and composing. *Reading and Writing*, *30*(2), 299–317. https://doi.org/10.1007/s11145-016-9674-4

Zimmerman, C. B. (2007). *Vocabulary learning methods*. Harvard University.

Zito, J. R., Adkins, M., Gavins, M., Harris, K. R., & Graham, S. (2007). Self-regulated strategy development: Relationship to the social-cognitive perspective and the development of self-regulation. *Reading & Writing Quarterly*, *23*(1), 77–95. https://doi.org/10.1080/10573560600837693

Zumbrunn, S., & Bruning, R. (2013). Improving the writing and knowledge of emergent writers: The effects of self-regulated strategy development. *Reading and Writing*, *26*(1), 91–110. https://doi.org/10.1007/s11145-012-9384-5