Read Writ (2019) 32:1431–1457
https://doi.org/10.1007/s11145-018-9836-7

The roles of writing knowledge, motivation, strategic behaviors, and skills in predicting elementary students' persuasive writing from source material

Kay Wijekumar[1] · Steve Graham[2] · Karen R. Harris[2] · Pui-Wa Lei[3] · Ashley Barkel[2] · Angelique Aitken[2] · Amber Ray[4] · Julia Houston[2]

Published online: 27 February 2018

Abstract A core tenet of the model of domain learning is that learning is shaped by cognitive and motivational forces. In writing, these catalysts include learners' knowledge, motivation, strategic behaviors, and skills. This study tested this proposition at two time points (Fall and Spring) with 179 fifth-grade students (52% were girls), examining if writing knowledge, motivation, strategic behavior, and skills each made a statistically significant and unique contribution to predicting writing quality and output on social studies persuasive writing tasks, after variance due to the other catalysts and reading comprehension were first controlled. Three of the four catalysts (writing knowledge, strategic behaviors, and skills) each accounted for statistically significant and unique variance in predicting writing quality, number of words, or both at each assessment point. These findings provided partial support for the model of domain learning as applied to writing.

Keywords Writing · Persuasive writing · Model of domain learning

Writing is a complex task involving a rich array of determinants. It is shaped and constrained by the communities in which it takes place (Graham, in press). In a typical fifth-grade class, for instance, the types of writing that children are assigned and produce conform to the purposes, goals, norms, and values of the class, which in turn are influenced by forces outside the classroom such as school, district, state,

✉ Steve Graham
steve.graham@asu.edu

[1] Texas A&M University, College Station, USA

[2] Arizona State University, Tempe, USA

[3] Pennsylvania State University, State College, USA

[4] University of Hawaii, Honolulu, USA

and national objectives (Bazerman, 2016). To illustrate, with the advent of the Common Core State Standards (CCSS, 2010), students are now expected to write persuasively in response to source material, with CCSS describing the form that such writing should take (see http://www.corestandards.org/ELA-Literacy/W/5/).

Context is not the only factor that influences students' writing. It is also shaped and constrained by the intentions and capabilities of those who produce it (Graham, in press). To continue with our example above, when fifth-grade students are asked to write a persuasive essay using source material, they must make a conscious decision to undertake the task, determine how much effort to commit, formulate personal intentions and goals for achieving it, and decide how to complete it. These decisions are likely fueled by their beliefs about writing, the assignment, as well as their presumed capabilities as writers. The quality of their arguments further depends on the knowledge they have at their disposal (e.g., knowledge of important features of text and their existing knowledge about the subject of the essay), mastery of basic writing skills such as spelling, their ability to extract useful information from source material, and the strategies and schemas they can bring to bear as they plan, draft, and revise text.

Given the complexity of writing, it is not surprising that it takes considerable time for children to become reasonably competent with the types of writing they are expected acquire at school (Bazerman et al., in press; Rijlaarsdam et al., 2012). Developing writers, such as the fifth-grade students referred to here, face multiple challenges. They are still in the process of acquiring needed know-how (Graham, 2006), as they have not yet mastered fundamental writing skills for translating and transcribing ideas into text, their knowledge about writing is incomplete, and they are still acquiring basic strategies for planning, drafting, revising, and editing text. This lack of know-how can impeded writing, as the cognitive actions these children undertake when composing can require a level of conscious attention that exceeds processing capabilities (McCutchen, 1988), resulting in cognitive overload or interference. Further complicating their writing efforts, they are still in the process of forming beliefs about writing and themselves as writers. These beliefs have specific consequences, as they influence how much effort is put forth and how fully children draw on available cognitive resources such as knowledge, fundamental writing skills, and the strategic processes needed to accomplish writing tasks successfully (Pajares, Johnson, & Usher, 2007).

The complexities of writing have raised questions about its development. While learning to write is undoubtedly shaped by the communities in which it occurs (Bazerman, 2016), growth in writing surely depends on processes that operate at the individual level too (Graham, in press). The present study is based on the assumption that individual differences play an important role in students' writing. It draws on Alexander's (2003) model of domain learning, and tests the application of this model to writing persuasively in response to source material in fifth-grade classes where students receive a common approach to writing instruction.

Model of domain learning

Alexander's (2003) model of domain learning proposed that learners move incrementally from a state of acclimation within a domain to a state of competence. She defined a domain as the formalized core conceptual knowledge of a recognized field of study. Acclimation involves a learner's initial orientation to the topography of a domain, whereas competence is obtained once a learner acquires a principled body of knowledge that can be applied to relevant situations and tasks within said domain.

Even though academic domains can differ on a number of dimensions (e.g., structure and abstractness) and learner's growth within a domain is individualistic and variable, Alexander (1997) hypothesized that there are three predictable catalysts that drive growth in all academic domains: learner's knowledge, motivation, and strategic behaviors. Accordingly, within a domain like writing, movement from acclimation to competence arises from writers becoming more knowledgeable (e.g., writers acquire more knowledge about text structure and their writing topics), motivated (e.g., writers become increasingly positive about writing and their capabilities as writers), and strategic (e.g., writers learn more sophisticated strategies for planning, drafting, and revising text). Alexander (1998) indicated that catalysts like knowledge, motivation and strategic behaviors work together to shape learners' development in a domain, but learners can also be propelled to greater competence by an enhancement in a single catalyst.

A previous review (Graham, 2006) provided initial support for Alexander's (1997, 1998) contention that changes in students' knowledge, motivation, and strategic behavior facilitate writing growth (although he did not limit himself to the same measures that were frequently studies in the model of domain learning such as interest for motivation). For all three catalysts, he examined if the following propositions were supported by available research: (1) skilled writers possess more of each catalyst than less skilled writers, (2) developing writers increasingly possess each catalyst with age and schooling, (3) individual differences in each catalyst predict writing performance, and (4) instruction designed to increase each catalyst improves writing performance. Graham argued that all four propositions should be supported by empirical evidence if a catalyst (e.g., motivation) impacts writing growth.

For strategic writing behavior, Graham (2006) found relatively strong and consistent support that it serves as a catalyst of writing growth. Skilled writers applied more sophisticated strategies for planning and were better at revising than less skilled writers. Students' planning and revising became more sophisticated with age. Individual differences in students' planning behaviors (but not necessarily revising) predicted writing performance, and teaching planning, and revising resulted in better writing.

Graham (2006) also found empirical support that writing knowledge and motivation served as catalysts for writing growth. Skilled writers were more knowledgeable about writing and more motivated than less skilled writers. Students became more knowledgeable about writing with age (although the predicted

changes in students' writing motivation were not consistent, as they were variable over time). Knowledge of writing and motivation predicted writing performance, and efforts to boost knowledge of writing and motivation enhanced writing performance. Graham indicated, however, that the evidence for knowledge and motivation was not as strong as it was for strategic behaviors, as it involved fewer investigations and less consistent results.

In addition to knowledge, motivation, and strategic behavior, Graham (2006) also examined if foundational writing skills (e.g., spelling and handwriting) played a role in students' growth as writers, extending the catalysts identified by Alexander (1997, 1998) at least for writing. He made this addition because writing skills require considerable cognitive resources until they become automatized (Berninger, 1999). This may influence students' writing knowledge (e.g., students who experience difficulty acquiring foundational writing skills may write less providing them with fewer opportunities to acquire knowledge about the characteristics of text), motivation (e.g., students who master these skills easily may be more confident about their writing capabilities), and strategic writing behaviors (e.g., until they are mastered, writing skills may be so demanding that students may minimize the use of other cognitively demanding processes such as strategic behaviors like planning; McCutchen, 1988).

