IMPROVING KINDERGARTEN STUDENTS' WRITING OUTCOMES USING PEER-ASSISTED STRATEGIES

ABSTRACT

The primary focus of this study was to determine the feasibility of teacher implementation of peer-assisted writing strategies (PAWS) in improving the writing outcomes of kindergarten children. Six classrooms were recruited, and 3 were randomly assigned to the experimental condition. Results indicated that the content, length, and formatting of the lessons were adequate for the teachers to deliver the lessons with fidelity. Students enjoyed PAWS, as reflected in the end-of-the-year surveys. Statistically significant differences between the experimental and control classrooms were noted for punctuation and sentence writing quality. In addition, preliminary results with our small sample size suggest that differences in writing performance between the PAWS and control classrooms were moderated by school type. In the medium-performing schools, differences between pre- and posttest scores were statistically significant for alphabet-writing fluency, punctuation, and sentence and essay curriculum-based writing measures, with effect sizes ranging from 0.69 to 1.96.

Cynthia S. Puranik
GEORGIA STATE UNIVERSITY

Yaacov Petscher
FLORIDA STATE UNIVERSITY

Stephanie Al Otaiba
SOUTHERN METHODIST UNIVERSITY

Christopher J. Lemons
VANDERBILT UNIVERSITY

THE ELEMENTARY SCHOOL JOURNAL
Volume 118, Number 4. Published online April 24, 2018
 0013-5984/2018/11804-0007$10.00

IN the past 3 decades, educational efforts have concentrated on improving reading instruction and intervention for children at risk for reading because educators and policy makers are cognizant of the grim short- and long-term consequences related to poor reading skills. In comparison, work on early writing instruction and intervention is sorely lacking. Although research on how to teach writing is sparse, this dearth is more pronounced in the early years of beginning to write. We were unable to come up with a single empirically based, integrated instructional study on how to teach writing to kindergarten (KG) children (also see Edwards, 2003). The scant research on writing has focused on children in first grade and beyond. Graham, McKeown, Kiuhara, and Harris (2012) conducted an extensive meta-analysis of writing instruction in elementary grades, which included 115 experimental and quasi-experimental studies. Of the 115 studies, only nine were conducted with first-grade students, and no studies included KG students. Although some research on teaching writing to beginning writers exists, it is primarily descriptive or based on individual case studies (e.g., Behymer, 2003; Hall, 1987; Lysaker, Wheat, & Benson, 2010; Sulzby, 1989) or has focused exclusively on teaching spelling primarily to improve decoding skills (Ehri & Wilce, 1987; O'Connor & Jenkins, 1995; Vandervelden & Siegel, 1997).

Not only is there little information on writing curricula that are effective, there is little research on the current state of writing instruction. Research studies including surveys and classroom observation studies concerned with answering questions about writing instruction indicate that very little time is spent teaching writing (Coker et al., 2016; Cutler & Graham, 2008; Graham et al., 2008; Puranik, Al Otaiba, Sidler, & Greulich, 2014) and that there is large variability in teacher practices. For example, surveys of primary grade teachers from first through third grades indicate that although the surveyed teachers reported teaching writing for 70 minutes/week on average, instructional times ranged from 2 to 60 minutes/day (Cutler & Graham, 2008; Graham et al., 2008).

Data obtained from surveys have been confirmed by direct classroom observations (Bingham, Quinn, & Gerde, 2017; Coker et al., 2016; Puranik et al., 2014). Relevant to the current study, Puranik et al. (2014) observed language arts instruction provided by 21 KG teachers across nine schools. Their observations indicated that, on average, teachers spent only 6.1 minutes in the fall and 10.5 minutes in the winter on any kind of writing instruction. Puranik et al. (2014) also noted large variability in the amount of writing instruction occurring in the classroom, the amount of time that KG teachers spent on writing, the amount of time that students spent writing, and student writing performance. Important to note, KG teachers did not report using any specific writing curriculum to teach writing. Most of the time spent on writing instruction was spent on students writing independently rather than on teachers providing instruction.

Despite the importance of teaching handwriting and spelling during the early years of learning to write, Puranik et al. (2014) found that in several classrooms, no systematic handwriting or spelling instruction was observed. Of the 21 teachers in the study, 6 of those observed did not teach handwriting. Although the remaining 15 teachers observed did teach handwriting, they were observed to spend on average less than 1 minute per day on handwriting instruction in the fall semester and less than 2 minutes per day in the winter semester. Experts recommend that

handwriting instruction for beginning writers such as children in KG take place every day and that teachers spend about 75 to 110 minutes per week on handwriting instruction (e.g., Graham & Miller, 1980; Troia & Graham, 2003). Despite the fact that KG is the time when children are learning to form and write letters of the alphabet, students do not appear to be receiving the recommended amounts of explicit and direct instruction needed. The same lack of attention and time devoted to teaching handwriting was true for spelling instruction, as teachers were observed devoting an average of less than 2 minutes to providing spelling instruction.

The large variability in teacher practices across KG and elementary school classes should come as no surprise, as most teachers report that they do not feel adequately prepared or trained to teach writing. For example, only 12% of teachers surveyed indicated that their college coursework adequately prepared them to teach handwriting (Graham et al., 2008). Research examining instructional practices in KG shows that a number of important instructional techniques such as the order in which letters should be taught, writing from memory, and fluency practice were not emphasized in the classrooms (Vander Hart, Fitzpatrick, & Cortesa, 2009). There is a general lack of knowledge regarding the most effective practices to teach writing (Cutler & Graham, 2008), perhaps because we currently do not have any universal or standardized curricula that teachers can follow and "no single writing curriculum that would successfully meet all the students' instructional needs" (Coker & Ritchey, 2015, p. 5).

This lack of research is especially disconcerting in light of the writing requirements on the Common Core State Standards for English Language Arts and Literacy (CCSS), which begin in KG (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). The CCSS for KG includes these items: (*a*) print many uppercase and lowercase letters; (*b*) write a letter or letters for most consonant and short-vowel sounds (phonemes); (*c*) spell simple words phonetically, drawing on knowledge of sound-letter relationships; (*d*) produce and expand complete sentences in shared language activities; and (*e*) use a combination of drawing, dictating, and writing to write about experiences, stories, people, objects, or events. The goal of this study is to address this dearth in writing instruction for beginning writers. We report on a writing intervention—peer-assisted writing strategies (PAWS)—conducted with KG children, thereby beginning writing instruction from the earliest years. We decided to focus on KG for two essential reasons. First, there is a clear lack of current research, as mentioned. Second, it is important to get students off to a good start because the consequences of poor educational experiences in writing are not unlike the consequences of poor educational experiences for reading: children who have a poor start have very little chance of catching up (Graham & Perin, 2007; Juel, 1988; Slavin, Karweit, & Madden, 1989).

Theoretical Foundations of PAWS

According to popular cognitive models of mature writing (Flower & Hayes, 1981; Hayes, 1996; Hayes & Berninger, 2009), the translation phase of writing involves encoding thoughts and ideas into language representations. Expanding this model to children, Berninger and her colleagues (e.g., Berninger, 1999; Berninger & Hooper,

1993) proposed that the translating process in children's writing includes two subcomponents: text generation and transcription. *Text generation* refers to the process by which the writer translates his or her planned ideas into meaningful chunks of sentences, phrases, and words, whereas *transcription* refers to the actual mechanics of converting sentences, phrases, and words into written symbols and includes spelling, handwriting, and punctuation. Thus, the scope and sequence of PAWS focused on teaching and practicing important transcription skills and text-generation skills in line with developmental cognitive models of writing. The specific transcription skills (letter writing, spelling, and punctuation) and text-generation skills (sentence construction and sentence combining) that are targeted are aligned with the KG CCSS (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010).

Transcription skills are a necessary component of writing, according to the developmental models of writing (Berninger et al., 2002; Berninger & Winn, 2006; Juel, Griffith, & Gough, 1986). Handwriting fluency and spelling significantly contribute to both writing quality and productivity (Graham, Berninger, Abbott, Abbott, & Whitaker, 1997; Kim et al., 2011; Puranik, Lombardino, & Altmann, 2008; Wagner et al., 2011). Together, spelling and handwriting fluency accounted for 66% of the variance in compositional productivity and 25% of the variance in students' writing quality in the early elementary grades (Graham et al., 1997). Transcription skills uniquely predict students' writing productivity as early as KG, even after accounting for oral language, reading, and IQ abilities (Puranik & Al Otaiba, 2012). This must explain why addressing transcription skills was the focus of all nine writing intervention studies with first-grade students identified in the meta-analysis conducted by Graham et al. (2012). Of the 115 studies in the meta-analysis, eight studies were identified as focusing on only transcription skills for students in grades 1–3 (Graham et al., 2012). Teaching transcription skills enhanced writing quality across studies for students in elementary grades (average weighted ES = 0.55).

As explained previously, text generation is another important component of the translation process of writing. Beginning or emergent writers start the process of text generation by writing at the letter or word level (Puranik & Lonigan, 2011, 2014). However, skilled or mature writing requires the generation and sequencing of increasingly larger units of written language—from letters in words to words in sentences, to sentences in paragraphs, and finally to paragraphs in written discourse. Based on extensive research, Berninger and colleagues (e.g., Berninger, 2008; Berninger, Mizokawa, Bragg, Cartwright, & Yates, 1994) recommended that even beginning writing instruction should integrate the three levels of language—word, sentence, and text—such that they function together as a system. Thus, a comprehensive writing curriculum should take into account instruction at the word, sentence, and text levels. Furthermore, empirical evidence on writing instruction with third-grade students indicates that teaching spelling and composing simultaneously is more effective in improving writing skills than teaching spelling or composing alone (Berninger et al., 2002).

In addition to training teachers to focus instruction on important transcription and text-generation skills, PAWS was conducted using a peer-assisted framework (peers helping one another to learn, to practice, and to learn themselves by teach-

ing; Topping & Ehly 1998). A peer-assisted framework was chosen because of the strong evidence base supporting the use of peer-assisted learning strategies (PALS; Fuchs & Fuchs, 2005; Fuchs et al., 2001; Fuchs, Fuchs, Mathes, & Simmons, 1997) as a successful teacher-directed, classwide general education practice for improving reading outcomes for students, including KG and primary-grade children. Equally good outcomes for peer-assisted learning have been noted in writing research. In the meta-analysis of writing instruction by Graham et al. (2012), four of the 115 studies used peer assistance with statistically significant effects (weighted ES = 0.89). In another meta-analysis that included 26 studies using peer assistance to teach writing to elementary school students, Hoogeveen and van Gelderen (2013) reported that in all but one of the studies, peer-assisted strategies were beneficial and resulted in improved writing outcomes. It is important to note that in both these meta-analyses, peer-assisted strategies were used much less often with younger writers compared with older writers. Most important, none of the studies in the meta-analyses were conducted with KG children, and none were conducted using a whole-class format.

