WRITING SKILLS, KNOWLEDGE, MOTIVATION, AND STRATEGIC BEHAVIOR PREDICT STUDENTS' PERSUASIVE WRITING PERFORMANCE IN THE CONTEXT OF ROBUST WRITING INSTRUCTION

ABSTRACT

This study tested whether writing skills, knowledge, motivation, and strategic behaviors (within the context of robust writing instruction) each made a statistically unique contribution to predicting fifth-grade students' (123 girls, 104 boys) composition quality and length on a persuasive writing task involving source material, after variance due to other predictors and control variables (reading comprehension, gender, class, and school effects) were controlled. With one exception, writing skills, knowledge, motivation, and strategic behaviors each accounted for statistically unique variance in predicting compositional quality. The exception involved writing knowledge, which did not make a unique contribution in the fall but did in the spring, when a topic knowledge measure was added. In addition, writing motivation, and strategic behaviors accounted for unique variance in composition length in the fall, and writing knowledge did so in the spring.

Steve Graham
ARIZONA STATE
UNIVERSITY

Kausalai Wijekumar
TEXAS A&M UNIVERSITY

Karen R. Harris
ARIZONA STATE
UNIVERSITY

Pui-Wa Lei
PENNSYLVANIA STATE
UNIVERSITY

Evan Fishman
ARIZONA STATE
UNIVERSITY

Amber B. Ray
UNIVERSITY OF HAWAII
AT MANOA

Julia Houston
ARIZONA STATE
UNIVERSITY

WRITING is a complex task. This is illustrated in a model of writing developed by Graham (2018, 2019). The basic assumption underlying this model is that writing is shaped and constrained by the community in which it occurs, as well as by the cognitive capabilities and resources of those who create it. Writing tasks assigned by fifth-grade teachers, for example, conform to the purposes, goals, norms, and values of the class in which they are created. What is written is also subject to the agency of the students producing it. They must make conscious decisions to undertake the task, determine how much effort to commit, formulate their own intentions and goals for achieving it, and decide how to complete it. These decisions are fueled by students' beliefs about the task, writing, and their capabilities as writers. The quality of what is produced depends on the knowledge they have at their disposal (e.g., topic and genre knowledge), mastery of basic writing skills such as handwriting and spelling, and the strategies and schemas they can bring to bear as they plan, draft, and revise text.

Because of its complexity, writing competence is not acquired quickly. It takes many years to develop (Bazerman, 2016; Rijlaarsdam et al., 2012) and requires guidance and instruction (Graham, 2006). Developing writers face multiple challenges when writing. First, they are still in the process of acquiring needed know-how (Graham, 2018). They have not yet mastered fundamental writing skills for translating and transcribing ideas into text. They are still acquiring strategies and schemas for planning, drafting, revising, and editing text. Their knowledge about writing is far from complete, as they are still learning about the characteristics of text and good writing as well as building topic knowledge for future writing projects.

Second, because they are still relatively early in the process of acquiring and mastering the needed know-how, the cognitive actions undertaken by young writers can require a level of conscious attention that exceeds their processing capabilities (McCutchen, 1988), resulting in cognitive overload or interference. For instance, when a younger writer has to consciously think about how to spell a word, this can interfere with other writing processes, as ideas or plans held in working memory may be forgotten (Berninger, 1999).

Third, students are still in the process of forming beliefs about writing and themselves as writers. Such beliefs are not without consequence. They not only influence how much effort is put forth but also provide the impetus for drawing on available cognitive resources such as knowledge, fundamental writing skills, and the strategic processes needed to accomplish writing tasks successfully (Graham, 2018; Pajares, Johnson, & Usher, 2007).

The challenges of learning to write have raised questions about how best to support its acquisition. On the one hand, learning to write is undoubtedly shaped by the communities in which it occurs (Bazerman, 2016; Graham, 2019), as these contexts influence how writing is taught and what is valued. It is unlikely, though, this is the only factor at play. Growth as a writer surely depends on processes that operate at the individual level too (Graham, 2018). As Alexander (2003) argued, learning in any domain is paved by individual changes in cognition and affect.

Cognitive and Affective Factors That Influence Writing Performance

The idea that growth depends on catalysts operating at the individual level is a central feature of the model of domain learning proposed by Alexander (2003). In this model, a domain involves the formalized core conceptual knowledge, including concepts and principles, of a recognized field of study. This includes a variety of academic domains like history, biology, mathematics, reading, or writing, which can differ on a variety of dimensions, such as structure and abstractness. Even though the model of domain learning is based on the assumption that learning is mostly domain specific (e.g., increased interest in history facilitates the acquisition of greater knowledge in this subject), she proposed that the same general types of catalysts advance development within each domain. This includes changes in a learner's knowledge, strategic behaviors, and motivation. For instance, increased motivation in science can lead to greater effort in acquiring knowledge about science as well as more goal-directed behavior in this area, whereas enhanced strategic behaviors relevant to science can result in more efficient science learning and more enthusiasm for this topic.

Alexander (2003) further proposed that learners move incrementally from a state of acclimation to competence within a domain, with a much smaller number of individuals becoming experts. Acclimation involves a learner's initial orientation to the topography of a domain, whereas competence is acquired when one obtains a principled body of knowledge that can be applied to relevant tasks and situations within the domain. In complex domains, competence can take many years to achieve (Graham, 2006).

According to Alexander's (2003) model, movement from acclimation to competence in writing, the focus of this investigation, is driven by changes in writing knowledge, motivation, and strategic behavior. As students become more knowledgeable (e.g., they know more about text structure and their writing topics), motivated (e.g., they become more positive about writing and their capabilities as writers), and strategic (e.g., they apply more sophisticated strategies to plan, draft, and revise their compositions), they become more competent writers. Like other cognitive models of writing (e.g., Hayes, 1996; Zimmerman & Reisemberg, 1997), the application of the domain model of learning to writing emphasizes the importance of writing specific knowledge, strategic behaviors, motivation, or some combination of them. It differs from these models, though, in that it assumes that positive changes in each of them facilitates writing growth.

A systematic review of the writing literature for grades 1–12, conducted by Graham in 2006, examined if there was empirical support for the proposition that changes in knowledge, motivation, and strategic behavior facilitate writing growth. He also examined if writing skills (e.g., handwriting and spelling) played a role in such growth, as they require considerable cognitive resources until they become automatized (Berninger, 1999). This may influence students' knowledge about writing (e.g., students who experience difficulty acquiring these skills may write less, providing them with fewer opportunities to acquire knowledge about the characteristics of text), motivation for writing (e.g., students who easily master these skills may be more confident about their writing capabilities), and strategic writing be-

haviors (e.g., the cognitive demands of writing may result in developing writers minimizing the use of other cognitively demanding processes, such as planning, evaluation, monitoring, and revising; McCutchen, 1988).

Graham (2006) reasoned that if these four factors (skills, knowledge, motivation, and strategic behavior) affected students' growth as writers, the following four tenets should be supported by empirical evidence demonstrating (*a*) that skilled writers possess more of each factor than less skilled writers, (*b*) that developing writers increasingly possess each factor with age and schooling, (*c*) that individual differences in each factor predict writing performance, and (*d*) that instruction designed to increase each factor improves writing performance.

Graham (2006) found relatively strong and consistent support that writing skills and strategic behaviors affect writing growth. Skilled writers were better spellers, were more fluent with handwriting, applied more sophisticated strategies for planning, and were better at revising than less skilled writers (e.g., Bereiter & Scardamalia, 1987; Bourdin & Fayol, 1993; Juel, 1988). In addition, spelling became more accurate, handwriting more fluent, and planning and revising more sophisticated with age (e.g., Bereiter & Scardamalia, 1987; Graham, Berninger, Weintraub, & Schafer, 1998). Individual differences in spelling, handwriting, and planning behaviors (but not necessarily revising) predicted writing performance (e.g., Berninger, Whitaker, Feng, Swanson, & Abbott, 1996; Graham, Berninger, Abbott, Abbott, & Whitaker, 1997). Teaching spelling, handwriting, planning, and revising resulted in better writing (e.g., Berninger et al., 1998; Englert et al., 1991; Jones & Christensen, 1999).