Again, Graham (Graham, 2006) found that available empirical evidence supported the role of writing skills in students' writing growth. Skilled writers are better spellers and more fluent with handwriting. Spelling becomes more accurate and handwriting more fluent with age. Individual differences in spelling and handwriting predict writing performance. Lastly, teaching spelling and handwriting improved students' writing.

The current study

Of particular importance to the current study is the third proposition in Graham's (2006) review: individual differences in each catalyst predict writing performance. Alexander's (1997, 1998) model of domain learning proposed that knowledge, motivation, and strategic behavior impact students' growth conjointly and individually. In terms of the third proposition examined by Graham, studies reviewed did not examine the individual effects of a catalyst (motivation) once variance due to the other catalysts were first controlled (knowledge, strategic behaviors, and skills). As a result, it is not clear if the third proposition would be supported if such studies had been available. This may have inflated the obtained associations reported in studies included in Graham's review, yielding support for a catalyst that may not been obtained if a more stringent test had been conducted.

More stringent tests, controlling for variance due to one or more catalysts, have been conducted more recently. For example, Olinghouse, Graham, and Gillespie (2015) found that genre knowledge and topic knowledge each made a unique contribution to predicting the quality of text produced by fifth-grade students once each form of knowledge, writing skills (spelling and handwriting), and motivation (topic interest) were first controlled. They further found that motivation (topic

interest) made a unique contribution beyond the other catalysts to predicting for some types of writing, but this was not the case for the two writing skill measures (handwriting and spelling). An earlier study by Olinghouse and Graham (2009) found that writing knowledge (genre knowledge), motivation (topic knowledge), and one writing skill measure (handwriting fluency) each made a unique contribution beyond the other two catalysts when predicting second and fourth grade students' writing performance, whereas Author reported that writing motivation (attitudes toward writing and self-efficacy) and strategic writing behavior (advanced planning and approach to writing) each made unique contributions to predicting the quality of fourth grade students' writing.

Particularly important to the present investigation was a recent study with fifth-grade students by Graham et al., (2018). They tested Alexander's (1997) model of domain learning, as adapted by Graham (2006) to include writing skills. They tested the model with fifth-grade students at two time points in the school year (Fall and Spring). They specifically examined Graham's third proposition (individual difference in each catalysts predict students' writing performance), but in contrast to previous investigations they included measures for all four catalysts. Writing performance in the Fall and Spring was assessed by asking students to write a persuasive essay after reading source material on a science topic. At both testing points, measures for writing knowledge (knowledge of discourse markers of text structure), writing motivation (self-efficacy and attitude towards writing), strategic writing behavior (pre-planning complexity and approach to writing), and writing skills (spelling) were administered. During the Fall assessment, students also completed a handwriting fluency measure, but this was dropped in the Spring as students composed with a word processor. At the second testing point, writing knowledge was expanded to include students' knowledge of the writing topic. All students in the study received the same basic writing instruction between assessments (Self-regulated Strategy Development instruction (SRSD; Harris & Graham, 2016, 2018) coupled with a web-based intelligent tutoring system (Wijekumar, Harris, Graham, & Meyer 2017), holding context constant at least to some degree.

Alexander's adapted model was generally supported by the Wijekumar et al., (2017) investigation. Each catalyst (writing knowledge, motivation, strategic behavior, and skills) made a unique contribution to predicting the quality of students' writing once the other three catalysts and reading skills were first controlled. The only exception involved writing knowledge in the Fall when only discourse knowledge of text structure was assessed, but it was statistically significant in the Spring when the topic knowledge measure was added as part of the assessment. The findings for the second writing outcome measure, number of words written, however, provided support for the unique contribution of three of the catalysts: writing motivation and strategic behavior in the Fall and writing knowledge in the Spring.

The present study replicated and extended the Graham et al., (2018) investigation. As was done in the previous study, we tested Alexander's (1997) model of domain learning as adapted by Graham (2006) by examining the unique contribution of writing knowledge, motivation, strategic behaviors, and skills in

predicting writing performance at the beginning and end of the school year, allowing us to determine if the obtained relationships remained relatively similar over the school year. As in the previous investigation, students received SRSD persuasive writing instruction combined with a web-based intelligent tutoring system which supported and extended teacher-led lessons. Measures for each of the catalysts were identical, except we did not administer a handwriting fluency assessment (as writing was done on a word processor), a topic knowledge measure was administered at both time points, and we added a measure of strategic writing behavior at each assessment (i.e., number of words written when students planned in advance). Instead of a single measure of reading performance, as was the case in the Author study, we included two measures of reading. We further changed the writing prompt in the current study to focus on writing a persuasive essay with source material in social studies and not science. The outcome measures of writing quality and length remained were retained however.

Writing quality provided an overall index of the caliber of a student's persuasive argument, and is viewed as an essential outcome measure by most writing scholars (see Graham & Perin, 2007). Writing length as measured by number of words is not as universally accepted. Even so, length of students' essays was of interest to us because it provided an index of students' text generation skills during writing. Scardamalia and Bereiter (1986) noted in their seminal review of writing research, text generation is challenging for young developing writers, and generating text is commonly included as a central element in models of early writing development such as the Not So Simple View of Writing (Berninger & Winn, 2006). Although text length does not ensure a paper is well written, this measure does account for a significant and sizable amount of variance in children's writing (Morphy & Graham, 2012; Page & Petersen, 1995).

The specific question addressed in this study with fifth-grade students was do individual differences in writing knowledge, motivation, strategic behavior, and writing skill uniquely and statistically predict writing quality and number of words written at the beginning and end of the school year when variance due to the other three set of predictors and reading skills are first controlled? Based on Alexander's (1997) model of domain learning, the findings from Graham's (2006) review, and more recent studies (e.g., Graham et al., 2018; Olinghouse & Graham, 2009; Olinghouse et al., 2015), we anticipated that each catalyst would uniquely predict the quality of students' writing. The prediction for number of words written were more tenuous given the findings from Graham et al., (2018).

Methods

Schools and teachers

This study involved 10 fifth-grade classes at three elementary schools in a single state in the United States. Two of the participating schools were public institutions (one with five classes and the other with four classes), and one was a private Catholic school (with one class). The two public schools were located in suburban

and urban areas. The private school was situated in an urban neighborhood and enrolled 174 students. The two public schools, in contrast, enrolled 694–636 students. The two public schools served only elementary grade students. Over 57.5% of children in the public schools were eligible for free or reduced lunch prices, and 58.5% of them were minority students. Data on free and reduced lunch or minority enrollment was not available for the Catholic school. All three schools were recruited to participate in a design experiment assessing the effects of a writing intervention. None of the schools or teachers declined to participate.

Students

The 241 students in the 10 fifth-grade classrooms were invited to participate in the study through a letter to the family. Parent consent and student assent was obtained for 180 students (75% of the students in the classrooms). Fifty-two percent of students were girls and 8% of them had an individualized education plan. Fifty-four percent of students were white; 35% were black. The remaining students were mostly Hispanic.

One of the 180 students moved before the start of the study. This left 179 students, with 172 of them completing all tests administered in the Fall (173 completed all tests but the motivation measures). One hundred forty-three students completed all tests in the Spring, with 148 students completing all but the approach to writing measure. Thus, 25 students tested in the Fall were not included in the analysis of data collected in the Spring. The reasons for their exclusion ranged from moving (N = 1), absence (N = 6), to not available due to extra-curricular activities (N = 18). There were no statistical differences between students who completed tests in the Fall when compared to the same students who completed the test in the Spring on any of the measures administered in this study (all p's $> .19$).