There are many reasons why peer assistance to teach writing works. First, and most important perhaps, commenting on a peer's work can make students aware of their own writing and help children build metacognitive/metalinguistic skills (Cazden, as cited in Hoogeveen & van Gelderen, 2013). To be able to provide feedback, students need to reflect on the writer's written output and provide feedback orally. These metalinguistic skills help writers eventually monitor and revise their own work. Second, getting feedback from a peer creates a real audience for writing. Beginning writers need ready access to audiences that can respond to their work (Graves, 1983), which, for kindergartners, are their peers and teachers. Third, working with a peer is seen as a teaching-learning opportunity for all students involved; both partners in a pair gain from the process of teaching and learning (Graham & Perin, 2007; Rijlaarsdam et al., 2008).

Context for the Current Study

This project was conducted as part of a 3-year Goal 2 development project funded by the Institute for Education Sciences. The purpose of a Goal 2 project is to demonstrate the feasibility and the promise of an intervention that can be evaluated later through an efficacy study with a more tightly controlled experimental design. In the first phase, we tested the feasibility, usability, fidelity, and dosage in a more controlled condition. Project staff delivered the intervention in a small-group pull-out fashion in five classrooms across two schools. Findings from the first phase indicated that lessons were generally feasible to implement within the time frame (30 minutes, three times/week) and that the components of the intervention functioned as intended. The staff members provided the instruction with a high degree of fidelity. Results of the pilot study in the first phase indicated that there were statistically significant differences between the PAWS participants and control children on the alphabet fluency and essay (correct word sequences) measure from pre- to posttest with large effect sizes of $d = 0.68$ and $d = 0.71$, respectively (Pu-

ranik, Patchan, Lemons, & Al Otaiba, 2017). Although not statistically significant, the PAWS participants also outperformed the control children on spelling and sentence writing measures ($d = 0.53$ and $d = 0.32$, respectively; Puranik, Patchan, Lemons, et al., 2017).

Whereas researcher-conducted intervention studies are a necessary first step, the eventual success of any instructional curriculum depends on whether classroom teachers can successfully implement it in an authentic educational setting. In the current study, we present findings in which we tested the feasibility of implementation with classroom teachers as end users. To summarize, the purpose of the current project was to address some of the gaps in the current research by examining the feasibility of training teachers to teach writing through PAWS and, in turn, the feasibility of PAWS in improving writing outcomes for KG children. Consequently, we had two primary research questions: (*a*) Does the content, length, and formatting of a lesson appear sufficient to enable teachers to conduct the lessons with fidelity and with adherence to the intent of the activities? (*b*) Was there an improvement in important transcription skills (handwriting automaticity and spelling) and text-generation skills (sentence completion and sentence combining) from pre- to posttest for the control versus experimental children? In addition, a third and related question arose in the process of disseminating the results of our first study: Did PAWS have a differential impact within schools that varied in their state ratings of reading performance (medium vs. high performing; explained in Participants and Setting)? This question was of interest because we noted differences in our initial pilot study, in the first phase, in the performance of students depending on the school type; students from high-performing schools showed greater gains compared with children from the low-performing school.

Method

Participants and Setting

The participants were 149 students ranging in age from 61 to 79 months ($M = 67.03$, $SD = 3.88$; 43% female). These participants were 69% White, 25% African American, 3% multiracial, 2% Hispanic, and 1% Asian. Participants were recruited from six KG classes in four schools. All four schools had medium socioeconomic status (i.e., 33%–67% of the students received free or reduced lunch). Two of the schools were medium performing (i.e., met the Adequate Yearly Progress measures as reported by the Pennsylvania Department of Education for 2 of the 3 preceding years), and the other two schools were high performing (i.e., met Adequate Yearly Progress measures for all 3 preceding years). We recruited one teacher each from three schools and three teachers from the fourth school for a total of six teachers. Three classes were then randomly assigned to the experimental condition (PAWS) and the other three served as control classrooms. We ended up with two teachers in the PAWS condition and one teacher serving as a control in the school with three teachers. In the other three schools, one teacher was assigned to one of the two conditions. Demographic information on participants by condition is included in Table 1.

Table 1. Participant Demographic Information by Condition

	PAWS	Controls
n	78	71
Age (months):		
Mean	66.35	67.78
SD	3.50	4.12
Gender (%):		
Male	61.03	52.80
Female	38.96	47.14
Ethnicity (%):		
African American	29.87	18.84
Asian	3.90	2.89
White	63.63	73.91
Other	2.60	4.34

Note.—PAWS = peer-assisted writing strategies.

Each of the three PAWS teachers had a bachelor's degree and between 6 and 10 years of teaching experience. Two control teachers had a bachelor's degree and one had a master's degree. One had between 6 and 10 years of teaching experience, whereas the other two had 10 or more years of experience. All six teachers were female and White. The six teachers used either Houghton Mifflin Harcourt's *Storytown* or Macmillan McGraw-Hill's *Treasures* as their adopted reading curricula. One control teacher reported using the *Peterson Directed Handwriting* program (Peterson, 1908) to teach handwriting; the other five teachers did not use a standard writing curriculum.

PAWS Intervention: Procedures and Design

In the initial feasibility study, writing instruction was provided to two KG classes for 8 or 12 weeks, and the findings demonstrated the feasibility and promise to help children improve their early writing skills (Puranik, Patchan, Lemons, et al., 2017). For the current pilot study, these initial lessons were expanded to create a curriculum that teachers could use in their classrooms throughout the whole school year (i.e., for 26 weeks). Thus, PAWS was delivered three times a week for 30 minutes at each session over 26 weeks. Students received instruction on how to play their roles as coach and writer during the peer interaction and how to provide feedback to their partners when they were coaches. During the lessons, student pairs took turns acting as both writers and coaches (i.e., coaches provided specific feedback to writers that could help improve their writing). To create the pairs, students were first ranked based on their Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Good & Kaminski, 2002) letter-naming fluency (LNF) scores. Then the class was divided in half (two lists), and the students at the top of each list were paired together. This model for pairing children was followed because it worked well in PALS. Finally, the teachers checked to ensure that the pairs of students would likely work well together (i.e., no potential personality conflicts).

All lessons followed the research-based instructional model: *I Do, We Do, You Do* (Fisher & Frey, 2008; Pearson & Gallagher, 1983). The teacher began each lesson

with a warm-up activity that briefly introduced the objective of the lesson. Then the teacher demonstrated the activity in the I Do portion of the lesson. The teacher thought aloud while performing both roles so that the students could hear her thought process. Next, the students completed the activity with their partners while the teacher guided them through each step (We Do). Afterward, the students repeated the activity with their partners at their own pace (You Do). The teacher monitored each pair and provided individual assistance as needed. Finally, the teacher provided closure to the lesson by recapping the lesson's objective. In line with the levels-of-language perspective (Whitaker, Berninger, Johnston, & Swanson, 1994), the PAWS instruction targeted writing at the letter, word, and sentence levels. Thus, the instruction involved three strands that slightly overlapped: letter writing, spelling, and sentence construction.

Letter writing. For the letter-writing strand, the teacher modeled each letter, providing visual and verbal cues for how to form the letter, similar to handwriting instruction used by Berninger et al. (2006) with first graders. First, in the letter-writing strand, students focused on letter formation. During the eight introductory lessons, they were introduced to the role of being a coach. They were taught to identify three types of errors that are commonly found in letter writing (e.g., place, size, and shape errors; see App. A for sample worksheets). Then they practiced writing simple strokes (i.e., arc, vertical line, circle, slant right), exchanging their papers, and providing each other with specific feedback. As part of the letter-writing strand, students completed 13 "Letter Learning" lessons. During each lesson, students were introduced to the name, sound, and formation of two letters. For each worksheet (We Do, You Do), students practiced writing the letter three times. To foster automatic letter production, children first traced and copied the letters before writing freehand. For the first attempt, the writer traced inside a bubble letter. For the second attempt, the writer traced a dotted line, and for the third attempt, the writer practiced writing the letter freehand. After each attempt, the students switched papers, and the coach provided the writer with specific feedback about the placement, size, or shape of the letter. If the attempt had an error, the writer would try again on a fresh line on the paper, which children called "the fix-it line."

To begin connecting the shape of the letter with the letter sound, students also completed five "Letter Sound" lessons. A Letter Sound lesson followed each of the first five Letter Learning lessons. For these lessons, the teacher said a word that started with one of the target letters, and the class repeated that word. Then the writer wrote the letter that he or she thought the word began with. The teacher told the class the correct letter, and if the writer wrote the incorrect letter, he or she would cross out the letter and write the correct letter. Similar to the Letter Learning lessons, the students switched papers, and the coach provided the writer with specific feedback about the placement, size, or shape of the letter. The writer attempted the letter again as needed. Students wrote the letters for four words per worksheet. As the PAWS curriculum progressed, the Letter Sound lessons were replaced with the "Missing Letter" lessons that focused on spelling rather than letter formation (see Spelling for more details).

Finally, after students learned four new letters, they completed a "Letter Learning Review" lesson. During this lesson, the teacher first modeled writing a letter

and guided the students to trace the letter twice. Then, the students practiced writing the letter freehand four more times. Once they completed the four attempts, the students coached themselves by writing a check mark next to the one they thought was best. This process was repeated for the four new letters. After all 26 letters of the alphabet were taught, students periodically completed additional Letter Learning Review lessons that focused on the more difficult letters (e.g., *b*, *d*, *e*, *g*, *p*, *q*, *s*). By the end of the PAWS curriculum, students had completed eight Letter Learning Review lessons.

Spelling. In the spelling strand, students focused on encoding using decodable words and spelling using sight words. Students completed 16 Missing Letters lessons, in which letters were strategically removed from consonant-vowel-consonant words to help students focus on different sounds of the word (i.e., initial, final, middle, and finally all three sounds; see app. A for sample worksheets). For each worksheet, the teacher reviewed the words for all the pictures. Then the writer wrote the letter that made the sound where the blank was located (e.g., *t* for _ub). The students switched their papers, and the coaches checked to see if the letter correctly matched the letter sound. If not, the coach helped the writer determine which letter should be written. If the correct letter was written, the coach checked the letter formation and provided specific feedback regarding the placement, size, or shape. Last, the writer used the fix-it line as needed either to write the correct letter or to attempt to write the letter more neatly. This process was repeated for three words on each worksheet.

As part of the spelling strand, students also completed seven "Sight Word" lessons. The teacher began these lessons by introducing two new sight words per lesson. As a class, students said the first word and spelled it together. Then they visualized it, or in child-friendly terms, "took a picture of it" with their imaginary cameras, before covering the sight word at the top of their worksheets. The writers immediately wrote the word from memory, and afterward the coaches checked to see if it was spelled correctly. Similar to the Missing Letter lessons, if the word was not spelled correctly, the coach helped the writer fix the mistake, and if the word was spelled correctly, the coach checked the handwriting. Again, the fix-it line was used as needed. This process was repeated two additional times. For the second and third attempts, the students were trained to count to 5 and 10 (respectively) before attempting to write the letter from memory. Both decodable and sight words were based on word lists for KG students (Fry, Polk, & Foutoukidis, 1984).

