2019

Writing-to-Learn in Secondary Science Classes: For Whom is It Effective?

Katherine Landau Wright
Boise State University

Tracey S. Hodges
University of Alabama

Wendy K. Zimmer
Texas A&M University

Erin McTigue
University of Stavanger

This is an Accepted Manuscript of an article published by Routledge, an imprint of Taylor & Francis Group, in *Reading & Writing Quarterly* in 2019, available online at https://doi.org/10.1080/10573569.2018.1541769

Writing-to-Learn in Secondary Science Classes: For Whom is It Effective?

Katherine Landau Wright
Boise State University
katherinewright@boisestate.edu

Tracey S. Hodges
University of Alabama

Wendi K. Zimmer
Texas A&M University

Erin M. McTigue
University of Stavanger

In a modern, literate society, scientific thought does not progress without writing (Norris & Phillips, 2002). Professional communities primarily share outcomes and knowledge learned from research through writing. Therefore, learning to communicate effectively for scientific purposes is essential for preparing students for success in science, technology, engineering, and mathematics (STEM) fields. Yet, despite the importance of developing scientific literacy, reading and writing rarely occur in middle and high school science classes. Furthermore, the most commonly used literacy tasks in secondary education tend to focus on short answer responses or fill-in-the-blank questions (Drew, Olinghouse, Faggella-Luby, & Welsh, 2017). Such activities require very little higher-order thinking from students, conveying an implicit message that science is little more than a series of facts to be memorized (Hohenshell & Hand, 2006), which is in direct contradiction to the beliefs of the professional STEM community.

One method for incorporating literacy and higher-order thinking into science classes is through writing-to-learn interventions. Writing-to-learn can describe a wide variety of classroom strategies, but generally involves writing activities in content-area classes with a dual aim of supporting content knowledge acquisition and writing ability (Bangert-Drowns, Hurley, & Wilkinson, 2004; Miller, Scott, & McTigue, 2016). Bangert-Drowns and colleagues' (2004) meta-analysis demonstrated that writing-to-learn interventions have a small, positive impact on academic achievement. However, when teasing out these findings further, these researchers found that interventions for students in grades six through eight demonstrated much smaller effects (weighted Cohen's d = -.03), with four out of the six middle grade studies actually yielding negative effect sizes. The authors could only speculate as to *why* this effect was found because few studies have tested the same intervention strategies at different grade levels. One possible explanation is that students' motivation for writing may vary at different grade levels, and that writing tasks may have "detrimental motivational consequences for some students who found writing tedious and problematic" (Banger-Drowns et al., 2004, p. 37).

The present study seeks to further examine the middle school student disparity, identified by Bangert-Drowns and colleagues (2004), by using parallel intervention procedures at various grade levels. This research will allow us to examine whether age or learner level variables explains why some students benefit more from writing-to-learn opportunities in science class than others. Furthermore, as no two students are alike, this study also seeks to determine for whom writing-to-learn in science is most effective, as well as how affective factors (including students' motivation for writing) can influence the impact of these instructional strategies. As such, this study seeks to answer the following research questions:

1. How does the impact of a writing-to-learn intervention vary across grade levels?
2. For whom are writing-to-learn activities in science class most effective?

Literature Review

In 1988, James Britton argued, "when talking, reading, and writing are orchestrated in the classroom in such a way that each can make its unique contribution to a single end, we have surely harnessed language to learning as powerfully as possible" (The National Institute of Education, 1988, p. 6). While reading has taken the dominant role in literacy education, writing research has grown as a field in recent decades. In fact, some argue that writing instruction should take precedence over reading as the former can be accomplished without "conscious comprehension" (Dunn, 2000, p. 169) and the writing product provides unequivocal evidence that a literacy task has been completed (Konopak, Martin, & Martin, 1987). Unfortunately, the effects of writing on learning have been inconsistent, and reasons for this phenomenon remain unclear (Klein, 1999). While many educators and researchers alike would agree that writing can make a unique contribution to learning, how to harness that contribution is still debated.