Writing instructional context

Between the Fall and Spring assessments, all 10 participating classes provided SRSD instruction for persuasive writing (Harris & Graham, 2016, 2018) combined with a web-based design intelligent tutoring system (Wijekumar et al., 2017). SRSD instruction focused on teaching students to write persuasive text. This teacher-led instruction was supported by three web-based extension modules that provided students with extra practice and assistance with specific skills, knowledge, and strategies presented via SRSD.

SRSD instruction

With SRSD, students learned strategies to help them plan and draft persuasive text. They learned to use these strategies as part of the writing process, integrating them with self-regulation procedures and developmentally appropriate genre knowledge. Students were taught a general writing process strategy represented by the mnemonic POW (**P**ick your ideas, **O**rganize your notes, **W**rite and say more), a strategy for helping them focus their writing efforts and use of genre knowledge

represented by the mnemonic TAP (**T**opic, **A**udience, **P**urpose), and a planning strategy for generating possible ideas for the a persuasive essay (this step helped them **O**rganize their notes—see POW above) represented by the mnemonic TREE (**T**opic, **R**easons [3 or more], **E**xplanations [1 or more for each reason], **E**nding). Students were further 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.

These strategies were taught through the following six teacher-led stages of instruction (see Harris & Graham, 2018; Harris, Graham, Mason, & Friedlander, 2008 for details): Develop Background Knowledge (e.g., vocabulary and knowledge needed to use writing and self-regulation strategies were taught, including knowledge about the basic characteristics of persuasive text), Discuss It (the writing strategies and self-regulation procedures were discussed as well as the differences between weaker and stronger persuasive texts), Model It (the teacher modeled using the writing strategies and self-regulation procedures to write persuasive text, involving students in the production of the plan and text), Memorize It (the steps of the strategy were memorized; this began earlier in instruction but was confirmed here), Support It (students moved from planning and creating persuasive text with others [teacher or peers] and support material [e.g., a graphic organizer for TREE] to independent use of the writing strategies and self-regulation procedures), and Independent Performance (students were able to use all procedures independently and appropriately). Rich dialogue between teacher and students and between peers was emphasized at each stage of instruction; teachers applied the stages recursively as needed (e.g., a teacher returned to the Model It stage after moving to Support It if additional scaffolding was needed).

Web-based SRSD extension modules

Teacher-led instruction was supported by web-based extension modules. For example, during the SRSD instruction, students engaged in teacher-led lessons that supported identifying persuasive elements in model essays (Develop Background Knowledge), assessing problems in weaker essays (Discuss It), or testing their understanding of each step of the writing strategies (Discuss It; Memorize It). The web-based modules supported both the "develop background knowledge" and "memorize it" stages by presenting interactive activities to identify parts of model essays. The web-based activities also included videos about POW, TAP, and TREE and a quiz about the mnemonics used to represent the strategies students were learning to use (supporting both the discuss it and memorize it stages of SRSD). The web-based modules allowed teachers to review student progress and customize their teacher-led follow-up lessons to ensure students mastered the targeted content.

Measures

All measures were administered in October (Fall assessment) or May (Spring Assessment).

Writing output and quality

Students were asked to write a persuasive essay during the Fall and Spring assessments. At each testing point, they were asked write a composition in response to a writing prompt where they were asked to state and defend their position on a topic in social studies (e.g., internet safety or saving money). For example, for internet safety the directions were: "Think about what it means to be safe on the internet. Write an essay to convince your classmates to agree with your position on internet safety." For saving money, students were directed; "Think about what it means to save money. Write an essay to convince your classmates to agree with your position on saving money."

Writing prompts were randomly assigned to classrooms in the Fall, and then counterbalanced in the Spring so that each class wrote about a different topic at this testing point. Counterbalancing served two purposes. First, it allowed us to determine if writing prompts were equivalent. Data for students from this study provided support for this proposition, as there was no statistical difference, F (2,175) = 1.44, p = .24) in the quality of students' papers (this measure is described below) written in response to these writing prompts. Second, counterbalancing of writing prompts reduced the likelihood that findings at each assessment point were due solely to the writing topic assigned.

Before writing their persuasive composition, students were asked to read an article on the assigned topic (e.g., internet safety) that provided relevant information and facts (e.g., There are over 30 billion web pages…Over 2 billion people use the internet…). Each article contained approximately 300 words and fit on a single page. The articles were written at a fourth grade readability level according to the Flesch Kincaid grade level readability formula. The articles are available from the authors.

Students were directed to write an essay to their classmates clearly stating and supporting their position on the assigned topic. They were asked to read the article about the topic, plan their essay in advance of writing (a blank page labeled Planning Sheet was provided), and then write their composition. Students were told they had 35 min to complete this assignment, and that the test administrator would not be able to help them. They were further informed they should spell all words as correctly as possible and write neatly so their essays could be read by others. Once 30 min had elapsed, the test administrator told students they had 5 remaining minutes to complete their composition.

Before essays were scored, all papers were typed in order to reduce presentation bias in scoring due to poor handwriting legibility (Graham, Harris, & Hebert, 2011). Number of words written was determined via the word count option in Microsoft Word.

Writing quality was assessed with a traditional holistic writing scale (Huot, 1990). Raters were directed to read each paper attentively, but not laboriously, to obtain a general impression of overall writing quality. Papers were scored using a 9-point scale, with higher scores representing higher writing quality. To determine the score for an essay, examiners were told that persuasiveness, ideation, organization, aptness of word choice, grammar, and sentence structure should all

be taken into account when forming a single judgment about writing quality, and that no single factor should receive undue weight.

When scoring essays, raters were provided with a representative paper for a score of 2, 4, 6, and 8. These representative papers were from essays written by students in three fifth-grade classes that did not participate in this experiment. Students in these classes wrote a persuasive paper using the same writing prompts and materials applied in this investigation. Three former intermediate grade teachers read the papers written by these students and independently selected multiple possible anchor papers for the scoring points identified above. They then met, examined all papers selected, and identified a single paper to represent each score. These papers were then randomly ordered, and three other former teachers independently sorted them from highest to lowest writing quality. These three former teachers placed the four compositions in the same rank order established by the first three former teachers.

All compositions written by students in the Fall and Spring were independently scored by two trained raters unfamiliar with the design and purpose of the study. Before scoring, all identifying information was removed. The Pearson product moment correlation between raters' scores was 0.85. The score for writing quality was the average score of the two raters.

Writing knowledge

Writing knowledge was assessed with two different measures. One measure assessed students' knowledge about the assigned writing topic (e.g., saving money). Topic knowledge, was solicited with an open-ended question asking students to tell everything they knew about assigned topic. Students' written responses were divided into idea units. An idea unit was a specific, single, and unique idea in a student's response (Olinghouse & Graham, 2009). For instance, the response, "the internet is not very old" was scored as one idea unit; whereas, "the internet is not very old, but over 2 billion people use it" was scored as two idea units (i.e., 1: the internet is not very old, 2, but over 2 billion people use it). Repeated information was not considered as a new idea unit (e.g., "the internet is not very old…the internet started a little while ago" was scored as one idea unit). The score for each sample was the number of unique ideas a student produced about the topic.

Two trained raters unfamiliar with the design and purpose of the study independently scored topic knowledge responses, with one rater scoring all the responses and the second rater scoring a random sample of 25% of them. The Pearson product moment correlation between raters' scores was 0.97.