The role of knowledge and motivation in writing was also supported by Graham's (2006) review. He found skilled writers were more knowledgeable about writing and more motivated than less skilled writers (e.g., Englert & Thomas, 1987; Shell, Colvin, & Bruning, 1995). Students became more knowledgeable about writing with age (e.g., Holliway & McCutchen, 2004), but changes in students' writing motivation were variable over time, increasing, decreasing, or remaining constant depending on the study and the type of motivation assessed (e.g., Knudson, 1991; Pajares, 2003). Knowledge of writing and motivation predicted writing performance (e.g., Mosenthall, Conley, Colella, & Davidson-Mosenthall, 1985; Pajares, 2003). Efforts to boost knowledge of writing and motivation enhanced writing performance (e.g., Fitzgerald & Markham, 1987; Schunk & Swartz, 1993). It is important to note that the evidence for knowledge and motivation was not as strong as for skills and strategic behavior, as it involved fewer investigations and less consistent results.

The Current Study

This study provided a test of the third tenet in Graham's (2006) review: individual differences in writing skills, knowledge, motivation, and strategic behaviors each predict writing performance. None of the studies reviewed by Graham controlled for variance related to the other three factors. This may have inflated the obtained associations reported in prior research, providing support for a factor that would not be supported when a more stringent test was conducted. We corrected this shortcoming in this study by controlling for variability associated with the other

factors (e.g., skills, motivation, strategic behaviors) when testing the association between a specific factor (e.g., knowledge) and writing performance. We also controlled for variance due to gender and students' reading comprehension skills, as girls tend to be better writers than boys are at young ages (Walberg & Ethington, 1991), and there is a relatively strong relationship between students' writing and reading (Shanahan, 2016).

Since Graham's (2006) review, several studies have examined some combination of writing skills, knowledge, motivation, and strategic behaviors to predict writing performance. As before, however, none of these studies considered all four of these factors at once. For example, Saddler and Graham (2007) assessed fourth-grade students' knowledge of the role of motivation and strategic behaviors in writing, demonstrating that both predicted the compositional quality and length of story writing for more skilled writers but not for less skilled writers. Olinghouse, Graham, and Gillespie (2015) found that fifth-grade students' writing knowledge (genre and topic knowledge) uniquely predicted composition quality of persuasive and informational text, after writing skills (handwriting and spelling), gender, topic interest, and compositional length were first controlled. Olinghouse and Graham (2009) reported that handwriting fluency (a writing skill) and genre knowledge, but not attitude toward writing (a motivational measure), each made statistically significant and unique contributions to predicting compositional quality and length of stories, after first controlling all other factors including grade, gender, spelling, and word reading. Finally, Graham, Kiuhara, Harris, and Fishman (2017) found that writing motivation (attitudes toward writing and self-efficacy) and strategic writing behaviors (advanced planning and approach to writing) each made statistically significant and unique contributions to predicting compositional quality and length of fourth-grade students' personal narratives, after variance due to the other factors and gender were first controlled.

Furthermore, no previous studies have examined if writing skills, knowledge, motivation, and strategic behaviors predicted students' writing performance at multiple points during the year. The relationship between these factors and writing performance may change from the start to the end of the school year, as students become better writers. It is also notable that most investigations examining such relationships provided little or no information on how writing was taught. When such information was provided (e.g., Graham et al., 2017; Olinghouse & Graham, 2009; Olinghouse et al., 2015; Saddler & Graham, 2007), it is likely that the instruction described (mostly via written surveys) had little impact on these relationships, as these studies were conducted early in the school year. In contrast to previous studies, we collected data at two time points (fall and spring), and teachers were taught how to administer a specific writing program during the interim. Teachers received professional development on how to implement a self-regulated strategy development (SRSD) intervention for persuasive writing that included a Web-based intelligent tutoring system to support teacher instruction (Wijekumar, Harris, Graham, & Meyer, 2017). SRSD is a robust writing intervention (Graham, Kiuhara, McKeown, & Harris, 2012), designed to enhance strategic writing behaviors, knowledge, and motivation (Harris, Graham, Mason, & Friedlander, 2008). This is the first study, to our knowledge, that examines if writing skills, knowledge,

motivation, and strategic behaviors predict students' writing performance before and after robust writing instruction. We also controlled for class and school effects in our statistical analyses.

Research Questions and Hypotheses

The specific question addressed in this study with fifth-grade students was, Do individual differences in writing skill, knowledge, motivation, and strategic behaviors each predict compositional quality and length at the beginning and end of the school year once the other factors, reading comprehension, gender, and class and school effects are first controlled?

Writing was measured in this study by having students plan and compose a persuasive text about a science topic after reading expository source material on it. This type of writing was emphasized in students' classrooms as well as with state-level writing competencies for grade 5 (Common Core State Standards Initiative, 2010). Consequently, this provided a relevant form of writing to test our hypotheses, as students were familiar with it and did such writing at school.

Our writing measures for the persuasive writing task included compositional quality, which provided an overall index of the caliber of a student's text (Huot, 1990), and compositional length. Although compositional quality is viewed as an essential outcome measure by most scholars (see Graham & Perin, 2007), compositional length as determined by number of words written is not as universally accepted. This second measure was of interest because it provided an index of students' text-generation skills during writing (text generation is challenging for developing writers; Scardamalia & Bereiter, 1986), it is commonly included as a central element in models of early writing development (not-so-simple view of writing; Berninger & Winn, 2006), and it accounts for a significant and sizable amount of variance in compositional quality (Morphy & Graham, 2012).

Hypotheses for writing skills. Two writing skills were assessed: spelling and handwriting fluency. Theoretically, both of these skills affect other writing processes and dispositions toward writing until they are mastered (Berninger, 1999). Handwriting fluency was only assessed at the beginning of the school year, as writing instruction during the school year involved digital tools. Furthermore, we added a new measure at the spring assessment (i.e., topic knowledge), and we needed to make room for it in our already lengthy testing protocol. We hypothesized that handwriting fluency and spelling skills would collectively make a statistically significant and unique contribution to predicting writing performance at the beginning of the school year and that spelling would continue to do so at the end of it. Students who demonstrate greater mastery of these skills are less likely to lose plans and ideas they are trying to hold in working memory as a result of slow handwriting or having to think about how to spell a word, resulting in better and longer essays (Berninger, 1999).

Hypotheses for writing knowledge. We assessed two forms of writing knowledge: knowledge of discourse markers in text and topic knowledge. The first measure was assessed at both time points, and we added the second measure (topic knowledge) to the spring assessment, reasoning that this new measure would be a powerful

addition to predicting writing performance (as was the case in Olinghouse et al., 2015). We hypothesized that the knowledge measures administered in this study would make a statistically significant and unique contribution to predicting writing performance at both testing points. Students who are more knowledgeable of discourse markers in text can use this knowledge to create better structured test (enhancing compositional quality). Students who more know about the writing topic should produce text with more and better connected ideas, increasing both compositional quality and length. This should be the case even when writing from source material, as greater topic knowledge should make it easier for students to take advantage of the source material on the topic.

Hypotheses for writing motivation. We assessed two aspects of motivation: attitude toward writing and writing self-efficacy. We hypothesized that these motivation measures would collectively make a statistically significant and unique contribution to predicting writing performance at both time points. Students who express a more positive attitude toward writing and are more efficacious about their writing ability are more likely to persist when writing and put more effort into the task, resulting in better and longer essays.

Hypotheses for strategic writing behaviors. In terms of strategic behaviors, we included a measure of advanced planning (students were asked to plan in advance of writing). We included a second strategic measure that was broader than advanced planning, though, as it assessed the degree to which students engaged in strategic behavior such as analyzing the writing assignment, planning, drafting, maintaining focus, and thinking about their audience. We hypothesized that these two measures of strategic writing behaviors would collectively make a statistically significant and unique contribution to predicting writing performance at both time points. Students who score higher on such measures are more likely to carry out strategic activities that helped them produce better (e.g., plan, consider the readers' needs, analyze the assignment) and longer (e.g., maintain focus) text.

Method

Schools and Teachers

This study involved 13 fifth-grade classes at four elementary schools in a single state in the United States. Three of the participating schools were public institutions, and one was a private Catholic school. Two of the public schools were located in suburban areas, and the other public school was in a mostly rural locale. The private school was situated in an urban neighborhood and enrolled 143 students. All three public schools, in contrast, enrolled 693 to 776 students. Two of the public schools served only elementary-grade students, whereas the third public school was a middle school (grades 5–8). More than 90% of children in one of the public schools were eligible for free or reduced lunch prices, and 75% of them were from ethnic minority groups. In the other three schools, 23% of students on average were eligible for free or reduced-price lunch, whereas 5% of them were from ethnic minority groups. Data on free or reduced-price lunch or ethnic minority enrollment were not available for the Catholic school.