Sentence writing and punctuation. In the sentence construction strand, students focused on constructing complex sentences, communicating a complete idea, and increasing sentence length. Similar to the other two strands, teacher-guided practice was followed by peer-directed independent practice. Students completed 10 "Sentence Starter/Copying" lessons. During these lessons, students learned to provide specific feedback on five aspects of writing sentences: (*a*) Did the sentence make sense and was it a complete thought (i.e., included an article, a subject, and a verb)? (*b*) Were all the words spelled correctly? (*c*) Was there spacing between all the words? (*d*) Was the first word of the sentence capitalized? (*e*) Did the sentence end with punctuation? In the earlier lessons, sentences were provided with the final word missing, and the writers chose and wrote the word that best completed the sentence. Then they practiced copying short sentences. Finally, they were again

given sentences that were missing the final word, and the writers copied the sentence along with the word that best completed the sentence. After the writer completed the sentence, the students switched papers. The coach checked to see if the sentence made sense and whether appropriate spelling, spacing, capitalization, and punctuation were used.

The next set of lessons in the sentence construction strand included five "Sentence Generating" lessons. In these lessons, a picture-word prompt was provided. The pair of students brainstormed ideas based on the prompt and then generated their own sentences. After the writer wrote a sentence, the students switched papers, and the coach provided specific feedback regarding a complete thought, spelling, spacing, capitalization, and punctuation. The final set of lessons in the sentence construction strand included five "Sentence Combining" lessons. In these lessons, the writer followed five steps to combine two sentences into a complex sentence (e.g., "She can tap. She can hop." combined into "She can tap and hop"; see App. A for sample worksheets). After the combined sentence was copied, the students switched papers. The coach checked the sentence to see whether it made sense and had appropriate spelling, spacing, capitalization, and punctuation. Saddler and Graham (2005) found that sentence combining, compared with traditional grammar instruction, resulted in better sentence-writing skills and had a positive impact on fourth graders' writing quality.

Teacher Training and Support: PAWS

To prepare the three teachers to implement the PAWS instruction with fidelity, they received 8 hours of professional development provided by lead project staff. The theoretical frame that guided our professional development (Bransford, Brown, & Cocking, 1999) ensured that the training was learner, knowledge, assessment, and community centered. Just before the school year started, the teachers attended a 5-hour PAWS instruction training course. This course involved three objectives. First, the teachers were introduced to the overall curriculum, which was connected to a brief research-based rationale. Second, the majority of the time was spent understanding the specific lessons. The teachers observed a model of an entire lesson, and they were provided an overview and demonstration for six types of lessons that would occur during the first half of the school year. Finally, the logistics of the study were discussed, including the importance of fidelity and how it would be measured.

At this time, the teachers also met individually with their assigned PAWS liaison, who would be their point of contact for the duration of the study. The PAWS liaisons were two staff members who had ample experience teaching (i.e., Liaison 1 had more than 20 years of experience teaching pre-KG through eighth grade and 4 years as a supervisor of student teachers; Liaison 2 had 4 years of experience teaching pre-KG through first grade), who helped develop the current PAWS curriculum materials (i.e., scripts and worksheets), and who implemented the PAWS instruction in the previous year. For each lesson during the first 2 weeks, the PAWS liaisons were present to observe and offer guidance as needed. Similarly, the PAWS liaisons were also present when new types of lessons were introduced. In addition, PAWS liaisons observed lessons every 2 weeks for fidelity (see Fidelity of Instruction for

more details). For the remaining 6 months, the PAWS liaison met with each teacher for half an hour once a month to discuss their observations and fidelity and to answer any questions (i.e., a total of 3 hours of ongoing professional development). At the end of the instruction period, the teachers were invited to meet with the research team for a focus group meeting to share their thoughts and impressions about the curriculum.

Control Group Instruction

The teachers in the control condition conducted their literacy instruction in their typical fashion. To contextualize the writing instruction taking place in the control classrooms, the teachers in the control condition were also observed six times (days) during the school year. The Classroom Observation Coding System (Coker et al., 2016) was used for these observations. The classroom observation included observation of materials used by the teacher, student groupings (whole class, small group), specific literacy focus (reading, writing), nature of student activity (worksheets, independent work), and so on. In addition to these observations, the project staff interacted with the control teachers during the pre- and posttest assessments. Control teachers were observed to provide reading and writing instruction. Reading instruction included activities to facilitate phonological and phonemic awareness, vocabulary, and reading comprehension. Writing instruction including activities to enhance handwriting and spelling.

Measures

All participants completed a battery of five curriculum-based measures of writing. These assessments took approximately 60 minutes to complete. They were administered to the whole class 1 week before (i.e., pretest) and 1 week after (i.e., posttest) the PAWS instruction by trained research assistants. All curriculum-based measures of writing were coded by trained research assistants who were blinded to condition.

Alphabet-writing fluency. For this task, children were asked to write the entire lowercase alphabet as quickly and as accurately as possible in 1 minute. The alphabet-writing fluency (AF60) task has been widely used by writing researchers to assess handwriting automaticity (e.g., Graham et al., 1997; Jones & Christensen, 1999; Puranik & Al Otaiba, 2012) with good concurrent validity (correlations ranging from .46 to .63 with criterion writing measures of handwriting, spelling, and composition for KG children; Puranik, Patchan, Sears, & McMaster, 2017). One point was awarded for each correctly written letter. Partial credit (i.e., 0.5 point) was awarded for letters that were uppercase, reversed, or recognizable but contained formation or control errors. The final score was the number of points earned in 1 minute.

Sentence writing. To measure students' ability to write sentences, a picture-prompt task was used. Students were given picture prompts and had to write a sentence describing the pictures. They were given 3 minutes to write as many sentences as possible describing the pictures. The sentences were coded for the number of correct word sequences (sentence writing, correct word sequence [SWCWS]) and

words written (sentence writing, words written [SWWW]) as quantitative indicators. A *correct word sequence* was two adjacent words that were correctly spelled and grammatically (e.g., capitalized and punctuated) and semantically acceptable within the context of the sentence. This measure has been widely used in previous studies with KG and first-grade students (e.g., Coker & Ritchey, 2010; McMaster, Du, & Petursdottir, 2009), with good reliability and validity (alternate-form reliability: correct word sequences, $r > .70$; criterion-related validity, $r = .50–.60$). *Words written* was the count of the total number of words written by the student without considering correct or incorrect usage in context or spelling. Intraclass correlation coefficients (ICCs) were used to determine interrater reliability, which was high at the beginning of KG (ICC = .97) and at the end of KG (ICC = .99).

In addition to the quantitative indicators, a qualitative score (sentence writing quality [SWQ]) was calculated to evaluate each student's ability to compose meaningful text using a rubric developed by Coker and Ritchey (2010) and used in previous studies with KG students with good concurrent validity correlations with standardized and unstandardized measures of writing ($r = .56–.65$; Puranik, Patchan, Sears, et al., 2017). The rubric includes five scoring categories: response type, spelling, mechanics, grammatical structure, and relationship to prompt. Each response was given a rating of 0 to 3 in each category. Ratings for each category were summed to determine the total qualitative score for each response (max. score = 15). Scoring was completed with high reliability both at the beginning of KG (ICC = .98) and at the end of KG (ICC = .97).

Essay. Finally, students' ability to write beyond the sentence level was assessed with an essay task in which participants were told that they had to write an essay about what they liked in KG. The same prompt was used at pre- and posttest. The assessor introduced the task and then helped the students brainstorm ideas. Next, the students had 5 minutes to complete their essays. Similar to coding for sentences, the essays were also coded for the number of correct word sequences (essay, correct word sequence [ESCWS]), words written (essay, words written [ESWW]), and essay quality (ESQ). The essay task has been used in previous studies with KG students (e.g., Puranik & Al Otaiba, 2012) with good concurrent validity correlations with standardized and unstandardized measures of writing ($r = .56–.73$; Puranik, Patchan, Sears, et al., 2017). The ICCs for ESCWS and ESWW were .99 at the beginning and at the end of KG; for quality (ESQ), the ICCs were .97 at the beginning and .98 at the end of KG.

Letter-naming fluency. In addition to the writing measures, participants also completed the LNF task from DIBELS (Good & Kaminski, 2002). This assessment was administered individually 3 weeks before the PAWS instruction The LNF task assessed students' ability to name letters. Following the DIBELS protocol, students had 1 minute to name randomly presented uppercase and lowercase letters. The LNF score was the number of correctly named letters in 1 minute.

Questionnaire. To address the social validity of PAWS and to examine how the students felt about the program, we administered a simple four-item questionnaire at the end of the instructional period. The instructor read a scripted introduction and provided instruction to ensure student understanding. The four items designed to examine students' levels of satisfaction were read aloud to the students, and students were required to respond with *yes*, *no*, or *sometimes*. Emoticons for

each response were included to help students. In addition to the four closed-ended questions, students were asked to respond in writing to two open-ended questions: "What did you like about PAWS?" and "What did you not like about PAWS?"

Fidelity of Instruction

A sample of 34 lessons was randomly selected to measure fidelity (i.e., one lesson every other week for each teacher). The PAWS liaisons observed the lesson in person to score the fidelity of instruction. The fidelity measure comprised eight general observations (e.g., Does the teacher gain and maintain students' attention? Does the teacher elicit responses from a variety of students, including students having difficulty with the task? Are the students attentive?) and eight observations specific to the PAWS intervention (e.g., Does the teacher state the purpose or objective of the lesson in the warm-up or I Do section? Does the teacher continually monitor students' understanding and performance during the You Do section? Does the teacher ensure that the writer and coach are working appropriately?). Each observation was scored using a 3-point scale: a score of 1 was given if the target observation never occurred, a score of 2 was given if the target observation occurred sometimes, and a score of 3 was given if the target observation occurred most of the time (i.e., 80% of the time). The fidelity score for each teacher was calculated (i.e., sum of points earned divided by total possible points, up to 48).

Analytic Approach

General linear modeling was used for testing the impact of PAWS on writing outcomes. We acknowledge that students were nested in classrooms and schools; however, the sample of classrooms and schools was too small for estimating robust coefficients. Rather, we opted to treat schools as fixed effects in the model, and they were dummy-coded to reflect medium-performing ($n = 3$) and high-performing ($n = 3$) schools. The general linear model (GLM) was used to estimate the average treatment effect when comparing posttest scores between the PAWS and control groups. Hedges's g was estimated as a calculation of effect size.

In addition to the primary impact models, given our prior study in which school type moderated intervention effects, we also conducted a set of exploratory analyses that tested whether treatment effects were moderated by school type (medium vs. high performing). GLM was used to test the interaction between treatment group and school type. Significant interaction terms were explored through simple slopes analyses (Preacher, Curran, & Bauer, 2006) to isolate the nature of effects.