The second knowledge measure assessed students' knowledge about discourse markers in text (i.e., linking words in the Common Core State Standards CCSS, signaling words within text structure; Wijekumar, Meyer, & Lei, 2012) using a fill-in-the-blanks task. With this instrument, students were asked to supply 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 between 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).

This task was used in previous literacy studies (Meyer et al., 2010) and served as a measure of students' ability to detect and utilize comparison text structures in reading comprehension. Students who were able to correctly identify most of the discourse markers were able to select and encode logical memory structures while reading (based on writing effective main ideas) and scored significantly higher on standardized tests of reading comprehension.

In this research study, we used this measure as a proxy for knowledge about discourse markers for writing. These markers are pertinent for persuasive writing. For example, if a child wants to persuade the parent to get a dog as a pet, they may use comparison to compare a dog versus a cat and suggest that the dog would make a better pet due to its friendliness and energy levels. A child writing this essay using linking words that signify this comparison (e.g., in contrast to a cat, dogs are much more friendly) will score higher due to the organization of the writing activity.

We used two equivalent forms of the fill-in-the blank task. One form compared emperor and Adelie penguins, whereas the other form compared pygmy and emperor monkeys. Both passages contained 15 sentences, 96 idea units, and 128 words. Student responses were transcribed and scored using a computer algorithm that checked for the closeness of the answer to the ideal response. A score of 7 indicated an exact match to the missing cloze item; a score of 6 represented 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 delineated the use of 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 involved a similar comparison signaling word with a different intent (e.g., "smaller than" when "larger" fit the context); a score of 3 indicated a signaling word that was not a comparison signaling word (e.g., "solution" for "same as"); a score of 2 typified that a word 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.

Writing motivation

Students were asked to complete two self-report measures of motivation. One measure assessed students' attitude towards writing and it contained 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 used a five-point Likert-type scale (strongly disagree = 1; strongly agree = 5) to express their agreement with these items. The first two items were from an instrument developed by Bruning, Dempsey, Kauffman, McKim, and Zumbrunn (2013), whereas the remaining three items were adapted from a scale created by Graham, Berninger, & Fan (2007). Collectively, the items asked students about their attitude towards writing in and out of school. Factor analysis of the 5 writing attitude items, using an oblique rotation,

produced a single-factor solution (coefficient alpha was 0.89). The score for attitude towards writing was the average score for the five items.

The second self-report measure assessed students' self-efficacy for writing. This scale included 13 items: (1) I can spell my words correctly; (2) I can write complete sentences; (3) I can punctuate my sentences correctly; (4) I can write grammatically correct sentences; (5) I can begin my paragraphs in the right spots; (6) I can quickly think of the perfect word; (7) I can think of many ideas for my writing; (8) I can think of a lot of original ideas; (9) I know exactly where to place my ideas in my writing; (10) I can focus on my writing for at least 35 min; (11) I can avoid distractions while I write; (12) I can start writing assignments quickly; (13) I know when and where to use writing strategies. Students indicated their confidence for each task using 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 items on the self-efficacy scale were from an instrument developed by Brunning et al. (2013). One of these items was modified from "I can focus on my writing for at least 60 min" to "35" minutes, as this was 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"), as each of these tasks are important aspects of composing (Rijlaarsdam et al., 2012).

An unconstrained principal factor analysis, using responses of the students in this study, produced two factors with an eigenvalues greater than 1.0 (5.59 and 1.32), accounting for 53% of the variance. Consequently, we ran a forced two-factor solution (with an oblique rotation), Nine items (I-9 above), focusing on self-efficacy for writing mechanics, loaded on the first factor at 0.64 or higher (coefficient alpha was 0.77). Four items (10–13) loaded on the second factor (coefficient alpha was 0.88): self-efficacy writing regulation. The score for each factor was the average score for all items loading on it.

Strategic writing behaviors

Three measures assessed strategic writing behavior. Two measure were based on the plans students produced for their assigned essay. This included the number of words students wrote when planning (plans were typed and then scored using the word count option in Microsoft Word) as well as a score for the sophistication of students' plan. This second score involved rating students' plans using a 5 point scale. A score of 0 was assigned when no plan was created; a score of 1 if the plan was a verbatim copy of the composition, a score of 2 if the plan was a partially written essay that was extended in the essay, a score of 3 if students listed several phrases or ideas on the planning sheet, and a score of 4 if the students used a sophisticated planning strategy such as a web, outline, genre specific planning strategy, and so forth.

Two trained raters unfamiliar with the design and purpose of the study independently scored plans using the 5-point scale, with one rater scoring all plans and the second rater scoring a random sample of 25% of them. The Pearson product moment correlation between raters' scores was 0.88.

The third strategic writing measure, approach to writing, was a self-report instrument asking students to indicate their agreement with the following 10 items: (1) I give a lot of detail when writing; (2) I make sure my writing is organized and easy to follow; (3) I closely examine what the writing assignment calls for; (4) I start with a fairly detailed outline; (5) I use a lot of examples and definitions to make things clear in my writing; (6) I easily find good words for what I want to say when writing; (7) I plan out my writing and stick with the plan; (8) I keep my topic or theme clearly in mind as I write; (9) I use my time wisely when writing; (10) I think about my readers while I write. Students indicated agreement with each item using a five-point Likert-type scale (strongly disagree = 1; strongly agree = 5).

The items in the approach to writing measure were originally part of an assessment developed by Lavale, Smith, and O'Ryan (2002) for secondary students. Graham, Kiuhara, Harris, & Fishman (2017) adapted this scale for elementary grade students by reducing the number of items and rewriting them so they so that the language was more appropriate for children.

An exploratory factor analysis using the data from this study was computed to determine if the approach to writing measure was unidimensional as reported by Graham et al., (2017). An unconstrained principal factor analysis produced two factors with an eigenvalues greater than 1.0 (3.47 and 1.03), accounting for 45% of the variance. As a result, we ran a forced two-factor solution (with an oblique rotation), but three items double loaded at 0.45 on both factors (i.e., items 5, 6, and 10 above). We reran this analysis, and the remaining seven items all loaded at 0.52 or greater on a single factor. The score for approach to writing measure was the average score for the seven items (coefficient alpha = 0.73).

Writing skills

Students' persuasive essays provided a single measure of students' writing skills. Spelling proficiency was measured as the proportion of incorrectly spelled words in students' essays. Misspelled words included any real or made up words not spelled correctly in a student's paper. To obtain a proportion of incorrect spellings, the number of misspelled words was divided by the total number of words in the paper. Two trained raters unfamiliar with the design and purpose of the study scored students' papers. One rater scored all papers; the second rater scored a random sample of 25%. The Pearson product moment correlation between raters' scores was 0.98.

Reading

Reading comprehension was assessed using procedures designed by Meyer's and colleagues (Wijekumar et al., 2012). With this instrument, students read a comparison text (pygmy vs. emperor monkeys or emperor vs. Adelie penguins described earlier). The two passages were equivalent in terms of words and readability statistics and counterbalanced across testing periods. After reading a passage, the student was directed to write a main idea for it with the passage in view. They then wrote a full recall of the text without the passage in view. Student

responses were scored for both competence (main idea response) and quality (recall response), using a computer program that applied Meyer's (1985) propositional analysis. The scores range from 1 to 8 for reading competence (a score of 8 represented the highest score), with the assigned scores assessing 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. Comprehension quality was assessed on a 6 point scale, with scores ranging from no comparison of the animals (score = 1) to a comparison that compared the animals on superordinate features such as habitat, diet, or characteristics (score = 6).