Students

The 286 students in 13 fifth-grade classrooms were invited to participate in the study through a letter and parental consent form sent to the family. Parent consent and student assent was obtained for 246 students (86% of the students in the classrooms). Two hundred and twenty-seven students completed all tests administered in the fall and the spring, and another 12 students completed just the posttests administered in the spring. These 239 students provided the data for this study. Fifty-four percent of these students were girls. Twelve percent of students had an Individualized Educational Program, and each student received assessment accommodations. Sixty-three percent of students were White; 35% were Black. Data for 8% of students in the fall and 3% in the spring were not available because of absences or extracurricular activities.

Writing Instruction in the Schools

Teachers in the 13 classrooms received professional development to implement a writing program that blended SRSD (Harris & Graham, 2016) with a Web-based intelligent tutoring system (Wijekumar et al., 2017). The focus of the SRSD instruction was to teach students to write persuasive text. Teacher-led SRSD instruction was supported through the use of an intelligent tutoring system that taught students how to use text structure to better comprehend text they read. Students must read and understand text to incorporate it when writing from source material. Two of the SRSD lessons were interfaced with a new Web-based intelligent tutoring system designed to support teacher-led instruction. Teachers delivered all instruction between the two assessments.

Using a gradual release model (Harris & Graham, 2016), the SRSD-instructed students were taught strategies and needed knowledge to help them plan and draft persuasive text. These strategies integrated the writing process; developmentally appropriate use of genre knowledge, including genre elements and characteristics of effective persuasive writing; and self-regulation of writing. Students were taught a general writing process strategy represented by the mnemonic POW (pick your ideas, organize your notes, write and say more); a strategy for helping them further focus their writing efforts and use of genre knowledge represented by the mnemonic TAP (topic, audience, purpose); and a planning strategy for generating possible ideas for the their persuasive text (this step helped them organize their notes— see POW, above), represented by the mnemonic TREE (topic, reasons [three or more], explanations [one or more for each reason], ending). Development in the use of TREE involved continued learning about effective writing to persuade. Students were also taught how to apply self-regulation strategies (goals setting, self-monitoring and recording, self-instruction, and self-reinforcement) to help them manage the three writing strategies as well as the writing process and their writing behaviors.

Teacher-led SRSD-instructed lessons were interfaced with a Web-based intelligent tutoring system. For example, Web-based supports included interactive activities to help students identify parts of model essays and assesses students' understanding of the writing strategies being taught. It also provided an assessment mechanism

where teachers could review student progress and customize their teacher-led follow-up lessons to the students' mastery of content.

Teacher-led and Web-based instruction also included instruction on how to use text structure to plan and organize reading and writing using the comparison and problem-and-solution text structures. This involved weaving text structure (e.g., problem and solution) into the reading, selecting, writing a main idea, and encoding tasks central to the reading in preparation for writing. The problem-and-solution strategy promoted in this instruction was designed to support selecting information for writing, encoding information in strategic memory structures, and organizing and signaling the persuasive essay.

This instruction occurred over a 3-month period, and each teacher was observed at least four times as they implemented SRSD and the intelligent tutoring system. Observers reported that procedures were implemented with fidelity (with only minor implementation problems occurring), although a formal assessment of fidelity was not conducted. It should further be noted that between the fall and spring assessments, teachers devoted instructional time to teaching sentence and paragraph skills.

Measures

All measures were administered in October (fall) and May (spring). Testers were taught to administer each measure until they could do so without error.

Compositional quality and length. Measures of compositional quality and length were obtained from essays written during the fall and spring assessments. At each testing point, students were directed to write a persuasive essay where they clearly stated and defended their position on a topic (e.g., wearing a helmet when riding a bike; saving water).

In the fall, writing prompts were randomly assigned to classes. The assignment of writing prompts was then counterbalanced by class in the spring so that each student wrote about a different topic at each testing point. Counterbalancing served two purposes. One, it allowed us to test if the writing prompts were equivalent in terms of students' performance on our primary writing outcome measure: writing quality. This was the case, as there was no statistically significant difference between prompts, $F(1, 2) = .83, p = .48$ (means differed by less than 0.25 point). Two, counterbalancing reduced the likelihood that findings at each time point were due to the topic tested.

Before writing their persuasive essays, students were directed to read an informational article on the topic (e.g., saving water). This text provided information and facts relevant to the topic (e.g., "Most of the earth is covered by water, but only 1% of the water on earth can be used by people"). The texts were approximately 330 words in length, fit on a single page, and were written at a level that grade 5 students should be able to read. The Flesch-Kincaid grade-level readability score for texts were at the middle of the second-grade level (2.5–2.7), whereas the ease-of-reading score for the Flesch-Kincaid was in the low 90s (92.4–93.4).

In administering the writing prompts, the test administrator told students they were going to write an essay to their peers clearly stating and supporting their position on a topic. Students were then assigned their topic (e.g., saving water), directed to read the short text about it, and plan their essay in advance of writing. They were

told they had 35 minutes to complete this task, and they received a 5-minute warning at 30 minutes.

All essays were typed, spelling errors corrected, and identifying information removed before scoring. Compositional length was number of words (determined with the "word count" option in Microsoft Word).

Compositional quality was measured using a holistic writing scale (Huot, 1990). Raters were asked to read each persuasive essay attentively, but not laboriously, to obtain a general impression of overall persuasive compositional quality. Students' papers were then scored on a 9-point scale, with higher scores representing higher writing quality. To determine the score for each essay, examiners were told persuasiveness, ideation, organization, aptness of word choice, grammar, and sentence structure should all be taken into account when forming a judgment about the quality of persuasive text, and that no single factor should receive undue weight.

To guide the scoring process, raters were provided with a representative paper or anchor point for scores of 2, 4, 6, and 8. These anchor-point papers were obtained from students in three fifth-grade classrooms that did not participate in this experiment. All students in these classes wrote a persuasive paper from source material following the same instructions applied in this study. After reading all essays in these three classes, three former intermediate-grade teachers independently selected multiple anchor papers for the scoring points identified above. They then worked together to identify a single paper to represent each score. These papers were randomly ordered, and three different former teachers independently sorted them from lowest to highest quality. All three teachers placed them in the same order identified by the first three teachers.

Two trained raters independently scored all essays after the spring assessment. Interrater reliability using Pearson r was .82.

Writing skills. Two measures assessed students' writing skills. Spelling proficiency was measured as the percentage of incorrectly spelled words in students' essays. The number of incorrectly spelled words was divided by the total number of words in a paper. Two trained graduate assistants (GAs) independently scored papers. One GA scored all papers; the second GA scored a random sample of 25%. The interrater correlation (Pearson r) was .96.

Handwriting fluency was assessed with a copying task where students were asked to copy a simple 59-word fourth-grade passage taken from a standardized reading test (Monroe & Sherman, 1996). The paragraph was read to students, and they were directed to copy the paragraph as quickly and accurately as possible. Students were told to stop copying the paragraph at the end of 60 seconds. One point was awarded for each letter copied correctly (substitutions, reversals, omissions, and additions were considered errors and received no points). The final score was the total number of correctly copied letters. Two trained GAs independently scored the copying task; one GA scoring all of them, and the other GA scored a random sample of 25% of them. The interrater correlation using Pearson r was .98.

Writing knowledge. Two measures assessed students' writing knowledge. One measure assessed students' knowledge of discourse markers using a cloze task. With this task, students supplied four missing words in a comparison text structure passage about two different animals. They were directed to supply the best word for each blank. To do so, they had to think about the logical connections be-

tween the ideas presented in text and generate words that signaled a compare and contrast between the two target animals. The first blank replaced a common word, "different," but succeeding blanks required students to draw inferences about text connections from material presented in text and use less common words ("unlike," "smaller," and "the same as").