Results

Missing Data and Descriptive Statistics

Missing data rates ranged from 10% to 13% across the selected measures. Little's test of data missing completely at random resulted in a nonsignificant effect, $\chi^2(27) = 39.02$, $p = .063$, suggesting that multiple imputation methods were reasonable to use in the present design. Although both full information maximum

likelihood and multiple imputation methods are available for addressing missing data, we opted to use the multiple imputation procedure in SAS software with 100 imputations because of the use of listwise deletion in quantile regression. Means, standard deviations, and correlations among the measures are reported in Table 2. Results indicate improvement in both the quantity and quality of writing skills from pre- to posttest. For example, on average, students wrote 3.4 letters in the AF60 task at pretest and 9.6 letters at posttest. Similarly, SWQ scores increased from 2.41 at pretest to 10.18 at posttest. Correlations ranged from .02 between the pretest ESWW and posttest ESQ to .84 between posttest SWQ and posttest SWCWS.

Research Question 1: Adequacy of the Content, Length, and Formatting of Lessons

The dosage across 26 weeks appeared adequate to cover transcription skills and sentence-generation skills. On average, teachers completed the lessons in 30 minutes, with the shortest lesson lasting 28 minutes and the longest lasting 36 minutes. All three teachers in the PAWS condition were able to deliver the lessons with high fidelity (97%, 98%, and 100%, respectively, for the three teachers; $M = 98\%$, $SD = 2\%$).

We were able to administer the posttest student survey to 67 of the 78 PAWS participants. In response to the question, "I liked the activities I did in PAWS," 56 (83.6%) of the students responded *yes*, 5 (7.5%) responded *sometimes*, and 6 (9%) responded *no*. In response to the question, "Did you like working with a classmate?" 59 (88.1%) responded *yes*, 6 (9%) responded *sometimes*, and 2 (3%) responded *no*. In response to the question, "What do you like most about PAWS?" 19 students made some reference to coaching or working with a partner. Some of their responses were "Coaching is fun," "Working with a partner," "I like everything with my classmate," and "It's fun when you work with a partner." Most important, in response to the question, "What do you not like about PAWS?" the responses were more about writing in general rather than something specific about PAWS. Examples include, "We have to do a lot of writing," "Writing sentences is hard," "My hands get tired when I hold the pencil," and "When my PAWS partner is not here." Some students' written responses are included in Appendix B.

Research Question 2: Improvement in Writing Skills

Table 3 shows pretest differences between the control and PAWS groups in writing outcomes. As the data indicate, the students in the PAWS condition had lower scores than the control students on most of the writing outcomes measures, and these differences were statistically significant. Hence, all analysis included covariates with the pretest version of each outcome. In addition, as part of testing for differences between the PAWS and control groups, several baseline covariates were included to raise power for detecting small effects. These included the DIBELS LNF, school type, and gender. Because only six classrooms were included in this quasi-experimental design, a student-level model was used to estimate effects. The test of intervention effects from the GLM showed a statistically significant effect

Table 2. Descriptive Statistics and Correlations for Observed Measures ($N = 144$)

Measure	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. PREAF60	–														
2. PRESWCWS	.53	–													
3. PRESWWW	.53	.81	–												
4. PRESWQ	.42	.65	.75	–											
5. PREESCWS	.48	.57	.50	.53	–										
6. PREESWW	.55	.44	.46	.36	.77	–									
7. PREESQ	.29	.62	.56	.51	.65	.39	–								
8. POSTAF60	.46	.25	.32	.34	.52	.25	.25	–							
9. POSTSWPUC	.13	.11	.15	.18	.45	.19	.13	.28	–						
10. POSTSWCWS	.18	.20	.19	.30	.40	.08	.22	.47	.55	–					
11. POSTSWWW	.19	.15	.18	.26	.30	.10	.19	.50	.29	.80	–				
12. POSTSWQ	.21	.15	.20	.29	.29	.09	.17	.44	.52	.84	.82	–			
13. POSTESCWS	.28	.14	.17	.10	.55	.32	.24	.42	.23	.40	.34	.32	–		
14. POSTESWW	.35	.15	.21	.13	.46	.32	.21	.44	.14	.39	.45	.35	.81	–	
15. POSTESQ	.21	.07	.10	.10	.12	.02	.14	.41	.25	.51	.49	.49	.66	.54	–
M	3.41	.58	1.35	2.41	2.34	7.79	1.08	9.60	1.41	9.70	14.07	10.18	5.20	12.40	7.91
SD	2.30	1.91	3.40	2.24	2.29	6.00	2.00	4.77	1.58	5.27	5.47	2.92	5.03	7.75	3.59
Range	.50–13.00	.00–16.00	.00–21.00	.00–11.30	.55–10.00	1.00–24.00	.00–9.50	.00–22.00	.00–4.00	.00–22.00	.00–28.00	.10–13.80	.00–22.00	.00–34.00	.00–15.00

Note.—PRE = pretest; POST = posttest; AF60 = alphabet writing fluency; SWCWS = sentence writing correct word sequence; SWWW = sentence writing words written; SWQ = sentence writing quality; ESCWS = essay correct word sequence; ESWW = essay words written; ESQ = essay quality.

Table 3. Pretest Differences between PAWS and Control Students on Writing Outcomes

	Intercept (Control)			PAWS			
	Estimate	*SE*	*p*	Estimate	*SE*	*p*	Hedges's *g*
AF60	3.99	.31	<.001	−1.08	.42	.0123	−.37
PUC	.34	.12	.008	.25	.17	.152	.19
SWCWS	1.03	.23	<.001	−.85	.31	.008	−.38
SWWW	2.03	.41	<.001	−1.27	.56	.026	−.37
SWQ	2.77	.27	<.001	−.67	.37	.079	−.30
ESCWS	2.63	.49	<.001	−.71	.77	.363	−.32
ESWW	8.81	1.05	<.001	−2.33	.147	.424	−.39
ESQ	1.12	.24	<.001	−.08	.34	.821	−.04

Note.—AF60 = alphabet writing fluency; PUC = punctuation; SWCWS = sentence writing correct word sequence; SWWW = sentence writing words written; SWQ = sentence writing quality; ESCWS = essay correct word sequence; ESWW = essay words written; ESQ = essay quality.

for PAWS compared with the control classrooms after applying a linear step up for multiple hypothesis correction (Table 4) for punctuation ($t = 5.50$, $p < .001$, $g = 0.90$) and SWQ ($t = 3.10$, $p = .001$, $g = 0.48$). No statistically significant effects were observed for the other outcomes; however, small effects were observed for AF60 ($g = 0.08$), SWCWS ($g = 0.28$) and ESQ ($g = 0.18$). A small, negative effect of PAWS relative to control was estimated on the ESWW ($g = -0.36$), and no effect was observed for ESCWS ($g = -0.01$) and SWWW ($g = 0.03$).

Research Question 3: Impact of School Type

The general linear model test of moderation showed that the effect of PAWS was significantly moderated depending on school type (medium vs. high performing). Significant interactions were observed for all outcomes with the exception of SWQ ($p = .059$). Because the sample was not powered for interaction effects at $p < .05$, simple slopes analyses were conducted for all eight outcomes. Graphic results from the test show that the nature of the interaction was such that differences in favor of PAWS were observed strictly within the medium-performing schools. The test of simple slopes (Table 5) confirmed this finding, in which PAWS students were consistently differentiated in their posttest scores from the control in medium-performing schools. Within this school type, the effect size differences ranged from 0.69 for AF60 to 1.96 for punctuation; however, different from the main impacts, stronger effects were estimated in the moderation model. For example, the main effect analysis of ESCWS showed a PAWS effect of $g = -0.01$, but in the moderation analysis, the estimated effect was $g = 1.03$ in medium-performing schools. Conversely, the test of differences between PAWS and control students in high-performing schools revealed moderate negative effects for the all outcomes except punctuation ($g = 0.76$) and SWQ ($g = 0.32$). The negative effect size differences within the high-performing schools ranged from $g = -0.14$ to $g = -0.73$.

Discussion

With the implementation of CCSS (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010), increased demands

Table 4. Main Effect Results from a General Linear Model of PAWS Impact

Parameter	Coefficient	*SE*	*t*	*p*	Parameter	Coefficient	*SE*	*t*	*p*
AF60:					SWQ:				
Intercept	6.91	1.03	6.68	<.001	Intercept	7.70	.65	11.89	<.001
Pretest	.02	.02	.91	.362	Pretest	.25	.11	2.24	.027
LNF	.14	.02	5.67	<.001	LNF	.06	.02	3.62	<.001
School perf	−.80	.80	−1.00	.320	School perf	−.19	.49	−.38	.704
Gender	.86	.73	1.19	.237	Gender	−.19	.45	−.42	.674
PAWS	−.53	.72	−.73	.468	PAWS	1.39	.45	3.10	.002
PUC:					ESCWS:				
Intercept	.36	.34	1.06	.290	Intercept	4.24	1.19	3.56	.001
Pretest	—	—	—	—	Pretest	.00	.00	.84	.401
LNF	.02	.01	3.02	.003	LNF	.08	.03	2.79	.006
School perf	−.23	.26	−.88	.380	School perf	−1.16	.92	−1.26	.210
Gender	−.04	.24	−.17	.869	Gender	.14	.84	.16	.869
PAWS	1.31	.24	5.50	<.001	PAWS	−.03	.83	−.04	.971
SWCWS:					ESWW:				
Intercept	7.07	1.18	5.99	<.001	Intercept	10.99	1.80	6.11	<.001
Pretest	.00	.00	.68	.498	Pretest	.001	.001	1.10	.274
LNF	.12	.03	4.39	<.001	LNF	.13	.04	3.07	.003
School perf	−1.71	.92	−1.86	.065	School perf	−1.38	1.40	−.99	.326
Gender	.63	.83	.75	.453	Gender	1.27	1.27	1.00	.318
PAWS	1.49	.83	1.81	.073	PAWS	−2.00	1.27	−1.59	.114
SWWW:					ESQ:				
Intercept	11.84	1.26	9.43	<.001	Intercept	6.35	.85	7.44	<.001
Pretest	.00	.00	1.34	.184	Pretest	.13	.16	.78	.435
LNF	.12	.03	3.98	<.001	LNF	.04	.02	2.02	.045
School perf	−.95	.97	−.97	.333	School perf	−.11	.66	−.16	.874
Gender	.24	.88	.27	.785	Gender	.39	.60	.65	.518
PAWS	.17	.88	.19	.848	PAWS	.66	.60	1.11	.271

Note.—AF60 = alphabet writing fluency; perf = performance; LNF = letter-naming fluency; PAWS = peer-assisted writing strategies; PUC = punctuation; SWCWS = sentence writing correct word sequence; SWWW = sentence writing words written; SWQ = sentence writing quality; ESCWS = essay correct word sequence; ESWW = essay words written; ESQ = essay quality.

have been placed on children to meet academic standards, including writing. Despite these standards, and the importance and the benefits of teaching writing, research on how to teach writing is sparse. The purpose of the current project was to address some of the gaps in the current research by examining the feasibility of training teachers to teach writing through PAWS based on adult learning principles (e.g., Bransford et al., 1999) and, in turn, the feasibility of PAWS in improving writing outcomes for KG children.