Procedures

During the Fall and Spring assessments, measures were administered on a single day by members of the research team in the presence of the classroom teacher. All testers had received training to administrator each assessment until they could do so without error.

The instructions for each assessment was read to students by the test administrator. Students completed these measures in the following order: topic knowledge, persuasive writing from source task, strategic approach to writing scale, attitude towards writing scale, writing self-efficacy scale, knowledge of discourse markers, and reading comprehension measure. The topic knowledge measure preceded the persuasive writing measure, as reading the source material (i.e., article) for the writing task provided students with information about the assigned topic. The strategic approach to writing, attitude towards writing, and writing self-efficacy scales were administered after students wrote their persuasive essay, as the essay writing activity provided students with information on which to base their judgments.

Results

Means and standard deviations for all writing measures are presented in Table 1. There were improvements in students' planning and writing from the start to the end of the school year, as scores for writing quality, $F(1,163) = 168.96$, $p < .001$, eta = 0.51, number of words written, $F(1,163) = 25.92$, $p < .001$, eta = 0.14, writing plan complexity, $F(1,163) = 83.45$, $p < .001$, eta = 0.34, and planning words, $F(1,163) = 5.87$, $p = .02$, eta = 0.05, each showed statistically significant gains between the Fall and Spring assessments. There were no statistically significant gains, however, for approach to writing or for the measures assessing writing motivation, knowledge, and skills.

Correlations between measures

Correlations between measures are presented in Table 2. At both assessments points, the two outcome measures (writing quality and number of words) were moderately to strongly related with each other, sharing 35% of the variance in Fall

Table 1 Means and standard deviations for all writing measures at Fall and Spring

Measure	Fall M (*SD*)	Spring M (*SD*)
Writing quality	3.35 (1.11)	4.86 (1.55)
Number of words	90.53 (46.63)	116.53 (61.41)
Attitude towards writing	3.23 (1.22)	3.07 (1.26)
Self-efficacy mechanics	77.85 (18.10)	78.7 (18.66)
Self-efficacy self-regulation	70.60 (20.51)	69.94 (21.76)
Topic knowledge	7.95 (4.83)	7.14 (4.57)
Discourse marker knowledge	12.48 (6.75)	12.96 (7.98)
Spelling errors	0.86 (0.57)	0.86 (0.55)
Approach to writing	3.90 (0.64)	3.92 (0.62)
Planning complexity	2.44 (1.56)	3.82 (1.65)
Planning words	46.56 (39.95)	69.87 (42.24)

M mean, *SD* standard deviation; spelling errors is the proportion of words misspelled in student's paper

scores and 59% in Spring scores. During the Fall assessment, attitude towards writing, the two knowledge measures (topic and discourse marker knowledge), and number of words generated when planning were statistically related to writing quality and number of words written. In addition, number of words written in the Fall was statistically related to self-efficacy for self-regulation, approach to writing, and proportion of spelling errors (as expected this last association was negative). The relations between outcome measures and predictors shifted somewhat in the Spring, with both self-efficacy measures, all three strategic measures (i.e., approach to writing, planning complexity, and planning words), and the measure for knowledge of discourse markers all statistically related to writing quality and number of words written.

It is interesting to note that the motivation measure self-efficacy for self-regulation was statistically related to attitude towards writing and self-efficacy for mechanics in the Fall and Spring. Likewise, planning complexity and planning words were statistically associated at both testing points, but the approach to writing and planning complexity measures only reached statistical significance in the Spring. Further, the two knowledge measures (topic and discourse marker knowledge) were only statistically related during the Spring assessment. While the transcription skill, proportion of spelling errors, was negatively related to approach to writing in the Spring, it did not evidence a statistical association with any other predictor variable at either of the two testing points. Finally, statistical associations between predictors were small to moderate, with the largest association between approach to writing and self-efficacy for self-regulation in the Fall accounting for 59% of the shared variance between the two measures.

Do knowledge, motivation, strategic approach and skills predict writing performance?

For both the Fall and Spring assessments, we examined if writing knowledge (topic and discourse marker knowledge), motivation (attitude towards writing, self-

Table 2 Correlations of writing measures at Fall and Spring assessments

Measures	1	2	3	4	5	6	7	8	9	10	11
1. Writing quality	–	.77**	.01	.32**	.18*	.10	.48**	.15	.26**	.31**	.39**
2. Number of words	.59**	–	− .04	.26**	.16*	.15	.34**	.00	.24**	.27**	.42**
3. Attitude towards writing	.15*	.26**	–	.06	.33**	− .16	− .17*	.13	.37**	.04	.03
4. Self-efficacy mechanics	.14	.13	.09	–	.73**	.22**	.33**	− .05	.45**	.10	.10
5. Self-efficacy self-regulation	.14	.20**	.42**	.59**	–	.22**	.15	.05	.60**	− .05	.04
6. Topic knowledge	.42**	.39**	.18*	.10	.14	–	.22**	.07	.15	.06	.04
7. Discourse marker knowledge	.17*	.17*	.04	.18*	.10	.06	–	− .03	.09	.16*	.11
8. Spelling errors	− .08	− .21**	− .08	− .12	.13	.08	− .08	–	.02	.00	− .06
9. Approach to writing	.09	.17*	.46**	.57**	.68**	.09	.16*	− .21*	–	.10	.19*
10. Planning complexity	.09	.11	.10	.09	.10	.11	.14	− .13	.12	–	.43**
11. Planning words	.24**	.18**	.13	− .01	− .01	.06	.12	− .21*	.04	.54**	–

Fall scores are below the diagonal and Spring scores are above the diagonal; spelling errors is proportion of spelling errors

*$p < .05$; **$p < .001$

efficacy for mechanics, and self-efficacy for self-regulation), strategic approach (approach to writing, planning complexity, planning words), and skills (proportion of spelling errors) each made a unique and statistically significant contribution to predicting both the quality of students' text and number of words written. To determine if each set of predictors (e.g., the three strategic measures) made a unique contribution to predicting an outcome measure (writing quality or number of words), we examined if there was a statistically significant difference between two models: a full model containing all predictors and control variables (reading comprehension and school) and a reduced model containing all but the predictors of interest (e.g., the set of strategic measures).

To account for the nested data structure of students within classrooms, we applied a mixed model procedure (multilevel model or random effects model) using the SAS Mixed procedure. First, we estimated an unconditional model with random student and classroom components to assess degree of dependency due to the different levels. During the Fall assessment, intra-class correlations (ICC) at the classroom level were large to moderate for writing quality (.14) and number of words (.07). During the Spring assessment, they were large for both measures (writing quality = .24; number of words = .24).

To determine the unique effect of a set of predictors (e.g., the motivational measures), we used the likelihood ratio Chi square test by examining deviance change between the full model (as described above) versus a reduced model that did not contain the target predictors (e.g., the motivation measures). Pseudo R-squares were also computed 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. The full model (Model 1) included the control variables (the two reading measures and class) and all the predictor variables. This was compared to four reduced models, each specified by dropping from the full model a different set of predictors. Model 2 dropped writing knowledge measures (topic and discourse marker knowledge), Model 3 dropped writing motivation measures (attitude towards writing, self-efficacy for mechanics, self-efficacy for self-regulation), Model 4 dropped the writing skills measure (proportion of spelling errors), and Model 5 dropped the strategic writing measures (approach to writing, planning complexity, and planning words). As noted earlier, all predictors were measured at the same time point as the corresponding outcome measures. Because fixed effect estimates were of primary interest, we used maximum likelihood estimation for all models. The Mixed model results for Fall and Spring assessments for writing quality and number of words are presented in Tables 3, 4, 5 and 6.