There were two equivalent forms of the discourse marker measure. One form was a passage comparing pygmy and emperor monkeys. The second passage compared emperor and Adélie penguins. Both passages were 128 words long, containing 15 sentences and 96 idea units. Student responses were scored using a computer algorithm designed to check for the closeness of the answer to the ideal response. A score of 7 was awarded when a response was an exact match to the missing cloze item; a score of 6 was an exact match spelled incorrectly or a response that contained part of the missing signaling word (e.g., "same" for "same as"); a score of 5 was a similar comparison signaling word that conveyed the same intent as the missing cloze item (e.g., "also like" for "same as"); a score of 4 was a similar comparison signaling word with a different intent (e.g., "smaller than" when "larger" fit the context); a score of 3 was a signaling word that was not a comparison signaling word (e.g., "solution" for "same as"); a score of 2 was a word that showed an understanding that two animals were being compared (e.g., "joining" instead of "same as"); a score of 1 was awarded for any word that did not meet the criteria above.

The second measure assessed knowledge about the assigned writing topic. Knowledge about the topic (e.g., saving water) was solicited via an open-ended question that asked students to tell everything they knew about it. To quantify students' knowledge about a topic, their written response was divided into idea units, and the number of unique ideas about the topic were tabulated. An idea unit was defined as a specific and single idea in a student's response. For example, a response such as "Bike helmets protect your head" was considered as one idea unit; whereas "Bike helmets protect your head because they are made of sturdy materials" was scored as two idea units (i.e., 1 = *bike helmets protect your head*; 2 = *because they are made of sturdy material*). Repeated information was not considered as a new idea unit. Furthermore, repetition of an idea that did not provide any new information about the topic was not scored as a new idea unit. To illustrate, "Bike helmets protect your head . . . helmets keep your head safe," was only scored as one idea unit as the second statement did not provide any new information. However, "Bike helmets protect your head . . . especially your skull," was scored as two idea units, as "especially your skull" presents new information that provides a more specific indication of what is protected. Two trained GAs independently scored responses, with one GA scoring all responses and a second GA scoring a random sample of 25% of responses. Interrater reliability using Pearson r was .89.

Writing motivation. Two measures assessed writing motivation. First, students' attitude toward writing was assessed using a self-report instrument with five items: "I enjoy writing"; "Writing is fun"; "I like to write at school"; "I like to write at home"; "Writing is a good way to spend my time." Students indicated their level of agreement with a 5-point Likert-type scale (1 = *strongly disagree*; 5 = *strongly agree*). The first two items were from the Bruning, Dempsey, Kauffman, McKim, and Zumbrunn (2013) scale. The remaining three items were adapted from a scale

by Graham, Berninger, and Fan (2007). Coefficient alpha for this scale for this study was .96.

The second motivation measure assessed self-efficacy for writing. This measure assessed perceived confidence to successfully apply writing conventions and skills, access ideas for writing, and regulate the writing process, asking students to indicate agreement with 13 items: "I can spell my words correctly"; "I can write complete sentences"; "I can punctuate my sentences correctly"; "I can write grammatically correct sentences"; "I can begin my paragraphs in the right spots"; "I can quickly think of the perfect word"; "I can think of many ideas for my writing"; "I can think of a lot of original ideas"; "I know exactly where to place my ideas in my writing"; "I can focus on my writing for at least 35 minutes"; "I can avoid distractions while I write"; "I can start writing assignments quickly"; "I know when and where to use writing strategies." Students responded with a 100-point Likert-type scale; a score of 0 = *no chance*, 15 = *very little chance*, 35 = *little chance*, 50 = *50/50 chance*, 65 = *good chance*, 85 = *very good chance*, and 100 = *completely certain*.

Eleven of the 13 items on the writing self-efficacy scale were from Bruning et al. (2013). We modified one item, changing "I can focus on my writing for at least 60 minutes" to "35" minutes (the time students had to plan and write an essay in our study). We added two items ("I can quickly think of the perfect word" and "I know when and where to use writing strategies"); both skills are important aspects of composing (Rijlaarsdam et al., 2012). Coefficient alpha for this scale for this study was .86. A confirmatory factor analysis conducted using data from this study confirmed that the two motivational measures described here were distinct and viable factors, as was approach to writing scale described above.

Strategic writing behaviors. Two measures assessed strategic writing behavior. One measure was based on the plans students produced for their essays. Plans were scored using a 0- to 5-point scale. A score of 0 was assigned if no plan was evident; a score of 1 if the essay was a verbatim copy of the words written on the planning sheet, a score of 3 if the essay was not an exact copy of the words written on the planning sheet, a score of 4 if students listed several phrases or ideas on the planning sheet, and a score of 5 if the students used a sophisticated planning strategy such as a web, outline, genre-specific planning strategy, and so forth. Two trained GAs independently scored all plans, with one GA scoring all plans and the second GA scoring a random sample of 25% of them. The interrater reliability coefficient (Pearson r) was 1.00.

The second measure, approach to writing, was a self-report measure (adapted from Lavelle, Smith, & O'Ryan, 2002) asking students to indicate if they agreed with the following 10 items: "I give a lot of detail when writing"; "I make sure my writing is organized and easy to follow"; "I closely examine what the writing assignment calls for"; "I start with a fairly detailed outline"; "I use a lot of examples and definitions to make things clear in my writing"; "I easily find good words for what I want to say when writing"; "I plan out my writing and stick with the plan"; "I keep my topic or theme clearly in mind as I write"; "I use my time wisely when writing"; "I think about my readers while I write." Students indicated agreement with a 5-point Likert-type scale (1 = *strongly disagree*; 5 = *strongly agree*). Coefficient alpha for this scale for this study was .79.

Reading. Reading comprehension was assessed using a measure designed by Meyer and colleagues (Wijekumar, Meyer, & Lei, 2012). With this assessment, students read a comparison text (pygmy vs. emperor monkeys or emperor vs. Adélie penguins described earlier). They then wrote a main idea for the passage with it in view.

The main idea generated by the student was scored from 1 to 8, with 8 representing a high score. The main idea was scored using a computer programmed with logic from Meyer's (1985) propositional analysis. The computer-assigned score was then verified by two trained human scorers. The assigned scores reflected students' ability to select important ideas from the text and summarize them by focusing on what was being compared and on what basis they were compared. The interrater reliability (exact agreement) for GAs was 96%; correlation between computer and human scores (exact agreement) was 98%. GAs scores were summed (scores ranged from 1 to 16).

Procedures

All measures described above were administered at each assessment point (fall and spring). The only exceptions were the handwriting fluency measure (administered only in the fall), and the topic knowledge measure (administered only in the spring).

At each testing point, measures were administered on a single day by members of the research team in the presence of the teacher. The test administrator read the instructions for each measure to students. Students completed these measures in the following order: handwriting fluency (topic knowledge was administered in the spring instead of handwriting fluency), persuasive writing from source task, strategic approach to writing, attitude toward writing, writing self-efficacy, text structure, and reading comprehension. We decided to administer the topic knowledge measure before the writing task, as the source material would likely provide students with new knowledge. It must be recognized, though, that this order of administration may have served to prime students' knowledge when they were administered the writing task. As a result, students' responses on the topic knowledge measure was removed as soon as this task was completed.

Analyses

Because students' performance was tested at two time points (fall and spring), we examined if students' writing skills, knowledge, strategic behavior, motivation, and performance when writing persuasively from source material changed across texting points. For each measure administered in the fall and spring, we conducted a one-way ANOVA with repeated measures to determine if there was a statistically significant change in scores over time. Before testing if writing skills, knowledge, motivation, and strategic behaviors uniquely predicted the composition quality and length of students' persuasive essays written with source material, we examined the correlations between outcome and predictor variables at each testing point (Table 1).

Table 1. Correlations between Writing Measures at Fall and Spring Assessments

Measure	1	2	3	4	5	6	7	8	9	10
1. Composition quality	–	.818**	.072	–	−.055	.425**	.306**	.212**	.227**	.408**
2. Composition length	.728**	–	.264**	–	−.058	.396**	.204**	.108	.171**	.268**
3. Spelling errors	−.009	.403**	–	–	.023	.069	.024	.134*	.082	−.050
4. Handwriting fluency	.269**	.219**	.049	–	–	–	–	–	–	–
5. Discourse markers	.227**	.033	−.235**	.100	–	.023	.094	.045	.053	.028
6. Topic knowledge	–	–	–	–	–	–	.170**	.279**	.218**	.262**
7. Approach to writing	.202**	.102	−.132	.062	.200*	–	–	.056	.488**	.658**
8. Planning	−.285**	−.360**	−.119	−.041	.142*	–	.075	–	.142*	.134*
9. Attitude	.135*	.117	.067	−.027	.095	–	.533**	.135*	–	.459**
10. Self-efficacy	.311**	.159*	−.134*	.058	.341*	–	.607**	.047	.371**	–

Note.—Fall scores are below the diagonal and spring scores are above the diagonal; spelling errors is the proportion of words misspelled in a student's paper.