Knowledge-Telling versus Knowledge-Transforming

This research is grounded in Bereiter and Scardamalia's (1987) concept of *knowledge telling* versus *knowledge transforming* writing. According to the authors, *knowledge telling* writing uses existing knowledge to report information, which constitutes most of the writing students do in science class (Choi, Notebaert, Diaz, & Hand, 2010). With this type of writing, the expectation is for students to restate correct information without additional commentary, synthesis, or analysis. Our goal, however, was to move students from *knowledge telling* to *knowledge transforming* writing, which involves using writing to engage in self-interactions to build new understandings (Bereiter & Scardamalia, 1987). This goal differs from prior writing research that focuses on a sociocognitive perspective of learning. That is, whereas sociocognitive theory posits that knowledge is built through social interactions with others (Vygotsky, 1980), the process of writing can serve as a self-interaction and help an individual organize, reformulate, and essentially transform existing knowledge into new concepts.

By using writing to transform knowledge, students are engaging in scientific literacy. Science literacy can be conceptualized under two distinct definitions: the *fundamental sense* and the *derived sense* (Norris & Phillips, 2002). The fundamental sense refers to an individual's ability to read and write when the subject is science, such as reading a scientific article and composing a response. In the derived sense, by contrast, science literacy refers to being knowledgeable and informed about science. However, a successful scientist's literacy skills must be strong in both domains, as there is a deep-seated relationship between scientific texts, literate thought, and scientific literacy (Norris & Phillips, 2002). We argue that this relationship implies that improving students' scientific literacy in the *fundamental sense*, and thus developing stronger readers and writers in science class, will make them better able to acquire the content necessary to become scientifically literate in the *derived sense.*

Integrating these two frameworks into scientific writing, we believe children learn more science content by focusing on knowledge-transforming writing to build both a fundamental and derived sense of science literacy. The outcome, theoretically, would be a population of young scientific thinkers who are able to share and communicate their knowledge. This study seeks to examine how knowledge-telling writing in science class can develop students' scientific literacy.

The Benefits of Writing in Science

According to the National Council of Teachers of English, writing "can both foster and demonstrate learning" in all disciplines through critical thinking and application of content specific knowledge (NCTE, 2008, p. 1). Such activities highlight the value of both knowledge and how professionals of the discipline think (Rainey & Moje, 2012). Science has its own process of knowledge creation and dissemination that require a unique discourse to function (McDermott & Hand, 2010). As explained by Hand, Park, and Suh (2018), "being literate in science [is] not simply a function of being able to replicate the language of science, but needed to incorporate the concept that understanding any science endeavor requires language" (p. 340). In essence, science is not possible without language (Norris & Phillips, 2003).

Overwhelmingly, research indicates that science instruction should integrate writing (Poe, Lerner, & Craig, 2010). Writing introduces students to the scientific process, and challenges them to reason and problem-solve, leading to the development of lasting science competencies (Otfinowski & Silva-Opps, 2015).

Furthermore, writing serves as a tool to develop students' conceptual understanding (McDermott & Hand, 2010). Writing-to-learn in science helps students make connections between familiar methods of learning, often received in literacy instruction, and more formal scientific discourse (Ritchie, Rigano, & Duane, 2008). Such connections possibly increase student participation, which lead to enhanced learning (Knipper & Duggan, 2006). Furthermore, when writing is applied to a specific discipline (e.g., science), ideas are strengthened by the process of building connections and describing concepts (Klein, 2006).

Review of Writing-to-Learn Research

While literature examining writing-to-learn dates back to well before Britton's early works in the 1970s, three recent reviews describe aspects of writing in contemporary classrooms. To begin, Bangert-Drowns and colleagues' (2004) meta-analysis examined the impact of writing-to-learn on academic achievement. This meta-analysis demonstrates that writing-to-learn interventions yield small, positive effect sizes (Bangert-Drowns, et al., 2004). However, as noted earlier, this effect was not consistent when examining middle-school participants.