The present study extends the existing research about PALS in reading and adds uniquely to writing research in several ways that have important implications for schools. First, the intended dosage of 26 weeks was adequate for students to learn both transcription and sentence-generation skills. Second, unlike previous research indicating that KG teachers provided very limited handwriting or spelling instruction (i.e., for less than 5 minutes a day; Puranik et al., 2014), on average our teachers found it feasible to complete 30-minute lessons. Third, in light of research regarding teachers' reported lack of preparedness to teach writing, it is encouraging that they could learn to implement PAWS effectively (i.e., with 97% fidelity) and that their students implemented PAWS adequately. Also encouraging was that teachers were able to ensure that their students learned and implemented their roles as coaches

Table 5. General Linear Model Simple Slopes Analysis

	Moderately Performing Schools					High-Performing Schools				
Test	Coefficient	*SE*	*t*	*p*	*g*	Coefficient	*SE*	*t*	*p*	*g*
AF60:										
Control	4.60	1.18	3.90	<.001		6.81	.86	7.91	<.001	
PAWS	3.37	1.32	2.55	.012	.69	−2.01	.80	2.51	.013	−.45
PUC:										
Control	−.16	.38	.41	.680		.32	.26	1.21	.229	
PAWS	2.42	.44	5.55	<.001	1.96	1.12	.28	3.95	<.001	.76
SWCWS:										
Control	4.73	1.35	3.51	<.001		6.19	.97	6.35	<.001	
PAWS	5.61	1.52	3.69	<.001	1.08	−.07	.94	−.07	.941	−.46
SWWW:										
Control	10.01	1.46	6.86	<.001		11.54	1.06	10.90	<.001	
PAWS	3.39	1.64	2.07	.040	.57	−1.05	1.02	1.03	.305	−.20
SWQ:										
Control	6.93	.76	9.18	<.001		7.76	.56	13.94	<.001	
PAWS	2.75	.84	3.26	<.001	.90	.88	.52	1.69	.093	.32
ESCWS:										
Control	1.12	1.32	.84	.520		4.19	.53	7.91	<.001	
PAWS	5.48	1.47	3.75	<.001	1.03	−2.12	2.27	.94	.520	−.47
ESWW:										
Control	6.44	2.00	3.22	.002		11.24	1.40	8.03	<.001	
PAWS	6.02	2.26	2.66	.009	.74	−5.06	1.38	3.66	<.001	−.73
ESQ:										
Control	4.69	.97	4.82	<.001		6.82	.73	9.37	<.001	
PAWS	3.67	1.10	3.34	.001	1.05	−.49	.68	−.72	.471	−.14

Note.—AF60 = alphabet writing fluency; PAWS = peer-assisted writing strategies; PUC = punctuation; SWCWS = sentence writing correct word sequence; SWWW = sentence writing words written; SWQ = sentence writing quality; ESCWS = essay correct word sequence; ESWW = essay words written; ESQ = essay quality.

and writers and worked together adequately with only 8 hours of professional development and some minimal support from liaisons. Fourth, and relatedly, kindergartners enjoyed PAWS, as reflected in the end-of-the-year student surveys and individual student responses (see App. B). So it is not surprising that during the focus group meeting at the end of the year, all three teachers in the experimental condition expressed satisfaction with the curriculum and a desire to continue using PAWS.

A fifth implication of our study related to the promise of PAWS for improving writing outcomes of kindergarteners; it is notable that PAWS led to statistically significant differences on two important writing measures: punctuation and SWQ. We found statistically significant differences for punctuation (ES = 0.90) and the SWQ outcome (ES = 0.48); these results exceed the effect size criteria of 0.25 for "substantively important" from the What Works Clearinghouse (2014). However, no statistically significant differences were detected across the two conditions on the other writing outcome measures, although a small effect size was noted for SWCWS (ES = 0.28) and ESQ (ES = 0.18).

Early results with our small sample size suggest that the impact of PAWS may be moderated by school type. In the high-performing schools ($n = 2$; classrooms in each condition), PAWS did not add substantial improvement beyond typical instruction. Statistically significant differences in favor of PAWS students compared with control students were noted only for punctuation (ES = 0.76) and SWQ (ES =

0.32). In contrast, in the medium-performing school ($n = 1$), the PAWS students (one classroom) outperformed the control students (one classroom) on all writing measures. Differences between pre- and posttest scores were significant, with large effect sizes for AF60 (ES = 0.69), punctuation (ES = 1.96), SWCWS (ES = 1.08), SWWW (ES = 0.57), SWQ (ES = 0.90), ESCWS (ES = 1.03), ESWW (ES = 0.74), and ESQ (ES = 1.05).

Reading researchers have noted for some time now that not all instructional practices are equally effective for all children. Foorman, Francis, Fletcher, Schatschneider, and Mehta (1998) noted that code-based instruction was more useful for first and second graders with weaker phonological awareness skills; that is, children with weaker phonological awareness skills at the beginning of the year demonstrated greater growth in decoding skills when they were in classrooms that emphasized code-based instruction compared with children with stronger phonological awareness skills. Similar results were noted by Juel and Minden-Cupp (2000) for first graders. Connor and her colleagues (Al Otaiba et al., 2011; Connor, Morrison, & Katch, 2004; Connor, Morrison, & Slominski, 2006) extended this line of work not only by accounting for children's initial skills but also by examining specific dimensions of instruction (teacher managed vs. child managed). Collectively, their body of work with children in preschool through second grade has shown child-instruction effects for reading interventions. For example, children with low initial decoding scores achieved greater decoding growth in classrooms with more time spent in teacher-managed explicit decoding instruction. In contrast, for children with initially high decoding scores, the amount of teacher-managed explicit decoding had no effect.

Although we were not adequately powered in this study to test for interactions, the child-instruction effects noted by reading researchers may help explain some of our findings of differential performance in the medium- versus high-performing schools. Perhaps the literacy (reading and writing) instruction in the high-performing schools was adequate, such that PAWS did not add value. Alternatively, the children from the high-performing schools may have had better readiness skills and did not benefit from time spent on transcription skills. Focusing less of the instruction on transcription skills and most of the writing instruction on challenging text-generation skills may have been more beneficial. Of course, these explanations are based on research in reading intervention, and we realize they are speculative. More research with larger samples is needed to understand the differences in performance between different types of schools: Is it related to counterfactual instruction, student abilities, or heterogeneity within classrooms? These are but a few of the variables that require further investigation.

Along the lines of examining whether PAWS works differently for different types of students, it is important to point out that in our initial pilot work in the first phase, students from high-performing schools showed greater gains compared with children from the low-performing schools. In the current study, children from the medium-performing school showed higher gains compared with the students from the high-performing schools. Although these findings appear disparate, it is important to point out that in the first phase, instruction was provided in a small-group format compared with a whole-class format in the current study. Moreover, the instruction was provided by research staff in the first phase, whereas the classroom

teachers delivered the intervention in this study. Examining management (teacher managed vs. researcher managed) and formatting (small group vs. whole class) of instruction would be important considerations in our future work.

Our observations in the control classrooms indicated no differences from the PAWS classrooms in classroom management and classroom environment; however, our observations in the control classrooms were generally superficial. Although there were differences in teacher training and experience (one of the control teachers had a master's degree, and two had 10 or more years of experience), we cannot be certain whether these differences were meaningful. Given our small sample size, it was not possible to conduct any statistical analysis. As pointed out by Lemons, Fuchs, Gilbert, and Fuchs (2014), we need to have as much knowledge about the counterfactual/control conditions as we do about the classrooms in which we perform our experiments. As these researchers have correctly pointed out, the nature of counterfactuals has changed on account of the increased emphasis on provision of evidence-based Tier 1 core reading programs by most schools. Whereas this is good news for educators and researchers, it results in an inability to have true control groups in schools; effect sizes are clearly impacted by the counterfactual. Having better knowledge of the control classrooms might lead to better understanding of the differences (or lack thereof) between control and experimental conditions, that is, the medium- and high-performing schools. Most recently, Greenberg and Abenavoli (2017) made an excellent argument against the "one-size-fits-all" outlook and suggested that researchers need to carefully consider the characteristics of the population when examining the effects of their intervention efforts. In future studies, we need to do a better job of contextualizing classroom instruction not only in the classes in which PAWS was used but also in control classrooms. To better understand the students for whom PAWS works (and does not), we need to carefully consider the characteristics of the students in both experimental and control classrooms.

Although our results indicate that the impact of PAWS was not generally significant in the high-performing schools, two writing outcomes were an exception. What might explain the statistically significant differences in favor of PAWS students compared with the control students for punctuation (ES = 0.76) and SWQ (ES = 0.32) even in the high-performing schools? Our observation of the control classrooms indicated that almost no time was spent on punctuation instruction, which could help explain the difference in performance on the punctuation outcome measure. One component of the SWQ coding was sentence type; students scored higher if they wrote complex or compound sentences as opposed to simple sentences. In PAWS, students worked on sentence combining, which may have had a direct bearing on the types of sentences they wrote.

Why did sentence-combining instruction not have a stronger impact on essay writing quality? One assumption underlying sentence-combining instruction is that this approach increases sentence complexity and makes the process of sentence construction less effortful, freeing up cognitive resources to generate ideas for larger chunks of writing. For the beginning writers in this investigation, the little time spent on composing only at the sentence level may not have been adequate to generalize their newly learned sentence-combining skills beyond the single-sentence level (i.e., while writing many sentences to write an essay).

In summary, our previous study showed that PAWS was feasible and promising in improving writing outcomes for KG children when delivered in a small-group format by research staff (Puranik, Patchan, Lemons, et al., 2017). The current results indicate that PAWS was feasible and showed promise for improving writing by KG children when delivered by classroom teachers, at least in medium-performing schools. This current work adds to the scant evidence base on writing instruction for young beginning writers by showing that PAWS was feasible and showed promise in improving writing outcomes for KG children when delivered by classroom teachers. Previous studies have shown that peer-assisted strategies represent an effective strategy to improve academic outcomes, namely, in reading and mathematics (Ames & Murray, 1982; Fuchs et al., 2001; Greenwood, Delquadri, & Hall, 1989; Silverman & Stone, 1972). The results of this study corroborate and extend previous work by showing that peer assistance or coaching is an effective strategy for promoting early writing skills, even among relatively young children. Being able to provide feedback to a peer requires students to be able to retain information and meaningfully convey that information to their partners when providing feedback. This appears to be an effective strategy for teaching writing, even for young beginning writers.