* $p < .05$.

** $p < .001$.

Most important, at both assessments (fall and spring), we examined if writing skills (handwriting fluency and spelling errors in the fall; spelling errors in the spring), knowledge (knowledge of discourse markers in the fall; topic knowledge and knowledge of discourse markers in the spring), motivation (attitude and self-efficacy), and strategic behaviors (approach to writing and planning) each made a unique and statistically significant contribution to predicting the compositional quality and length. We tested the unique contribution of each predictor (e.g., motivational measures) at each testing point (fall and spring) and for each writing outcome (compositional quality and length) by examining if there was a statistically significant difference between two models: a full model containing all predictors and control variables (gender, reading comprehension, school, and class) and a reduced model containing all but the predictor of interest (e.g., writing skills).

We used mixed models (multilevel model or random-effects model) with the SAS Mixed procedure to account for the nested data structure of students in classes in schools. We first estimated an unconditional model with random student, class, and school components to assess degree of dependency due to different levels. Intraclass correlations (ICCs) at the classroom level were moderate for compositional length (.09 for both fall and spring assessments) and large for compositional quality (.17 for fall and .18 for spring), and ICCs at the school level were 0 for both fall writing outcomes but small (.03 for composition length) to moderate (.09 for compositional quality) for spring outcomes. Given the nonignorable class-level ICCs, we modeled classroom effects as random in all models. Because number of schools was small, we included a school indicator (or dummy) variable to control for school effects (as fixed effects).

To determine the unique effect of a predictor (e.g., motivation measures), we used the likelihood ratio chi-square test by examining deviance change between two nested models described above (i.e., a full model vs. a reduced model). Pseudo R squares were also calculated for student-level outcomes as proportion of total variance (i.e., sum of Level 1 and Level 2 variance estimates from the unconditional null model) explained by the model in question (Snijders & Bosker, 1999). The full model (Model 1) included the control variables and all the predictor variables. This was compared with four reduced models, each specified by dropping from the

Table 2. Multilevel Models for Composition Quality on the Fall Assessment

Fixed Effect	Model 1 (n = 190)		Model 2 (n = 190)		Model 3 (n = 201)		Model 4 (n = 190)		Model 5 (n = 198)	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Intercept	−.28	.16	−.28	.16	−.17	.15	.15	.16	−.39*	.17
Gender	−.40**	.12	−.40**	.12	−.51***	.12	−.53***	.13	−.42***	.12
Reading	−.01	.02	−.01	.02	.01	.02	.01	.02	.0005	.02
Discourse markers	.01	.01	−	−	.01	.01	.02	.01	.002	.01
Handwriting fluency	.003*	.001	.003*	.001	.003*	.001	−	−	.003*	.001
Spelling errors	−6.31***	1.10	−6.49***	1.08	−6.20***	1.11	−	−	−6.72***	1.14
Attitude	.08	.07	.08	.07	−	−	.02	.07	.04	.06
Self-efficacy	.013*	.005	.014**	.005	−	−	.014*	.006	.013**	.005
Approach to writing	−.05	.14	−.05	.14	.24*	.11	.05	.16	−	−
Planning	−.16***	.04	−.16***	.04	−.17***	.04	−.20***	.04	−	−
Variance components:										
Class	0	−	0	−	0	−	0	−	.01	.03
Residual	.57***	.06	.57***	.06	.60***	.06	.68***	.07	.62***	.07
Deviance	432.1		432.9		468.8		467.1		470.4	
Δ deviance (vs. M1)[a]	−		.8		36.7***		35.0***		38.3***	
R^2	.45		.45		.42		.34		.39	

Note.—School effects are controlled for but not shown. All predictor variables except gender are grand-mean centered. $R^2 = 1 −$ [(Model Level 1 variance + Model Level 2 variance) / (null Model Level 1 variance + null Model Level 2 variance)].

  [a] $df = 1$ for comparison with Model 2; $df = 2$ for comparison with other models.

  * $p < .05$.

  ** $p < .01$.

  *** $p < .001$.

full model a different predictor. Model 2 dropped writing knowledge measures, Model 3 dropped writing skills measures, Model 4 dropped writing motivation measures, and Model 5 dropped strategic writing measures. Because fixed-effect estimates were of main interest, we used maximum likelihood estimation for all models. The mixed-model results for fall and spring writing quality and number of words are presented in Tables 2–5.

Results

Changes in Students' Scores from Fall to Spring

As expected, students became better writers between the two assessments, as compositional quality (fall $M = 4.40$, $SD = 1.61$; spring $M = 5.15$, $SD = 1.76$), $F(1, 226) = 39.21$, $p \le .001$, $\eta_p^2 = .15$; knowledge of discourse markers (fall $M = 13.02$, $SD = 7.07$; spring $M = 15.40$, $SD = 6.86$), $F(1, 196) = 11.88$, $p = .001$, $\eta_p^2 = .06$; and planning (fall $M = 1.52$, $SD = 1.53$; spring $M = 2.78$, $SD = 2.26$), $F(1, 226) = 6.08$, $p = .015$, $\eta_p^2 = .26$, improved. Spelling errors also statistically decreased from fall to spring (fall $M = 0.12$, $SD = 0.23$; spring $M = 0.04$, $SD = 0.05$), $F(1, 226) = 49.78$, $p \le .001$, $\eta_p^2 = .18$. There was, however, a drop in compositional length (fall $M = 117.03$, $SD = 51.60$; spring $M = 106.53$, $SD = 47.43$), $F(1, 226) = 7.73$, $p \le .006$, $\eta_p^2 = .03$, and attitude toward writing (fall $M = 3.54$, $SD = 1.08$; spring $M = 3.05$, $SD = 1.02$), $F(1, 209) = 37.01$, $p \le .001$, $\eta_p^2 = .15$. There was no statistical significance in the difference over time for ap-

proach to writing (fall $M = 3.79$, $SD = 10.55$; spring $M = 3.75$, $SD = 0.56$) or self-efficacy (fall $M = 73.99$, $SD = 15.42$; spring $M = 75.29$, $SD = 16.65$). The fall mean for handwriting fluency was 88.76 ($SD = 40.47$). The spring mean for topic knowledge was 9.35 ($SD = 5.74$).

Correlations between Measures

Correlations between measures are presented in Table 1. At both assessments points, the two outcome measures (compositional quality and length) were strongly correlated, sharing 53% of the variance in fall scores and 67% in spring scores. With one exception (i.e., spelling errors), all predictor variables were statistically related to compositional quality at both time points. However, planning was negatively related to compositional quality in the fall but not the spring. Four predictors (i.e., handwriting fluency, spelling errors, planning, and self-efficacy) were statistically related to compositional length at the fall assessment, and five predictors (i.e., spelling errors, topic knowledge, strategic approach to writing, attitude, and self-efficacy) to this outcome measure at spring. Statistically significant correlations between predictors and each outcome measure shared no more than 18% of variance in common in any instance.

Correlations between predictors at each testing point were small to moderate in magnitude (see Table 1), sharing no more than 43% of variance in any instance. The two motivation measures (attitude and self-efficacy) were statistically related at both assessment points. In the fall, the knowledge of discourse markers was statistically related to all but one predictor (i.e., attitude), whereas planning and attitude as well as spelling errors and self-efficacy were statistically related. In the spring, topic knowledge was statistically related to the two motivation variables, strategic approach to writing, and planning, whereas planning was also statistically related to the two motivation measures as well as spelling errors. Strategic approach to writing was also statistically related to the two motivational measures.

Do Writing Skills, Knowledge, Motivation, and Strategic Behaviors Predict Writing Performance?