Other researchers considered the characteristics of successful writing instruction. Graham and Perin's (2007) meta-analysis describes characteristics of writing interventions that tend to yield the greatest effect sizes. Specifically, Graham and Perin noted that 75% of studies using writing-to-learn strategies for content-area learning yielded small, positive effect sizes. While some interventions were more consistently successful than others, students required scaffolded and explicit instruction to develop lasting writing skills. Moreover, the authors acknowledge that they could not make recommendations about which students, such as low-performing writers, writing-to-learn strategies worked best for as this information was omitted from most studies.

Most recently, Miller and colleagues (2016) synthesized the literature on writing in content-area classes to provide practical classroom strategies for implementing writing instruction. Additionally, Miller and her colleagues fill a gap in previous reviews by included qualitative studies, which the previous two meta-analyses had to omit. When examining the context of writing to learn tasks, Miller and colleagues found that writing connected to specific learning goals was more effective than writing for the sake of a grade. Additionally, evidence-based tasks and prompts generally supported learning better than imaginative writing. Furthermore, students needed to engage in these types of writing tasks regularly for significant growth to be observed. Additionally, these researchers noted the prevalence of metacognitive writing tasks. When students engaged in metacognitive writing (such as describing what was understood or not understood from a day’s lesson), learning increased. However, how to consistently prompt metacognitive writing with younger learners remains unclear (Miller et al., 2016). This review indicates that children perform better when they write informationally about topics learned in class, aimed at helping them evaluate, synthesize, and reflect on their learning.

Taken together, the research suggests that instruction focused on developing metacognitive writing can support academic achievement. Furthermore, in science classes, interventions allowing children to write for authentic audiences and use evidence to form arguments also yield positive effects. However, these studies leave room to question the extent that writing-to-learn interventions can be successful with middle grade students and, when successful, which students benefit most. The purpose of this study is to utilize these research-based best practices for writing-to-learn in science classes at both the middle and high school levels. Doing so allows us to make direct comparisons between grades, as well as examine student-level variables that may moderate for whom writing-to-learn in science is effective.

Purpose

While the research on writing-to-learn in science shows the importance of integrating writing and science, a gap still exists. First, research has not emphasized which writing-to-learn methods help students of different abilities. For instance, in other areas of education, motivation has been positively correlated with academic achievement (Vecchione, Alessandr, & Marsicano, 2014). Therefore, a students' motivation for writing may moderate the effectiveness of a writing-to-learn intervention. However, the field has not fully explored the relationship between motivation for writing and the impact of a writing-to-learn intervention.

Furthermore, research has not specifically compared metacognitive writing prompts to argumentative writing prompts and how they impact student writing and science learning. In the present study, we culminate the recommendations from three prior reviews on writing-to-learn (Bangert-Drowns et al., 2004; Graham & Perin, 2007; Miller et al., 2016) to develop a writing-to-learn intervention in science class that uses metacognitive and argumentative writing prompts and identifies for whom writing-to-learn interventions is effective.

Methods

This study utilized a mixed-methods quasi-experimental intervention design to describe for whom writing-to-learn in science is effective.

Intervention Design

Using established best-practices for writing in science classes, author (2017) developed classroom procedures aimed at providing a minimally-intrusive method for science teachers to incorporate writing-to-learn into their classes. Author (2017) tasks students with two types of writing (described below) at select intervals throughout the intervention. The intervention took place over eight weeks. We began by administering assessments to understand participants' pre-intervention writing skills, writing motivation, and science background knowledge. Following, students completed 10 short metacognitive writing assignments and three longer argumentative writing assignments.