Limitations and Future Directions

This work makes many contributions. It was guided by integrating two theoretical and empirical frameworks. It focused on important transcription and text-generation skills and used peer-assisted strategies to teaching writing to KG students. PAWS was conducted in a whole-classroom format with teachers delivering the instruction and with minimal support from liaisons. However, as with most research, our study has its share of limitations, perhaps the biggest of which are the relatively small sample size and the number of classes. Small sample sizes, related to the nature of Goal 2 studies with limited resources available for pilot studies, not only restrict generalizability but also reduce power to detect significant differences. These findings need to be replicated in a large-scale randomized controlled trial to further examine the efficacy of PAWS. Although the six classes in this study were randomly assigned to condition, we ended up with a single class each in the PAWS and control medium-performing conditions; therefore, we need to temper our exploratory results regarding school type with caution. It is imperative that these results be replicated with a larger sample before drawing any definitive conclusions. As we consider a larger scale efficacy trial of PAWS, it will be important to include a wider range of schools and to ensure that we have the power to test moderation.

There are many ways in which we can improve on the instructional components of this work. First, instruction occurred only three times per week. This specific duration was chosen based on teacher availability; however, it would be worthwhile to examine whether greater intensity of instruction might result in stronger writing outcomes. We were not able to spend any time on writing beyond the sentence level. Adding story and narrative writing or essay writing would be a useful exercise. It would be beneficial to examine whether differentiating instruction for the higher performing schools—spending less time on handwriting and spelling

and more time on writing beyond the sentence level—on stories, essays, and so forth, would result in increased performance.

Although it is a critical issue, the composition of pairs in studies using peer-assisted strategies has, surprisingly, received very little attention (see Hoogeveen & van Gelderen, 2013). On account of a lack of clear guidance in the literature, student pairs were modeled after PALS. This ensured that there were no large differences between pairs (based on their DIBELS LNF scores). We do not know if this was the best way to pair students when it comes to writing. Several other options such as pairing high-performing students may have yielded different results and would be a good subject for a future project.

We were not equipped to explore whether any teacher-level or other contextual factors contributed to writing outcomes. Another important aspect that will need to be addressed in the future is better understanding the instruction taking place in the counterfactual/control classrooms. Finally, we need to extend findings by examining effects with other vulnerable populations of learners (e.g., English-language learners, children with language impairments) and with low-performing schools. Findings also need to be extended to higher grades. Despite these limitations, the results of the current study provide preliminary evidence of the feasibility, promise, and success of a unique peer-assisted program that requires participants to practice various metacognitive strategies to improve writing outcomes in medium-performing schools. These results are noteworthy because they were implemented by classroom teachers in authentic educational settings.

Appendix A

Sample Student Worksheets

Letter Writing

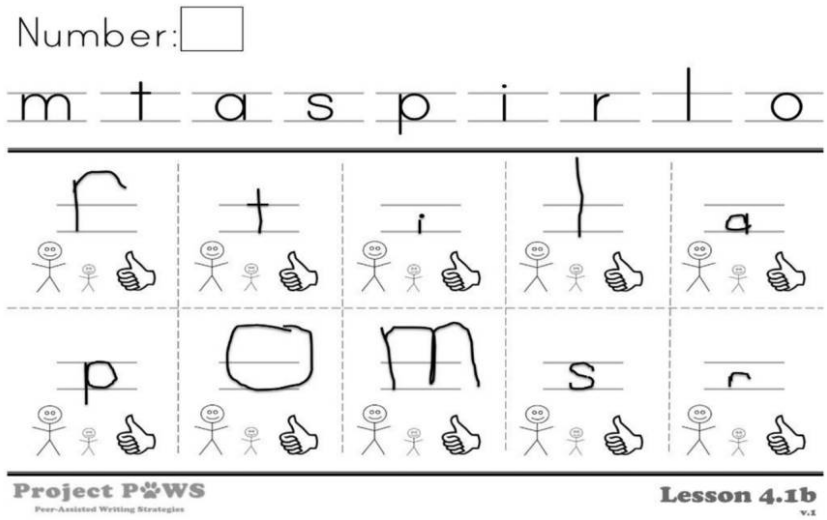

Figure A1. Students were taught how to identify errors of letter size.

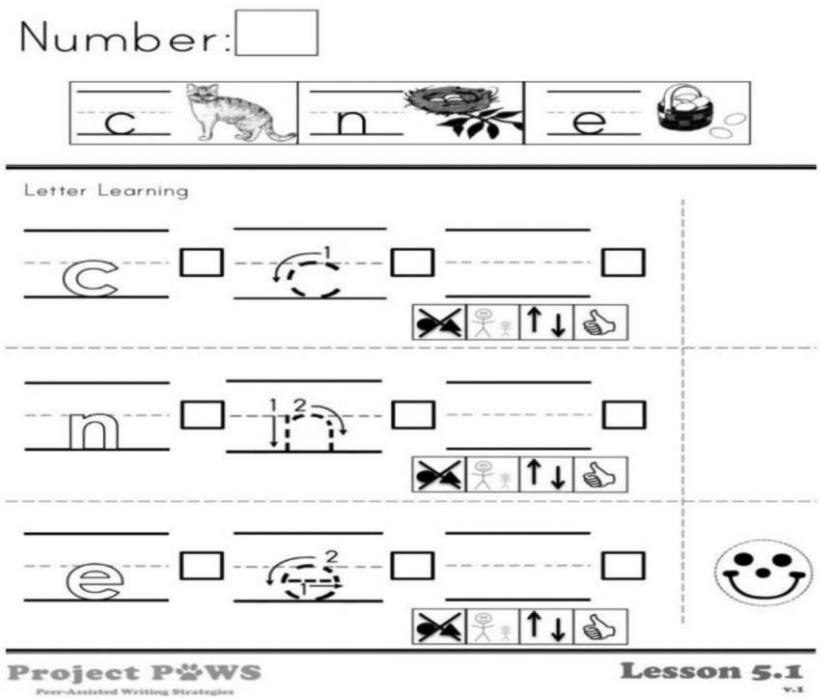

Figure A2. Students provided feedback on letter shape, size, and place.

Spelling

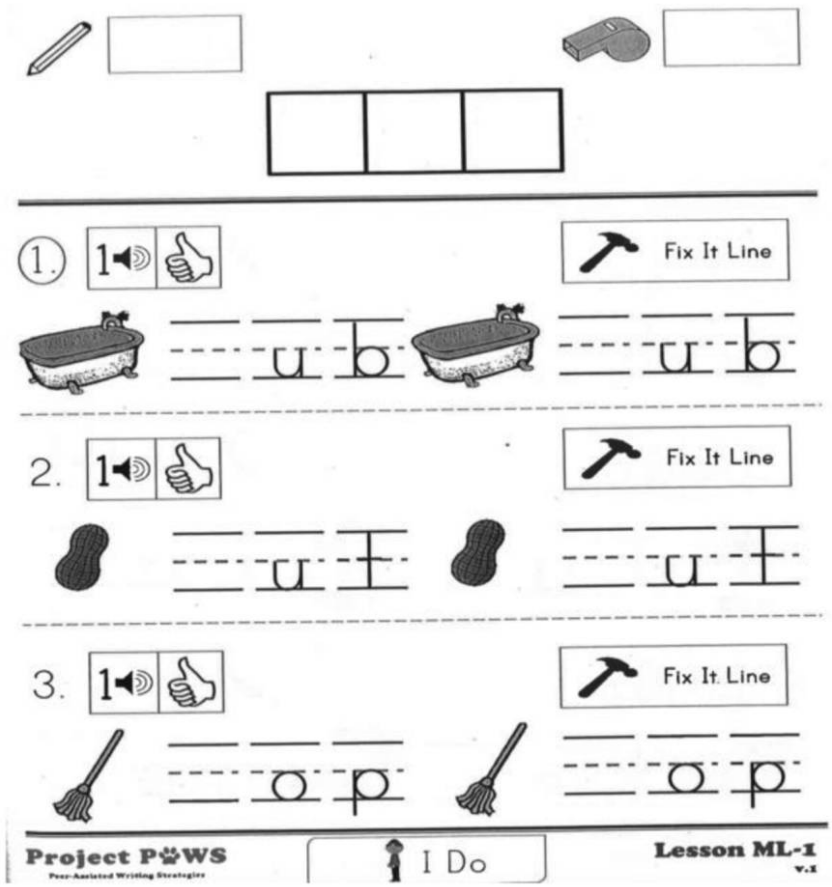

Figure A3. Missing letters, word initial.

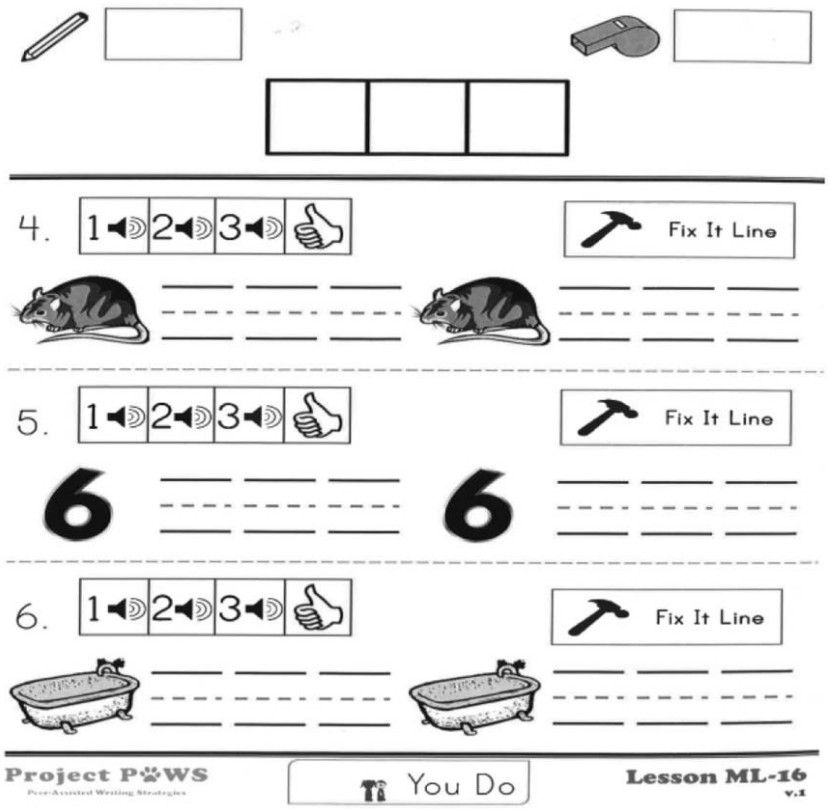

Figure A4. Spelling decodable words.