Fall assessment

During Fall, the full model with all predictor variables (Model 1) explained 30% of the total variance in writing quality (see Table 3). Writing knowledge (Chi square = 28.4, $df = 2$, $p < .001$; Δ r-square = 0.16) and strategic writing behaviors (Chi square = 8.7, $df = 2$, $p = .04$; Δ r-square = 0.06) each made a statistically significant unique contribution, beyond the control variables and other predictor variables, in predicting the quality of students' writing. However, writing

Table 3 Multilevel models for writing quality for the Fall assessment

Fixed effects	Model 1 (n = 172)		Model 2 (n = 172)		Model 3 (n = 173)		Model 4 (n = 172)		Model 5 (n = 172)	
	Estimates	Std err	Estimates	Std err	Estimates	Std err	Estimates	Std err	Estimates	Std err
Intercept	− 0.49***	0.07	− 0.52***	0.09	− 0.49***	0.07	− 0.49***	0.07	− 0.49***	0.07
Reading—main idea competence	0.22	0.12	0.21	0.13	0.24*	0.12	0.21	0.12	0.20	0.12
Reading—main idea quality	− 0.21	0.18	− 0.18	0.20	− 0.23	0.18	− 0.19	0.18	− 0.14	0.19
Attitude towards writing	0.03	0.05	0.07	0.05			0.03	0.05	0.03	0.04
Self-efficacy mechanics	0.004	0.004	0.004	0.004			0.004	0.004	0.003	0.003
Self-efficacy self-regulation	0.002	0.003	0.004	0.004			0.002	0.003	0.000	0.003
Topic knowledge	0.06***	0.01			0.06***	0.01	0.05***	0.01	0.06***	0.01
Discourse marker knowledge	0.005	0.008			0.006	0.007	0.006	0.008	0.006	0.008
Spelling errors	− 0.06	0.05	− 0.02	0.05	− 0.06	0.05			− 0.07	0.05
Approach to writing	− 0.10	0.11	− 0.12	0.12	0.03	0.08	− 0.08	0.11		
Planning complexity	− 0.05	0.04	− 0.04	0.04	− 0.05	0.04	− 0.05	0.04		
Planning words	0.004**	0.001	0.004**	0.002	0.004**	0.001	0.004**	0.001		
Variance components										
Class	0.02	0.02	0.05	0.04	0.02	0.02	0.02	0.02	0.02	0.02
Residual	0.35***	0.04	0.40***	0.04	0.35***	0.04	0.35***	0.04	0.37***	0.04
Deviance	312.9		341.3		317.0		314.2		321.6	
Δ deviance (vs. M1)			28.4***		4.1		1.3		8.7*	
R-square	0.30		0.14		0.30		0.30		0.26	

All predictor variables are grand-mean centered. R-square = 1 − ((model level-1 variance + model level-2 variance)/(null model level-1 variance + null model level-2 variance))

*$p < .05$; **$p < .01$; ***$p < .001$

Table 4 Multilevel models for number of words written for the Fall assessment

Fixed effects	Model 1 (n = 172)		Model 2 (n = 172)		Model 3 (n = 173)		Model 4 (n = 172)		Model 5 (n = 172)	
	Estimates	Std err	Estimates	Std err	Estimates	Std err	Estimates	Std err	Estimates	Std err
Intercept	− 0.24**	0.06	− 0.28*	0.09	− 0.24**	0.06	− 0.23**	0.06	− 0.24**	0.06
Reading—main idea competence	0.10	0.14	0.12	0.15	0.12	0.14	0.07	0.15	0.09	0.14
Reading—main idea quality	− 0.11	0.22	− 0.11	0.24	− 0.12	0.22	− 0.06	0.22	− 0.07	0.22
Attitude towards writing	0.10	0.06	0.14*	0.06			0.10	0.06	0.10*	0.05
Self-efficacy mechanics	0.001	0.004	0.001	0.005			0.001	0.004	0.000	0.004
Self-efficacy self-regulation	0.002	0.004	0.004	0.004			0.002	0.004	0.001	0.004
Topic knowledge	0.06***	0.01			0.07***	0.01	0.06***	0.01	0.06***	0.01
Discourse marker knowledge	0.01	0.01			0.01	0.01	0.01	0.01	0.01	0.01
Spelling errors	− 0.17**	0.06	− 0.14*	0.06	− 0.17**	0.06			− 0.18**	0.06
Approach to writing	− 0.04	0.14	− 0.06	0.14	0.11	0.09	0.02	0.14		
Planning complexity	− 0.02	0.04	− 0.01	0.05	− 0.02	0.04	− 0.03	0.04		
Planning words	0.002	0.002	0.003	0.002	0.003	0.002	0.003	0.002		
Variance components										
Class	0	–	0.04	0.03	0	–	0	–	0	–
Residual	0.53***	0.06	0.59***	0.06	0.54***	0.06	0.55***	0.06	0.53***	0.06
Deviance	377.5		403.0		383.1		386.2		379.4	
Δ deviance (vs. M1)			25.5***		5.6		8.7**		1.9	
R-square	0.25		0.11		0.23		0.22		0.25	

All predictor variables are grand-mean centered. R-square = 1 − ((model level-1 variance + model level-2 variance)/(null model level-1 variance + null model level-2 variance))

*$p < .05$; **$p < .01$; ***$p < .001$

Table 5 Multilevel models for writing quality for the Spring assessment

Fixed effects	Model 1 (n = 143)		Model 2 (n = 144)		Model 3 (n = 143)		Model 4 (n = 143)		Model 5 (n = 148)	
	Estimates	std err	Estimates	Std err	Estimates	std err	Estimates	Std err	Estimates	Std err
Intercept	0.30	0.17	0.35	0.16	0.30	0.16	0.32	0.15	0.26	0.19
Reading—main idea competence	0.07	0.14	0.17	0.14	0.09	0.14	0.06	0.14	0.08	0.15
Reading—main idea quality	0.06	0.25	0.05	0.26	0.04	0.25	0.04	0.26	0.11	0.27
Attitude towards writing	0.02	0.05	− 0.003	0.05			− 0.004	0.05	0.07	0.05
Self-efficacy mechanics	0.001	0.005	0.01	0.01			0.004	0.005	0.002	0.005
Self-efficacy self-regulation	0.003	0.004	0.003	0.005			0.001	0.004	0.005	0.004
Topic knowledge	0.02	0.01			0.02	0.01	0.02	0.01	0.033*	0.014
Discourse marker knowledge	0.025**	0.009			0.03**	0.01	0.03**	0.01	0.02*	0.01
Spelling errors	− 0.18**	0.06	− 0.18**	0.06	− 0.17**	0.06			− 0.20**	0.06
Approach to writing	0.29*	0.12	0.32*	0.12	0.38***	0.09	0.33**	0.12		
Planning complexity	0.02	0.04	0.04	0.05	0.02	0.04	0.01	0.04		
Planning words	0.005***	0.001	0.005***	0.002	0.005***	0.001	0.006***	0.002		
Variance components										
Class	0.23*	0.13	0.22*	0.12	0.23*	0.12	0.18*	0.10	0.31*	0.16
Residual	0.36***	0.04	0.42***	0.05	0.37***	0.05	0.39***	0.05	0.45***	0.05
Deviance	281.5		302.0		283.5		290.8		323.0	
Δ deviance (vs. M1)			20.5***		2.0		9.3**		41.5***	
R-square	0.42		0.37		0.41		0.44		0.26	