Fall assessment. For the fall assessment, the full model with all predictor variables (Model 1) explained 45% of the total variance in compositional quality (see Table 2). Writing skills, $\chi^2 = 35.0$, $df = 2$, $\Delta r^2 = .11$; writing motivation, $\chi^2 = 36.7$, $df = 2$, $\Delta r^2 = .03$; and strategic writing behaviors, $\chi^2 = 38.3$, $df = 2$, $\Delta r^2 = .06$, each made a statistically significant unique contribution (all $ps < .001$), beyond the control variables and other predictor variables, in predicting compositional quality. Writing knowledge as measured by discourse markers, however, did not make a unique contribution to predicting compositional quality in the fall, $\chi^2 = 0.8$, $df = 1$, $p > .05$, $\Delta r^2 = .00$. As expected, spelling errors had a unique and statistically negative relation with compositional quality (as did planning unexpectedly), whereas handwriting fluency and self-efficacy each showed a unique and statistically positive relation with compositional quality.

Table 3. Multilevel Models for Composition Length on the Fall Assessment

Fixed Effect	Model 1 ($n = 190$)		Model 2 ($n = 190$)		Model 3 ($n = 201$)		Model 4 ($n = 190$)		Model 5 ($n = 198$)	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Intercept	.28	.17	.28	.17	.40*	.17	.36*	.16	.08	.20
Gender	−.61***	.13	−.61***	.13	−.68***	.12	−.67***	.13	−.63***	.13
Reading	.01	.02	.01	.02	.02	.02	.01	.02	.02	.02
Discourse markers	−.01	.01	−	−	−.002	.01	−.003	.01	−.01	.01
Handwriting fluency	.003	.001	.002	.001	.003	.001	−	−	.003	.002
Spelling errors	−1.06	1.19	−.94	1.17	−.90	1.19	−	−	−1.86	1.24
Attitude	.04	.07	.04	.07	−	−	.02	.07	−.003	.07
Self-efficacy	.01	.01	.01	.01	−	−	.01	.01	.007	.005
Approach to writing	−.08	.16	−.08	.16	.07	.11	−.04	.16	−	−
Planning	−.21***	.04	−.22***	.04	−.23***	.04	−.22***	.04	−	−
Variance components:										
Class	0	−	0	−	0	−	0	−	.05	.04
Residual	.68***	.07	.68***	.07	.69***	.07	.69***	.07	.72***	.08
Deviance	464.6		464.8		497.2		468.5		505.8	
Δ deviance (vs. M1)[a]	−		.2		32.6***		3.9		41.2***	
R^2	.31		.31		.30		.30		.22	

Note.—School effects are controlled for but not shown. All predictor variables except gender are grand-mean centered. $R^2 = 1 −$ [(Model Level 1 variance + Model Level 2 variance) / (null Model Level 1 variance + null Model Level 2 variance)].

  [a] $df = 1$ for comparison with model 2; $df = 2$ for comparison with other models.

  * $p < .05$.

  ** $p < .01$.

  *** $p < .001$.

For compositional length (see Table 3), the full model (Model 1) explained 31% of the total variance. Writing motivation, $\chi^2 = 32.6$, $df = 2$, $\Delta r^2 = .01$, and strategic writing behaviors, $\chi^2 = 41.2$, $df = 2$, $\Delta r^2 = .09$, both made a statistically significant unique contribution (all $ps < .001$), beyond the control variables and other predictor variables, in predicting compositional length (all $ps < .001$). Writing knowledge, $\chi^2 = .2$, $df = 1$, $p > .05$, $\Delta r^2 = .00$, and writing skills, $\chi^2 = 3.9$, $df = 2$, $p > .05$, $\Delta r^2 = .01$, did not make a statistically significant unique contribution to predicting compositional length. Planning had a unique and statistically negative relation with number of words written.

Spring assessment. At spring (see Table 4), the full model (Model 1) explained 57% of the total variance. Writing knowledge $\chi^2 = 23.6$, $df = 2$, $p < .001$, $\Delta r^2 = 0.05$; writing skills, $\chi^2 = 15.0$, $df = 1$, $\Delta r^2 = .05$; writing motivation, $\chi^2 = 10.4$, $df = 2$, $\Delta r^2 = .03$; and strategic writing behaviors, $\chi^2 = 13.7$, $df = 2$, $\Delta r^2 = .04$, all made a statistically significant unique contribution (all $ps < .01$), beyond control and other predictors, in predicting compositional quality. Spelling errors had a unique and statistically negative relation with compositional quality. Topic knowledge, self-efficacy, approach to writing, and planning had a unique and statistically positive relation with compositional quality.

For composition length (see Table 5), the full model (Model 1) explained 34% of the total variance. Only writing knowledge, $\chi^2 = 23.2$, $df = 2$, $\Delta r^2 = .06$, made a statistically significant unique contribution ($p < .001$), beyond control and other predictors, in predicting compositional length. Topic knowledge had a statistically positive relation with compositional length.

Table 4. Multilevel Models for Composition Quality on the Spring Assessment

Fixed Effect	Model 1 (n = 205)		Model 2 (n = 207)		Model 3 (n = 206)		Model 4 (n = 205)		Model 5 (n = 205)	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Intercept	.24	.13	.37*	.13	.21	.13	.27	.13	.43***	.12
Gender	−.44***	.10	−.51***	.10	−.41***	.10	−.49***	.10	−.48***	.10
Reading	.03*	.01	.026	.015	.031*	.015	.043**	.015	.033*	.015
Attitude	−.04	.05	−.02	.05	−	−	−.07	.05	−.01	.05
Self-efficacy	.012**	.004	.014***	.004	−	−	.015***	.004	.018***	.003
Topic knowledge	.04***	.01	−	−	.05***	.01	.05***	.01	.05***	.01
Discourse markers	−.01	.01	−	−	−.01	.01	−.01	.01	−.01	.01
Spelling errors	−4.45***	1.13	−5.05***	1.18	−5.26***	1.12	−	−	−3.97***	1.16
Approach to writing	.30*	.12	.27*	.12	.48***	.09	.27*	.12	−	−
Planning	.08**	.03	.09**	.03	.09**	.03	.07*	.03	−	−
Variance components:										
Class	.00	−	.006	.018	.002	.02	.003	.019	.00	−
Residual	.45***	.04	.49***	.06	.47***	.05	.49***	.05	.49***	.05
Deviance	420.2		443.8		430.6		435.2		433.9	
Δ deviance (vs. M1)[a]	−		23.6***		10.4**		15.0***		13.7**	
R²	.57		.52		.54		.52		.53	

Note.—School effects are controlled for but not shown. All predictor variables except gender are grand-mean centered. $R^2 = 1 -$ [(Model Level 1 variance + Model Level 2 variance) / (null Model Level 1 variance + null Model Level 2 variance)].

[a] $df = 1$ for comparison with model 4; $df = 2$ for comparison with other models.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Table 5. Multilevel Models for Composition Length on the Spring Assessment

Fixed Effect	Model 1 (n = 205)		Model 2 (n = 207)		Model 3 (n = 206)		Model 4 (n = 205)		Model 5 (n = 205)	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Intercept	.11	.15	.24	.15	.09	.15	.12	.15	.23	.14
Gender	−.36**	.12	−.44***	.12	−.35**	.12	−.39**	.12	−.39**	.12
Reading	.02	.02	.01	.02	.01	.01	.02	.02	.01	.02
Attitude	−.01	.06	.01	.06	−	−	−.03	.06	.01	.06
Self-efficacy	.007	.004	.01*	.005	−	−	.01	.005	.01**	.004
Topic knowledge	.05***	.01	−	−	.05***	.01	.05***	.01	.05***	.01
Discourse markers	−.01	.01	−	−	−.01	.01	−.01	.01	−.01	.01
Spelling errors	−2.49	1.36	−3.25*	1.40	−3.03*	1.31	−	−	−2.17	1.36
Approach to writing	.22	.14	.20	.14	.33**	.11	.20	.14	−	−
Planning	.05	.03	.07*	.03	.06	.03	.04	.03	−	−
Variance components:										
Class	0	−	0	−	0	−	0	−	0	−
Residual	.66***	.06	.72***	.07	.66***	.07	.67***	.07	.67***	.07
Deviance	495.9		519.1		499.8		499.2		500.1	
Δ deviance (vs. M1)[a]			23.2***		3.9		3.3		4.2	
R²	.34		.28		.34		.33		.33	

Note.—School effects are controlled for but not shown. All predictor variables except gender are grand-mean centered. $R^2 = 1 -$ [(Model Level 1 variance + Model Level 2 variance) / (null Model Level 1 variance + null Model Level 2 variance)].