Metacognitive Writing. The metacognitive writing took place twice a week at the end of a class period. We provided a template and asked students to describe what they had learned that day to a friend who was absent from class. The goal of this task was to prompt metacognitive writing (that is, writing about one's own learning and thought process) while providing an authentic audience (a peer who needs to know what they missed). These metacognitive assignments formed the basis of the intervention, as they provided students opportunities to practice writing about scientific content. Furthermore, the teachers made instructional decisions based upon the writing, addressing areas of confusion and re-teaching when necessary.

Argumentative Writing. The longer argumentative writing occurred three times throughout the intervention – once after the first week, once at the midpoint, and once at the end of the intervention. This assignment varied slightly for each class, depending upon the content taught at that time; however, the general format remained the same to allow for pre- and post-intervention and grade level comparisons. Students were provided an inaccurate scientific statement published in a mock-journal, and were asked to write a "letter to the editor" describing what was incorrect and providing the correct information (See Figure 1). These argumentative assignments provided students further opportunities to practice writing, engage in higher-order thinking, and acquire content knowledge. To specify the unique content for different classes, we collaborated with the teachers to ensure the information reflected the instruction.

Figure 1 about here

Participants

We included two sets of participants in this study: two teachers and their 54 students in grades six through 11. Mrs. James (all names are pseudonyms), taught grades six and seven and 10th grade chemistry. Mr. Devin taught 11th grade physics. Demographic information for the student participants is included in Table 1. Information and permission slips were sent home to all parents. A total of 54 students (representing an 86% response rate of the possible participants) returned signed permission slips to participate in the intervention.

Table 1 about here

Pre-Intervention Assessments

We administered three pre-intervention assessments to better understand the dispositions, knowledge, and skills our participants brought to the study.

Motivation for Writing. We administered the *Self-beliefs, Writing beliefs, and Attitude Survey* (SWAS; Author, 2018), which has been validated for this age group, to the students at the beginning and end of the intervention. The Cronbach's α reliability estimations for these administrations were .947 and .937, respectively. While the two administrations did not yield any statistically significant differences, we averaged the student scores for analysis. Averaging the scores increases the overall validity by lessening the impact of one-time measurement error.

Scientific Knowledge. An author's knowledge of content greatly influences the overall quality of writing (Graham, 2006). Furthermore, writing without content knowledge will not contribute to learning (Willingham, 2007). Therefore, we created a measure of general scientific content to evaluate students' level of background knowledge in science. Our measure was based upon multiple choice test questions from released versions of the Trends in International

Mathematics and Science Study (IEA, 2013) and the National Assessment of Educational Progress (NAEP, 2014). We reviewed recent tests and selected questions that required students to apply their knowledge or use scientific reasoning (higher-order thinking), rather than simply repeating content information. Prior to administering these measures, two certified science educators (independent of the study) reviewed the items to ensure the questions were all valid and pertinent to relevant scientific information. Additionally, the selected questions were designed to target the knowledge of 4th, 8th, and 12th grade students. As our participants were in grades six through 11, this variety of questions helped prevent ceiling and floor effects.

Writing Skills. Individuals are more motivated and engaged in a task when they believe they can be successful at that task (Eccles & Wigfield, 2002). Therefore, we hypothesized that students' general writing skills, prior to the intervention, would impact their engagement in the activities. For the purpose of this study, general writing skills refer to adherence to English conventions, organization of writing, and writing style, which are necessary for effective written communication in all genres.

We used the spontaneous writing subtest (form A) of the fourth edition of the Test of Oral Written Language (TOWL) (Hammill & Larsen, 2009) to assess students' overall writing ability. This subtest presents a picture and allocates twenty minutes for students to write a story based upon the visual. The writing samples are scored for contextual conventions and story composition, and these scores make up the spontaneous writing composite score.

Fidelity Measures

One researcher or a research assistant attended each class period where writing took place. The researcher assisted the teacher in classroom activities (such as handing out papers) so the students quickly became accustomed to the presence of the additional adult in the room. At the end of each class period, the researcher completed a fidelity form that recorded general classroom activities, student engagement during class and writing time, and amount of time allotted for student writing. Additionally, we took field notes and recorded any unusual disturbances or activities.