All predictor variables are grand-mean centered. R-square = 1 − ((model level-1 variance + model level-2 variance)/(null model level-1 variance + null model level-2 variance))

*$p < .05$; **$p < .01$; ***$p < .001$

Table 6 Multilevel models for number of words written for the Spring assessment

Fixed effects	Model 1 (n = 143)		Model 2 (n = 144)		Model 3 (n = 143)		Model 4 (n = 143)		Model 5 (n = 148)	
	Estimates	Std err	Estimates	Std err	Estimates	Std err	Estimates	Std err	Estimates	Std err
Intercept	− 0.001	0.20	0.08	0.18	0.01	0.19	0.10	0.19	− 0.04	0.22
Reading—main idea competence	0.001	0.16	0.08	0.17	0.03	0.16	− 0.003	0.16	0.05	0.17
Reading—main idea quality	0.07	0.29	0.09	0.31	0.03	0.29	0.06	0.29	0.05	0.30
Attitude towards writing	0.02	0.06	− 0.01	0.06			0.01	0.06	0.05	0.06
Self-efficacy mechanics	− 0.001	0.006	0.005	0.006			0.000	0.006	− 0.001	0.006
Self-efficacy self-regulation	0.007	0.005	0.006	0.005			0.006	0.005	0.008	0.005
Topic knowledge	0.05***	0.01			0.05***	0.01	0.05**	0.01	0.06***	0.02
Discourse marker knowledge	0.02	0.01			0.02*	0.01	0.02*	0.01	0.02	0.01
Spelling errors	− 0.09	0.07	− 0.07	0.07	− 0.07	0.07			− 0.12	0.07
Approach to writing	0.17	0.14	0.23	0.15	0.33**	0.11	0.19	0.14		
Planning complexity	− 0.06	0.05	− 0.04	0.05	− 0.06	0.05	− 0.07	0.05		
Planning words	0.008***	0.002	0.008***	0.002	0.008***	0.002	0.008***	0.002		
Variance components										
Class	0.32*	0.17	0.25*	0.13	0.30*	0.15	0.30*	0.15	0.41*	0.21
Residual	0.50***	0.06	0.59***	0.07	0.52***	0.06	0.51***	0.06	0.57***	0.07
Deviance	328.7		349.7		332.2		330.5		360.4	
Δ deviance (vs. M1)			21.0***		3.5		1.8		31.7***	
R-square	0.26		0.25		0.26		0.27		0.12	

All predictor variables are grand-mean centered. R-square = 1 − ((model level-1 variance + model level-2 variance)/(null model level-1 variance + null model level-2 variance))

*$p < .05$; **$p < .01$; ***$p < .001$

skills, as measured by proportion of spelling errors, and writing motivation, as measured by attitude towards writing and two self-efficacy measures, did not make a unique contribution to predicting writing quality in the Fall. Only two specific measures, topic knowledge and planning words, accounted for a unique and statistically significant proportion of the variance once variance due to all other variables was controlled.

For number of words written on the Fall writing prompt (see Table 4), the full model with all predictor variables (Model 1) explained 25% of the total variance. Writing knowledge (Chi square = 25.5, $df = 2$, $p < .001$; Δ r-square = 0.14) and writing skills as measured by proportion of spelling errors (Chi square = 8.7, $df = 2$, $p < .01$; Δ r-square = − 0.03) both made a statistically significant unique contribution, beyond the control variables and other predictor variables, in predicting the number of words written by students. The sets of variables measuring writing motivation and writing strategies did not make a statistically significant unique contribution to predicting words written. Only two specific measures, topic knowledge and proportion of spelling errors, made a unique contribution to predicting number of words once variance due to all other variance was controlled.

Spring assessment

At Spring (see Table 5), the full model with all predictor variables (Model 1) explained 42% of the total variance. Writing knowledge (Chi square = 20.5, $df = 2$, $p < .001$; Δ r-square = 0.05), writing skills as measured by proportion of spelling errors (Chi square = 9.3, $df = 1$, $p < .01$; Δ r-square = − 0.02), and strategic writing behaviors (Chi square = 41.5, $df = 2$, $p < .001$; Δ r-square = 0.16) each made a statistically significant unique contribution, beyond the control variables and other predictors, in predicting writing quality. The set of measures assessing writing motivation did not make a statistically significant unique contribution to predicting writing quality in the Spring. Statistically unique contributions to predicting writing quality were made by four specific writing measures: discourse marker knowledge, proportion of spelling errors, approach to writing, and planning words.

For number of words (see Table 6), the full model with all predictor variables (Model 1) explained 26% of the total variance. Writing knowledge (Chi square = 21.0, $df = 2$, $p < .001$; Δ r-square = 0.01) and strategic writing approach (Chi square = 31.7, $df = 2$, $p < .001$; Δ r-square = 0.14) each made a statistically significant unique contribution, beyond the control variables and other predictor variables, in predicting number of words written. However, writing skills, as measured by proportion of spelling errors, and writing motivation, as measured by attitude towards writing and two self-efficacy measures, did not make a unique contribution to predicting number of words written in the Spring. Only two specific measures, topic knowledge and planning words, accounted for a unique and statistically significant proportion of the variance once variance due to all other variables was controlled.

Control variables

Reading comprehension as measured by the two control variables (main idea competence and main idea quality) did not make a statistically unique contribution to predicting writing quality or number of words for the Fall or Spring assessments. Mean performance on the competence measure in the Fall and Spring was 1.93 ($SD = 0.97$) and 2.06 ($SD = 0.90$), respectively. Mean performance on the quality measure was 1.60 ($SD = 0.63$) and 1.64 ($SD = 0.51$), respectively.

Discussion

Do writing knowledge, motivation, strategic behaviors, and skills predict writing performance?

Alexander's (1997) model of domain learning proposed that growth in an academic domain is shaped by cognitive and motivational catalysts. She argued that knowledge, motivation, and strategic behaviors collectively and uniquely promote students' growth as they move from initial acclimation in a domain to competence. The current study tested this proposition with writing, adding a fourth catalyst, writing skills, as proposed by Graham (2006).

At the first assessment point in the Fall, these four catalysts along with two measures of students' reading skills accounted for 30 and 25% of the variance in writing quality and number of words written, respectively. With one exception (writing motivation), our prediction that each catalyst would make a statistically unique contribution to predicting writing performance, after first controlling for variance due to all other variables, was supported. Writing knowledge uniquely predicted writing quality and number of words written in the Fall. Writing strategies uniquely predicted writing quality, whereas writing skills uniquely predicted number of words written.

The Spring assessment, which occurred after students were taught how to plan and write persuasive text, yielded somewhat similar findings. The four catalysts along with two measures of reading accounted for 42 and 26% of the variance in writing quality and number of words, respectively. In terms of writing quality, three catalysts, writing knowledge, strategic behaviors, and skills, accounted for unique variance in the quality of students' writing, whereas two of the catalysts, writing knowledge and strategic behaviors uniquely predicted number of words written. As in the Fall, writing motivation did not uniquely predict either writing outcome measure.

It is important to note that there were several notable differences between Fall and Spring outcomes in terms of statistically significant predictors. In the Fall, writing strategic behaviors did not make a unique contribution to predicting number of words, but it did in the Spring. This may be a consequence of the SRSD and web-based instruction students' received over the course of the school year in how to plan and draft persuasive text. These strategies may have helped students access

more relevant information from their own memory or the source material they read before writing. Additional research is needed, however, to verify this explanation.