Sentence Copying

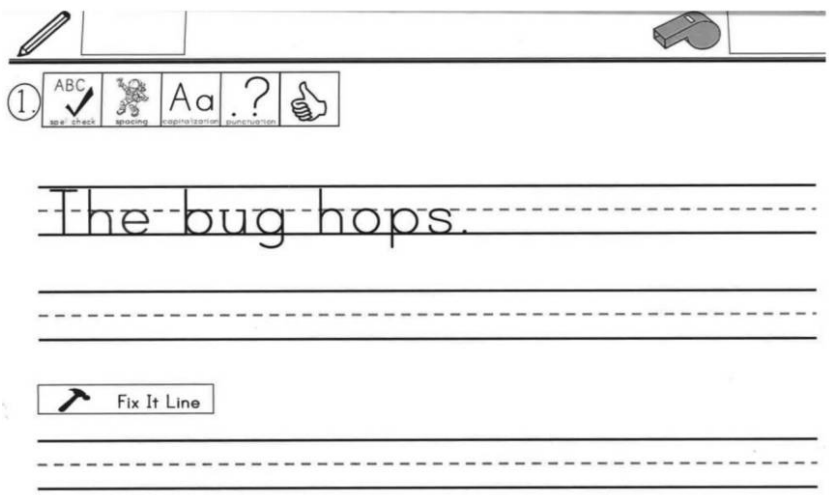

Figure A5. Student coaches were taught to provide feedback on checking spelling, spacing between words, and punctuation.

Sentence Combining

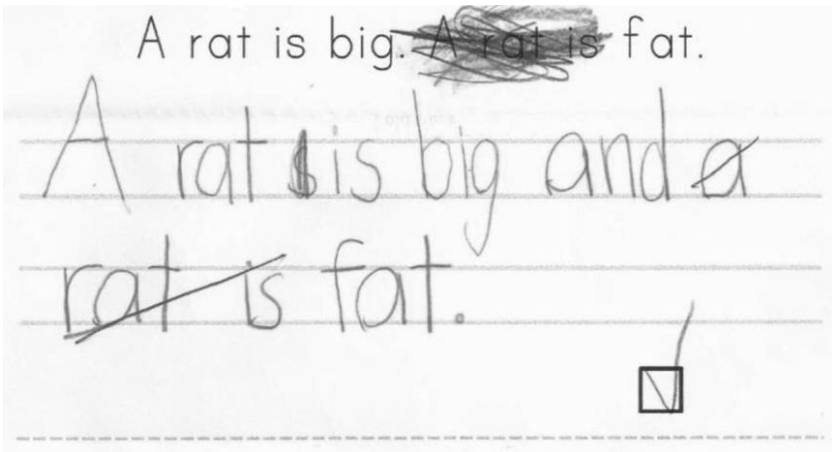

Figure A6. Student coaches provided feedback on combining sentences. Color version available online.

Appendix B

Written Feedback from PAWS Students

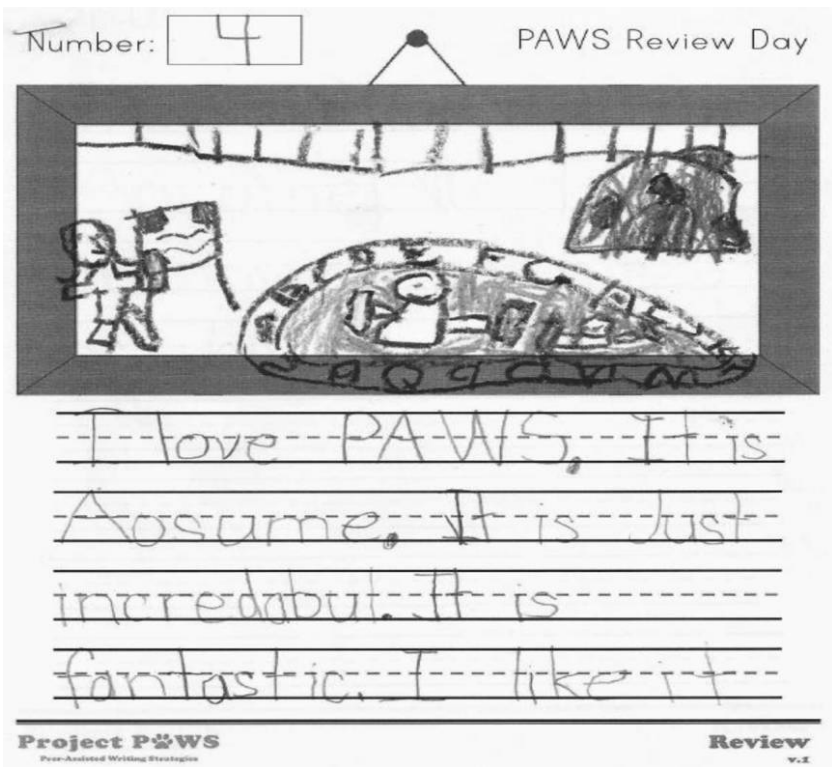

Figure B1. Posttest student survey response number 4 to the question, "What do you not like about PAWS?" Color version available online.

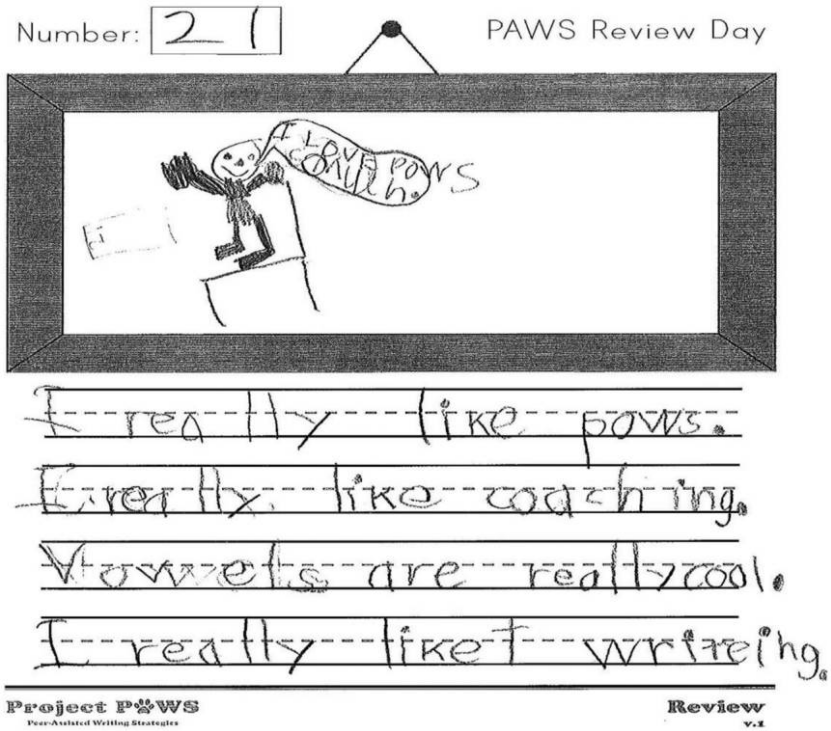

Figure B2. Posttest student survey response number 21 to the question, "What do you not like about PAWS?"

Note

Support for carrying out this research was provided in part by grant R305A120368 from the Institute of Education Sciences. The opinions expressed are those of the authors and do not represent views of the funding agencies. We are very grateful to the teachers, the students, and their parents for participating in this research project. Special thanks to Melissa Patchan, Leah Reiter, Mary Sears, and Kelly Stasik for their invaluable contributions to this project. Without their work, this project would not have been possible. Cynthia Puranik is an associate professor in communication science and disorders and an affiliate of the Center for Language at Literacy at Georgia State University; Yaacov Petscher is the associate director at the Florida Center for Reading Research, Florida State University; Stephanie Al Otaiba is a professor and Patsy and Ray Caldwell Centennial Chair in teaching and learning at Southern Methodist University; Christopher Lemons is an associate professor in special education at Vanderbilt University. Correspondence should be addressed to Cynthia S. Puranik, Georgia State University, 30 Pryor Street, Suite 850, Atlanta, GA 30303; cpuranik@gsu.edu.

References

Al Otaiba, S., Connor, C. M., Folsom, J. S., Greulich, L., Meadows, J., & Li, Z. (2011). Assessment data-informed guidance to individualize kindergarten reading instruction: Findings from a cluster-randomized control field trial. *Elementary School Journal*, **111**(34), 535–560.

Ames, G. J., & Murray, F. B. (1982). When two wrongs make a right: Promoting cognitive change by social conflict. *Developmental Psychology*, **18**(6), 894–897. doi:10.1037/0012-1649.18.6.894

Behymer, A. (2003). Kindergarten writing workshop. *Reading Teacher*, **57**(2), 85–88.
Berninger, V. W. (1999). Coordinating transcription and text generation in working memory during composing: Automatic and constructive processes. *Learning Disabilities Quarterly*, **22**, 99–112.
Berninger, V. W. (2008). Written language instruction during early and middle childhood. In R. Morris & N. Mather (Eds.), *Evidence-based interventions for students with learning and behavioral challenges* (pp. 215–235). New York: Taylor & Francis.
Berninger, V. W., & Hooper, S. (1993). Preventing and remediating writing disabilities: Interdisciplinary framework for assessment. *School Psychology Review*, **22**, 590–595.
Berninger, V. W., Mizokawa, D. T., Bragg, R., Cartwright, A., & Yates, C. (1994). Intraindividual differences in levels of written language. *Reading & Writing Quarterly*, **10**(3), 259–275. doi:10.1080/1057356940100307
Berninger, V. W., Ruthberg, J. E., Abbott, R. D., Garcia, N., Anderson-Youngstrom, M., Brooks, A., . . . Fulton, C. (2006). Tier 1 and Tier 2 early intervention for handwriting and composing. *Journal of School Psychology*, **44**, 3–30. doi:10.1016/j.jsp.2005.12.003
Berninger, V. W., Vaughan, K., Abbott, R. D., Begay, K., Coleman, K. B., Curtin, G., . . . Hawkins, J. M. (2002). Teaching spelling and composition alone and together: Implications for the simple view of writing. *Journal of Educational Psychology*, **94**, 291–304. doi:10.1037/0022-0663.94.2.291
Berninger, V. W., & Winn, W. (2006). Implications of advancements in brain research and technology for writing development, writing instruction, and educational evolution. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 96–114). New York: Guilford.
Bingham, G., Quinn, M., & Gerde, H. (2017). Examining early childhood teachers' writing practices: Associations between pedagogical supports and children's writing skills. *Early Childhood Research Quarterly*, **39**, 35–46. doi:10.1016/j.ecresq.2017.01.002
Bransford, J., Brown, A. L., & Cocking, R. R. (Eds.). (1999). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academies.
Coker, D. L., Jr., 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**, 793–832.
Coker, D. L., Jr., & Ritchey, K. D. (2010). Curriculum-based measurement of writing in kindergarten and first grade: An investigation of production and qualitative scores. *Exceptional Children*, **76**(2), 175–193.
Coker, D. L., Jr., & Ritchey, K. D. (2015). *Teaching beginning writers*. New York: Guilford.
Connor, C. M., Morrison, F. J., & Katch, E. L. (2004). Beyond the reading wars: The effect of classroom instruction by child interactions on early reading. *Scientific Studies of Reading*, **8**, 305–336. doi:10.1207/s1532799xssr0804_1
Connor, C. M., Morrison, F. J., & Slominski, L. (2006). Preschool instruction and children's emergent literacy growth. *Journal of Educational Psychology*, **98**, 665–689. doi:10.1037/0022-0663.98.4.665
Cutler, L., & Graham, S. (2008). Primary grade writing instruction: A national survey. *Journal of Educational Psychology*, **100**, 907–919. doi:10.1037/a0012656
Edwards, L. (2003). Writing instruction in kindergarten: Examining an emerging area of research for children with writing and reading difficulties. *Journal of Learning Disabilities*, **36**, 136–148.
Ehri, L. C., & Wilce, L. S. (1987). Does learning to spell help beginners learn to read words? *Reading Research Quarterly*, **22**(1), 37–65. doi:10.2307/747720.
Fisher, D., & Frey, N. (2008). *Better learning through structured teaching: A framework for the gradual release of responsibility*. Alexandria, VA: Association for Supervision and Curriculum Development.
Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, **32**(4), 365–387. doi:10.2307/356600
Foorman, B. R., Francis, D. J., Fletcher, J. M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology*, **90**, 37–55.
Fry, E., Polk, J., & Foutoukidis, D. (1984). *The reading teacher's handbook of lists*. Englewood Cliffs, NJ: Prentice Hall.