[a] $df = 1$ for comparison with M4; $df = 2$ for comparison with other models.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Control variables. Across all analyses (see Tables 2–5), male students had lower scores on compositional quality and length than female students did. Schools (their fixed-effect estimates are not presented in the tables to conserve space) were significantly different on fall outcome writing measures but not on spring outcome writing measures. Reading comprehension was unique and positively associated with compositional quality in the spring but not in the fall.

Discussion

Do Skills, Knowledge, Motivation, and Strategic Behaviors Predict Writing Performance?

We tested the predictive validity of writing skills, knowledge, motivation, and strategic behaviors at two time points (fall and spring). For the fall assessment, our hypotheses that writing skills, motivation, and strategic behaviors would uniquely predict writing quality were mostly supported. Motivation (attitudes and self-efficacy) and strategic behaviors (planning and approach to writing) both made a unique and statistically significant contribution to predicting compositional quality and length, whereas writing skills (spelling and handwriting fluency) accounted for unique and statistically significant variance in compositional quality but not in compositional length. A surprising finding was that there was a negative correlation between planning and what students wrote (this was not the case in the spring after students' planning skills improved). This suggests that at the start of fifth grade, planning in advance of writing may not be an effective tactic for many children. If this finding is replicated in future research, it is important to determine why this is the case.

Following robust writing instruction, all four sets of predictors (writing skills, knowledge, motivation, and strategic behaviors) uniquely and statistically predicted compositional quality for the spring assessment, but only writing knowledge accounted for unique and statistically significant variance in compositional length in the spring. It is important to remember that spring and fall assessments differed in two important ways, and this may have influenced the different patterns of results obtained at the two times points. We added a measure of topic knowledge in the spring, while dropping the handwriting fluency measure.

In the fall, writing knowledge as measured by our knowledge of discourse markers assessment did not uniquely predict either compositional quality or length. When topic knowledge was added in the spring, the two knowledge measures collectively accounted for variance in both persuasive compositional measures (quality and length) beyond the other predictor and control variables entered into the statistical analysis. The most probable explanation for the increased predictive power of writing knowledge was the addition of topic knowledge to this construct. In the spring, the topic knowledge measure alone made a statistically significant and unique contribution to predicting compositional quality and length; knowledge of discourse markers did not. Although students did evidence progress in their knowledge of discourse markers between fall and spring, it is possible that this measure was not sufficiently aligned with the persuasive writing task, as it did not measure discourse markers in persuasive writing specifically. Although additional research is needed to verify these results, it must be noted that the predictive value of topic

knowledge in this study was consistent with findings from previous investigations (Graham, 2006; Olinghouse et al., 2015). However, this study extended these findings to writing from source material.

Writing skills, as measured by spelling and handwriting fluency in the fall and just spelling in the spring, uniquely predicted composition quality at both assessment points but not composition length at either. One possible reason why the removal of handwriting fluency from the spring assessment did not eliminate the predictive value of writing skills for composition quality is the finding that students became better spellers over the course of the school year, and this reduced possible interference of spelling on writing quality.

Over the course of the school year, students wrote better arguments but did so using fewer words. Although SRSD instruction typically produces longer text, there are SRSD studies with persuasive writing where text become better but more concise (Graham, Harris, & McKeown, 2013), as was the case in this study. This may be one reason why skills, motivation, and strategic behavior did not account for unique and statistically significant variance in composition length in the spring. For example, as teachers taught strategies for planning, drafting, and regulating their writing, students may have placed greater emphasis on composition quality than on composition length, reducing the predictive value of this factor for writing output.

A unique aspect of this study was that it examined if writing skills, knowledge, motivation, and strategic behaviors predicted students' persuasive writing performance at the beginning of the school year and at the end of it following robust writing instruction. Although some prior studies provided information on how students were being taught to write at the beginning of the year when data for the investigations were being collected (e.g., Graham et al., 2017; Olinghouse & Graham, 2009; Olinghouse et al., 2015; Saddler & Graham, 2007), they did not provide professional development to teachers that might influence the course of the relationships tested. This did occur in this study, as teachers were taught how to administer SRSD for persuasive writing accompanied by a Web-based intelligent tutoring system to support teacher instruction (Wijekumar et al., 2017). Because this instruction was designed to increase strategic writing behaviors, knowledge, and motivation, this may have influenced the relationship between these variables and students' writing performance on the spring assessment. For example, the negative relation between planning and writing quality in the fall changed to a positive relationship in the spring, and the control and predictor variables accounted for 12% more variance in the spring versus the fall assessment. It was possible, however, that such changes were because the set of predictors in the fall and spring were not identical (i.e., the fall handwriting measure was replaced with a spring topic knowledge measure). In any case, there is a need to better describe the writing instruction students receive in studies such as the current one, as writing context does make a difference (Bazerman, 2016; Graham, 2019), as well as a need to examine more directly the pattern of relationships observed under specific instructional arrangements, as was done in this investigation.

In summary, the evidence from our study provided support for the importance of writing skills, knowledge, motivation, and strategic behaviors for developing writers. For the most important outcome, composition quality, three of the factors tested (writing skills, motivation, and knowledge) were statistically unique predic-

tors at the beginning and the end of the school year. The fourth catalyst, writing knowledge, accounted for statistically significant unique variance once topic knowledge was added to the assessment in the spring. The findings for composition length were less supportive, as only two factors (writing motivation and strategic behaviors) accounted for unique variance in the fall and just one catalyst (writing knowledge) did so in the spring. If these findings are replicated in future research, then it would suggest that a more nuanced approach to considering the impact of skills, knowledge, motivation, and strategic behaviors should be applied, as the predictive value of each depends on how factor is measured.

Limitations and Recommendations for Research

This study replicated and extended findings from prior research. It replicated findings showing that writing skills (e.g., Graham et al., 1997), knowledge (e.g., Olinghouse & Graham, 2009; Olinghouse et al., 2015), motivation (e.g., Graham et al., 2017; Pajares et al., 2007), and strategic behaviors (e.g., Lavelle et al., 2002) predict students' writing performance. It extended this research by showing that this is still the case when variance due to other predictors and control variables (reading comprehension, gender, and class and school effects) are controlled. It also extended prior research by examining these relationships at two time points: before and after robust writing instruction was provided.

Nevertheless, there are several limitations to the current study. This investigation was conducted with only fifth-graders, did not extend beyond a single year, and involved just one type of writing. Thus, the relationships tested in our study need to be examined with younger and older students, over a longer period of time, and with other forms of writing. It is especially important to examine other forms of writing, as students' abilities to write in one genre does not necessarily reflect their abilities to write in a different one (Rijlaarsdam et al., 2012). Thus, our findings that writing knowledge (particularly topic knowledge), writing skills, motivation, and strategic behaviors predict writing performance on a source-based persuasive writing task may or may not replicate with a different genre. To illustrate, Olinghouse et al. (2015) found that topic knowledge predicted persuasive and informative writing performance but not story writing performance.

As in all studies involving writing, we made deliberate decisions about which measures to use for both predictors and outcomes. Replication and extension of the findings from the current study is needed to determine if the same pattern of findings hold when these as well as other measures of writing skills, knowledge, motivation, and strategic behavior are applied.

Because we did not administer exactly the same measures at each time point (handwriting fluency only in the fall and topic knowledge only in the spring), we were unable to statistically compare models across the two time points. We are hopeful that future research will be able to do so.

It is also important to note that this study took place in a specific writing environment. Because context can shape and bound writing performance (Graham, 2018), additional research is needed to determine if the findings can be replicated in other environments and with other forms of robust writing instruction. At a

minimum, we encourage researchers to better describe the type of writing instruction students do receive, so that findings can be interpreted in light of the context in which they occur.

Finally, caution must be exercised when drawing instructional implications from correlational studies such as this one, treating the findings as if they are causal. The current findings, however, are consistent with previous research showing that students become better writers when efforts are made to enhance their writing skills, knowledge, motivation, and strategic behaviors (Graham et al., 2013; Graham & Perin, 2007). What we clearly need to understand better is how to design instruction that brings these catalysts together in productive and optimal ways, in different contexts, and with different students.