We also found that writing skills (as exemplified through students' spelling errors) made a statistically unique contribution to predicting number of words in the Fall as well as writing quality in the Spring. The same statistically significant relationships were not obtained at the other assessment point. It is possible that writing skills became less important to predicting number of words written in the Spring, as the strategies for planning and drafting text students' acquired over the course of the school year mitigated the potentially negative effect of spelling errors. There was a statistically negative relation in the Fall between spelling errors and number of words written, but not in the Spring. This also raises the possibility that students' spelling improved during the school year, making spelling a positive factor in students' writing, especially in terms of quality. However, such an improvement was not observed, drawing into question the veracity of this explanation.

The most robust single predictor of writing performance in this study was one of the knowledge measures: topic knowledge. It uniquely predicted the number of words written in the Fall and Spring as well as writing quality in the Fall. While other researchers have found topic knowledge to be important to writing performance (Kellogg, 1987; McCuthchen, 1986; Olinghouse et al., 2015), our findings suggest that how much a student initially knows about a writing topic is still important even when they are provided with relevant topic information through source material. This prior knowledge likely helps them better access, select, and evaluate the information contained in source material as well as provide additional information for their writing.

The findings from the current investigation replicated in part the outcomes reported in Graham et al., (2018). In the previous study, all four catalysts uniquely predicted writing quality, number of words written, or both on the Fall and Spring assessments. The only exception involved writing knowledge in the Fall, which was only assessed with the discourse marker knowledge measure (topic knowledge was not assessed until Spring in the prior study). In the current study, three of the four catalysts, writing knowledge, strategic behavior, and skills, uniquely predicted one or both of the writing outcomes (i.e., quality and number of words) at both testing points. Notably, however, writing motivation did not account for unique variance in either writing outcome at either assessment point, even though the bivariate relationship between attitude towards writing and each writing outcome measure were statistically significant in the Fall, as were the bivariate relationships between the two self-efficacy measures and writing measures in the Spring.

While the current study and Graham et al., (2018) employed the same writing motivation measures, the self-efficacy measure in our study was best represented by two factors and not one as in the prior study. This may be responsible, at least in part, for the differences in the predictive value of writing motivation in the two studies, as bivariate relations between self-efficacy and the two writing outcomes (quality and number of words) were not only statistically significant at both time points in the prior investigation, but larger as well. Even so, we cannot rule out that differences in the writing tasks (science vs. social studies writing prompts), the participants themselves, or other factors were responsible for these differences.

Additional research is needed to replicate these findings and to determine which factors contribute to observed differences, assuming such differences are verified in the future.

While writing knowledge, strategic behaviors, and skills each accounted for unique variance in students' writing performance on a persuasive writing task involving source material in this study and the prior Graham et al., (2018), there were differences between the two studies on which catalysts uniquely predicted which writing outcomes at which assessment point. For instance, in the present study writing strategic behaviors predicted number of words, but this was not the case in the prior investigation. Likewise, writing knowledge and skills predicted number of words written in the Fall in this study, but not in the previous one. These differences may be due, at least in part, to differences in the assessments used in each study (e.g., this study added the measure planning words at both assessment points, and unlike the prior investigation included topic knowledge at the Fall assessments). Additional research is needed, however, to determine if differences in measures, writing prompts, students, or some other factor are responsible for these disparate findings across the two studies.

In summary, the present study provided partial support for Alexander's (2003) model as adapted by Graham (2006). Three of the four catalysts tested (writing knowledge, strategic behaviors, and skills) were statistically unique predictors of writing performance at the start and end of the school year. While this study did not provide evidence to support the unique role of one of Alexander's catalysts, motivation, this has been found to be a unique contributor to predicting writing performance in other investigations (Graham, 2006; Graham et al., 2018; Olinghouse & Graham, 2009; Olinghouse et al., 2015). Thus, the collective findings across this and prior studies support the view that a nuanced approach be used when applying Graham's (2006) adaptation of the Alexander's model of domain learning to writing, as the impact of specific catalysts likely depend on how a catalyst is defined and measured, the writing outcomes assessed, the capabilities of the writers, and the context of the study.

Limitations and recommendations for research

Limitations of the present study included it only involved fifth-grade students, a single type of writing was measured, it focused on writing in a single content area, and it did not extend beyond a single year. Consequently, the relationships tested in this investigation need to be examined with younger and older students, different forms of writing, in different academic domains, and over longer periods of time.

We also made specific decisions about which measures would be used to assess each catalysts as well as writing outcomes. As a result, replication and extension is needed to determine if the same pattern of findings are obtained when additional or even different measures are used to define each construct. This should include the use of behavioral measures in place of the self-report measures we applied for assessing efficacy, attitudes, and approach to writing. This may be especially important for self-report measures involving strategic behaviors during writing, as it may be difficult for some students, especially those who find writing challenging or

have limited knowledge of how to write, to adequately assess their capabilities in this area.

Lastly, caution must be exercised when drawing instructional implications from correlational studies like this one. 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, and strategic behaviors (Graham, Kiuhara, Mckeown, & Harris, 2012). Hopefully, future research will test instructional methods that are designed to bring these catalysts (including motivation) together in productive ways.

References

Alexander, P. (1997). Mapping the multidimensional nature of domain learning: The interplay of cognitive, motivational, and strategic forces. In M. Maehr & P. Pintrich (Eds.), *Advances in motivational achievement* (Vol. 10, pp. 213–250). Greenwich, CT: JAI.

Alexander, P. (1998). The nature of disciplinary and domain learning: The knowledge, interest, and strategic dimensions of learning from subject-matter text. In C. Hynd (Ed.), *Learning from text across conceptual domains* (pp. 55–76). Mahwah, NJ: Erlbaum.

Alexander, P. (2003). The development of expertise: The journey from acclimation to proficiency. *Educational Reearcher, 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). Guilford, NY: Guilford.

Bazerman, C., Berninger, V., Brandt, D., Graham, S., Langer, J., Murphy, S., et al. (in press). *The lifespan development of writing*. Urbana, IL: National Council of English.

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., & 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: The Guilford Press.

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: National Governors Association and Council of Chief School Officers. (2010). Downloaded from http://www.corestandards.org/.

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

Graham, S. (in press). 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*. Urbana, IL: National Council of English.

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., Harris, K. R., & Hebert, M. (2011). It is more than just the message: Analysis of presentation effects in scoring writing. *Focus on Exceptional Children*, *44*(4), 1–12.

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.

Graham, S., Wijekumar, K., Harris, K. R., Lei, P., Fishman, E., Ray, A., et al. (2018). *The relationship between writing skills, knowledge, motivation, and strategic behavior and writing performance with developing writers*. (**submitted**).

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

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

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

Kellogg, R. (1987). Effects of topic knowledge on the allocation of processing time and cognitive effort of writing processes. *Memory & Cognition, 15,* 256–266.

Lavale, 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 metacognitive control. *Written Communication, 5,* 306–324.

McCuthchen, D. (1986). Domain knowledge and linguistic knowledge in the development of writing ability. *Journal of Memory and Language, 25,* 431–444.

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.

Meyer, B. J. F., Wijekumar, K., Middlemiss, W., Higley, K., Lei, P., Meier, C., et al. (2010). Web-based tutoring of the structure strategy with or without elaborated feedback or choice for fifth- and seventh-grade readers. *Reading Research Quarterly, 41,* 62–92.

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

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.

Page, E. B., & Petersen, N. S. (1995). The computer moves into essay grading: Updating the ancient test. *Phi Delta Kappan, 76,* 561–566.

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., et al. (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.

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

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, NY: 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.