Fuchs, D., & Fuchs, L. S. (2005). Peer-assisted learning strategies: Promoting word recognition, fluency, and reading comprehension in young children. *Journal of Special Education*, **39**, 34–44. doi:10.1177/00224669050390010401

Fuchs, D., Fuchs, L. S., Mathes, P. G., & Simmons, D. C. (1997). Peer-assisted learning strategies: Making classrooms more responsive to diversity. *American Education Research Journal*, **34**(1), 174–206. doi:10.3102/00028312034001174

Fuchs, D., Fuchs, L. S., Thompson, A., Al Otaiba, S., Yen, L., McMaster, K. L., . . . Braun, M. (2001). *Peer-assisted learning strategies: Kindergarten reading*. Nashville: Kennedy Center for Research on Human Development, Vanderbilt University.

Good, R. H., & Kaminski, R. A. (Eds.). (2002). *Dynamic Indicators of Basic Early Literacy Skills* (6th ed.). Eugene, OR: Institute for Development of Educational Achievement.

Graham, S., Berninger, V. W., Abbott, R. D., Abbott, S. P., & Whitaker, D. (1997). Role of mechanics in composing of elementary school students: A new methodological approach. *Journal of Educational Psychology*, **89**, 170–182. doi:10.1037/0022-0663.89.1.170

Graham, S., Harris, K. R., Mason, L., Fink-Chorzempa, B., Moran, S., & Saddler, B. (2008). How do primary grade teachers teach handwriting? A national survey. *Reading and Writing*, **21**, 49–69. doi:10.1007/s11145-007-9064-z

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. doi:10.1037/a0029185

Graham, S., & Miller, L. (1980). Handwriting research and practice: A unified approach. *Focus on Exceptional Children*, **13**, 1–16.

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

Graves, D. (1983). *Writing: Teachers and children at work*. Exeter, NH: Heinemann.

Greenberg, M. T., & Abenavoli, R. (2017). Universal interventions: Fully exploring their impacts and potential to produce population-level impacts. *Journal of Research on Educational Effectiveness*, **10**, 40–67. doi:10.1080/19345747.2016.1246632

Greenwood, C. R., Delquadri, J. C., & Hall, R. V. (1989). Longitudinal effects of classwide peer tutoring. *Journal of Educational Psychology*, **81**(3), 371–383. doi:10.1037/0022-0663.81.3.371

Hall, N. (1987). *The emergence of literacy*. Portsmouth, NH: Heinemann.

Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 1–27). Mahwah, NJ: Erlbaum.

Hayes, J. R., & Berninger, V. W. (2009). Relationships between idea generation and transcription: How the act of writing shapes what children write. In R. K. Bravermen, K. Lunsford, S. McLeod, S. Null, & A. S. P. Rogers (Eds.), *Traditions of writing research*. New York: Taylor & Francis/Routledge.

Hoogeveen, M., & van Gelderen, A. (2013). What works in writing with peer response? A review of intervention studies with children and adolescents. *Educational Psychology Review*, **25**, 473–502. doi:10.1007/s10648-013-9229-z

Jones, D., & Christensen, C. (1999). The relationship between automaticity in handwriting and students' ability to generate written text. *Journal of Educational Psychology*, **91**, 44–49. doi:10.1037/0022-0663.91.1.44

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, **80**(4), 437–447. doi:10.1037/0022-0663.80.4.437

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**, 243–255. doi:10.1037/0022-0663.78.4.243

Juel, C., & Minden-Cupp, C. (2000). Learning to read words: Linguistic units and instructional strategies. *Reading Research Quarterly*, **35**, 498–492. doi:10.1598/RRQ.35.4.2

Kim, Y., Al Otaiba, S., Puranik, C., Folsom, J. S., Greulich, L., & Wagner, R. K. (2011). Componential skills of beginning writing: An exploratory study. *Learning and Individual Differences*, **21**, 517–525. doi:10.1016/j.lindif.2011.06.004

Lemons, C. J., Fuchs, D., Gilbert, J., & Fuchs, L. (2014). Evidence-based practices in a changing world: Reconsidering the counterfactual in education research. *Educational Researcher*, **43**, 242–252. doi:10.3102/0013189X14539189

Lysaker, J. T., Wheat, J., & Benson, E. (2010). Children's spontaneous play in writer's workshops. *Journal of Early Childhood Literacy*, **10**(2), 209–229. doi:10.1177/1468798410363835

McMaster, K. L., Du, X., & Petursdottir, A. L. (2009). Technical features of curriculum-based measures for beginning writers. *Journal of Learning Disabilities*, **42**, 41–60. doi:10.1177/0022219408326212

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects*. Washington, DC: Author. Retrieved from http://www.corestandards.org/ELA-Literacy/

O'Connor, R. E., & Jenkins, J. R. (1995). Improving the generalization of sound/symbol knowledge: Teaching spelling to kindergarten children with disabilities. *Journal of Special Education*, **29**(3), 255–275. doi:10.1177/002246699502900301

Pearson, P. D., & Gallagher, M. C. (1983). The instruction of reading comprehension. *Contemporary Educational Psychology*, **8**(3), 317–344.

Peterson, P. O. (1908). *Peterson directed handwriting*. Greensburg, PA. Retrieved from http://www.peterson-handwriting.com/index.html

Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interaction effects in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics*, **31**, 437–448. doi:10.3102/10769986031004437

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. doi:10.1007/s11145-011-9331-x

Puranik, C. S., Al Otaiba, S., Sidler, J. F., & Greulich, L. (2014). Exploring the amount and type of writing instruction during language arts instruction in kindergarten classrooms. *Reading and Writing*, **27**(2), 213–236. doi:10.1007/s11145-013-9441-8

Puranik, C. S., Lombardino, L. J., & Altmann, L. J. (2008). Assessing the microstructure of written language using a retelling paradigm. *American Journal of Speech-Language Pathology*, **17**, 107–120. doi:10.1044/1058-0360(2008/012)

Puranik, C. S., & Lonigan, C. J. (2011). From scribbles to scrabble: Preschool children's developing knowledge of written language. *Reading and Writing*, **24**, 567–589.

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

Puranik, C. S., Patchan, M., Lemons, C., & Al Otaiba, S. (2017). Using peer assisted strategies to teach early writing: Results of a pilot study to examine feasibility and promise. *Reading and Writing*, **30**, 25–50. doi:10.1007/s11145-016-9661-9

Puranik, C. S., Patchan, M., Sears, M., & McMaster, K. (2017). Alphabet writing fluency in kindergarten: Exploring the issue of time on task. *Assessment for Effective Intervention*, **42**(2), 81–96. doi:10.1177/1534508416665324

Rijlaarsdam, G., Braaksma, M., Couzijin, M., Janssen, T., Raedts, M., Van Steedam, E., . . . Van Den Bergh, H. (2008). Observation of peers in learning to write: Practice and research. *Journal of Writing Research*, **1**, 53–83.

Saddler, B., & Graham, S. (2005). The effects of peer-assisted sentence-combining instruction on the writing performance of more and less skilled young writers. *Journal of Educational Psychology*, **97**, 43–54. doi:10.1037/0022-0663.97.1.43

Silverman, I. W., & Stone, J. M. (1972). Modifying cognitive functioning through participation in a problem-solving group. *Journal of Educational Psychology*, **63**, 603–608. doi:10.1037/h0034082

Slavin, R., Karweit, N., & Madden, N. (1989). *Effective programs for students at risk*. Boston: Allyn & Bacon.

Sulzby, E. (1989). Assessment of writing and of children's language while writing. In L. Morrow & J. Smith (Eds.), *The role of assessment and measurement in early literacy instruction* (pp. 83–109). Englewood Cliffs, NJ: Prentice-Hall.

Topping, K., & Ehly, S. (1998). Introduction to peer-assisted learning. In K. Topping & S. Ehly (Eds.), *Peer-assisted learning* (pp. 1–25). Mahwah, NJ: Erlbaum.

Troia, G., & Graham, S. (2003). Effective writing instruction across the grades: What every educational consultant should know. *Journal of Educational and Psychological Consultation*, **14**, 75–89. doi:10.1207/S1532768XJEPC1401_04

Vander Hart, N., Fitzpatrick, P., & Cortesa, C. (2009). In-depth analysis of handwriting curriculum and instruction in four kindergarten classrooms. *Reading and Writing*, **23**, 673–699. doi:10.1007/s11145-009-9178-6

Vandervelden, M. C., & Siegel, L. S. (1997). Teaching phonological processing skills in early literacy: A developmental approach. *Learning Disability Quarterly*, **20**(2), 63–81. doi:10.2307/1511215

Wagner, R. K., Puranik, C. S., Foorman, B., Foster, E., Wilson, L. G., Tschinkel, E., & Kantor, P. T. (2011). Modeling the development of written language. *Reading and Writing*, **24**, 203–220. doi:10.1007/s11145-010-9266-7

What Works Clearinghouse. (2014). *WWC procedures and standards handbook* (Version 3.0). Washington, DC: Institute of Education Sciences. Retrieved from http://ies.ed.gov/ncee/wwc/DocumentSum.aspx?sid=19

Whitaker, D., Berninger, V., Johnston, J., & Swanson, H. L. (1994). Intraindividual differences in levels of language in intermediate grade writers: Implications for the translating process. *Learning and Individual Differences*, **6**(1), 107–130. doi:10.1016/1041-6080(94)90016-7