Note

References

Alexander, P. (2003). The development of expertise: The journey from acclimation to proficiency. *Educational Researcher*, **32**(8), 10–14.

Bazerman, C. (2016). What do sociocultural studies of writing tell us about learning to write? In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (Vol. **2**, pp. 11–23). New York: Guilford.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Erlbaum.

Berninger, V. (1999). Coordinating transcription and text generation in working memory during composing: Automatic and constructive processes. *Learning Disability Quarterly*, **22**, 99–112.

Berninger, V., Vaughn, K., Abbott, R., Brooks, A., Abbott, S., Rogan, L., . . . 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**, 587–605.

Berninger, V., Whitaker, D., Feng, Y., Swanson, L., & Abbott, R. (1996). Assessment of planning, translating, and revising in junior high writers. *Journal of School Psychology*, **34**, 23–32.

Berninger, V., & 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.

Bourdin, B., & Fayol, M. (1993). *Comparing speaking span and writing span: A working memory approach*. Paper presented at the meeting of the European Association for Research in Learning and Instruction, Aix-en-Provence, France.

Bruning, R., Dempsey, M., Kauffman, D., McKim, C., & Zumbrunn, S. (2013). Examining dimensions of self-efficacy for writing. *Journal of Educational Psychology*, **105**, 25–38.

Common Core State Standards Initiative. (2010). Preparing America's students for success. Retrieved from http://www.corestandards.org/

Englert, C., Raphael, T., Anderson, L., Anthony, H., Stevens, D., & Fear, K. (1991). Making writing strategies and self-talk visible: Cognitive strategy instruction in writing in regular and special education classrooms. *American Educational Research Journal*, **28**, 337–372.

Englert, C., & Thomas, C. (1987). Sensitivity to text structure in reading and writing: A comparison between learning disabled and non-learning disabled students. *Learning Disability Quarterly*, **10**, 93–105.

Fitzgerald, J., & Markham, L. (1987). Teaching children about revision in writing. *Cognition and Instruction*, **4**, 3–24.

Graham, S. (2006). Writing. In P. Alexander & P. Winne (Eds.), *Handbook of educational psychology* (pp. 457–478). Mahwah, NJ: Erlbaum.

Graham, S. (2018). A writer(s) within community model of writing. In C. Bazerman, V. Berninger, D. Brandt, S. Graham, J. Langer, S. Murphy, P. Matsuda, D. Rowe, & M. Schleppegrell (Eds.), *The lifespan development of writing* (pp. 272–325). Urbana, IL: National Council of Teachers of English.

Graham, S. (2019). The writer(s)-within-community model of writing. *Educational Psychology*.

Graham, S., Berninger, V., Abbott, R., Abbott, S., & Whitaker, D. (1997). The role of mechanics in composing of elementary school students: A new methodological approach. *Journal of Educational Psychology*, **89**, 170–182.

Graham, S., Berninger, V., & Fan, W. (2007). The structural relationship between writing attitude and writing achievement in young children. *Contemporary Educational Psychology*, **32**, 516–536.

Graham, S., Berninger, V., Weintraub, N., & Schafer, W. (1998). Development of handwriting speed and legibility. *Journal of Educational Research*, **92**, 42–51.

Graham, S., Harris, K. R., & McKeown, D. (2013). The writing of students with LD and a meta-analysis of SRSD writing intervention studies: Redux. In L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed., pp. 405–438). New York: Guilford.

Graham, S., Kiuhara, S., Harris, K. R., & Fishman, E. (2017). The relationship between strategic behavior, motivation, and writing performance with young, developing writers. *Elementary School Journal*, **118**, 82–104.

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

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

Harris, K. R., & Graham, S. (2016). Self-regulated strategy development in writing: Policy implications of an evidence-based practice. *Policy Insights from Behavioral and Brain Sciences*, **3**, 77–84.

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

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

Holliway, D., & McCutchen, D. (2004). Audience perspective in young writers' composing and revising. In L. Allal, L. Chanquoy, & P. Largy (Eds), *Revision: Cognitive and instructional processes* (pp. 87–101). Boston: Kluwer.

Huot, B. (1990). The literature of direct writing assessment: Major concerns and prevailing trends. *Review of Educational Research*, **60**, 237–263.

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.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grade. *Journal of Educational Psychology*, **80**, 437–447.

Knudson, R. (1991). Development and use of a writing attitude survey in grades 4 to 8. *Psychological Reports*, **68**, 807–816.

Lavelle, E., Smith, J., & O'Ryan, L. (2002). The writing approaches of secondary students. *British Journal of Educational Psychology*, **72**, 399–418.

McCutchen, D. (1988). "Functional automaticity" in children's writing: A problem of meta-cognitive control. *Written Communication*, **5**, 306–324.

Meyer, B. J. F. (1985). Prose analysis: Purposes, procedures, and problems. In B. K. Britton & J. Black (Eds.), *Understanding expository text: A theoretical and practical handbook for analyzing explanatory text* (pp. 11–64, 269–304). Hillsdale, NJ: Erlbaum.

Monroe, M., & Sherman, E. (1996). *Group diagnostic reading aptitude and achievement tests*. Bradenton, FL: Nevins.

Morphy, P., & Graham, S. (2012). Word processing programs and weaker writers/readers: A meta-analysis of research findings. *Reading and Writing*, **25**, 641–678.

Mosenthall, P., Conley, M., Colella, A., & Davidson-Mosenthall, R. (1985). The influence of prior knowledge and teacher lesson structure on children's production of narratives. *Elementary School Journal*, **85**, 621–634.

Olinghouse, N., & Graham, S. (2009). The relationship between the writing knowledge and the writing performance of elementary-grade students. *Journal of Educational Psychology*, **101**, 37–50.

Olinghouse, N., Graham, S., & Gillespie, A. (2015). The relationship of discourse and topic knowledge to writing performance. *Journal of Educational Psychology*, **107**, 391–406.

Pajares, F. (2003). Self-efficacy beliefs, motivation, and achievement in writing: A review of the literature. *Reading and Writing Quarterly*, **19**, 139–158.

Pajares, F., Johnson, M., & Usher, E. (2007). Sources of writing self-efficacy beliefs of elementary, middle, and high school students. *Research in the Teaching of English*, **42**, 104–120.

Rijlaarsdam, G., Van den Bergh, H., Couzijn, M., Janssen, T., Braaksma, M., Tillema, M., . . . Raedts, M. (2012). Writing. In K. R. Harris, S. Graham, & T. Urdan (Eds), *APA educational psychology handbook* (Vol. **3**, pp. 189–227). Washington, DC: American Psychological Association.

Saddler, B., & Graham, S. (2007). The relationship between writing knowledge and writing performance among more and less skilled writers. *Reading and Writing Quarterly*, **23**, 231–247.

Scardamalia, M., & Bereiter, C. (1986). Written composition. In M. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 778–803). New York: Macmillan.

Schunk, D., & Swartz, C. (1993). Goals and progress feedback: Effects on self-efficacy and writing achievement. *Contemporary Educational Psychology*, **18**, 337–354.

Shanahan, T. (2016). Relationships between reading and writing development. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 194–207). New York: Guilford.

Shell, D., Colvin, C., & Bruning, R. (1995). Self-efficacy, attribution, and outcome expectancy mechanisms in reading and writing achievement: Grade-level and achievement-level differences. *Journal of Educational Psychology*, **87**, 386–398.

Snijders, T., & Bosker, R. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London: Sage.

Walberg, H., & Ethington, C. (1991). Correlates of writing performance and interest: A U.S. national assessment study. *Journal of Educational Research*, **84**, 198–203.

Wijekumar, K., Harris, K. R., Graham, S., & Meyer, B. (2017). We-Write: A teacher and technology supported persuasive writing tutor for upper elementary students. In S. Crossley & D. McNamara (Eds.), *Handbook of educational technologies for literacy* (pp. 184–203). New York: Taylor & Francis.

Wijekumar, K., Meyer, B. J. F., & Lei, P. (2012). Large-scale randomized controlled trial with 4th graders using intelligent tutoring of the structure strategy to improve nonfiction reading comprehension. *Journal of Educational Technology Research and Development*, **60**, 987–1013.

Zimmerman, B., & Reisemberg, R. (1997). Becoming a self-regulated writer: A social cognitive perspective. *Contemporary Educational Psychology*, **22**, 73–101.