As the study progressed, teachers asked to make slight adjustments to the procedures to better adapt the intervention to their classroom. As the goal of this study was to examine how this intervention would work in a real classroom setting, teachers were encouraged to "take charge" and modify the writing to fit the needs of their students. These small changes, such as using an argumentative writing task in place of a scheduled quiz, allowed the teachers to take ownership of the intervention. We explained the principles supporting the intervention, and ensured that the writing assignments still adhere to those best practices identified through previous research (i.e., using evidence to form arguments; writing for authentic audiences; multiple opportunities to write; metacognitive writing).

Outcome Variable

We used the students' argumentative writing assignments as our primary outcome variable. Using the *Rubric for Scientific Writing* (RSW; Authors, 2016), an independent science educator and a literacy educator evaluated each argumentative writing assignment, and we averaged their ratings to calculate individual scores. The *RSW* measures six aspects of quality scientific writing: (1) Claim/Warrant, (2) Evidence/Support, (3) Analysis of Content, (4) Organization, (5) Audience, and (6) Presentation of Writing. These six aspects yield three composite variables: Overall Quality, Scientific Rhetoric, and English Composition. Scientific writing skills are understood to be the ability to make an effective scientific argument in writing.

We examined students' growth from the first to the third argumentative writing assignments to describe the impact of the intervention on students' writing.

Data Analysis

We first conducted ANOVAs to compare students by grade level. Next, we conducted a cluster analysis to better describe for whom the writing-to-learn intervention was most effective.

Grade-level Comparisons. We first examined the data by grade level. We combined the students in grades six and seven to form a "middle school" group, and likewise combined the older students to form a "high school" group. We then conducted ANOVAs using the two groups' scores on the argumentative writing assignments as dependent variables to examine the impact of grade level. We included students in this analysis if they completed the first and last argumentative writing assignments. Two students were absent for either the first or last argumentative writing assignment, resulting in a sample size of 52 students.

Cluster Analysis. As one aim of this study was to determine *for whom* writing-to-learn interventions in science are most effective, we also conducted a cluster analysis (Everitt, Landau, Leese, & Stahl, 2011). Cluster analysis is a variable-centered methodology that connects participant scores to create profiles. The purpose of a cluster analysis is to organize "cases" (in this instance, students) by features so the resulting clusters exhibit high internal homogeneity. Cluster analysis is descriptive and atheoretical, however we can use group membership to describe the cases using statistical analyses such as ANOVA (Hair & Black, 2000). Cluster analysis has been used in previous literacy research for purposes such as describing different approaches to shared book readings (e.g., Haden, Reese, & Fivush, 1996; Hammett, Van Kleeck, & Huberty, 2003), and identifying how writing traits can predict other variables (Glogger, Schwonke, Holzapfel, Nuckles, & Renkl, 2012; Roid, 1994).

We used the pre-measures from the SWAS, General Science Knowledge test scores, and the students' scaled scores on the TOWL as cluster variables (See Table 2 for list of variables and rationale). Again, two students did not complete either the first or the last argumentative writing task, and were not included in this analysis. Furthermore, one student transferred to the school just after the intervention began, and never completed all pre-assessments, and three others changed classes and were unable to finish the post-intervention assessments. Therefore, 48 students were included in the cluster analysis.

Table 2 about here

We also examined the metacognitive writing samples to develop student profiles. While all students were provided the same prompts, how they responded varied greatly. Therefore, we coded these writing samples for the following: (1) number of words in each writing samples, (2) number of activities listed, (3) key scientific vocabulary used, and (4) number of scientific facts explained. We also noted whether students created some sort of visual, graphic, or formula to represent the information (see table 3).

Table 3 about here

Results

In the following section, we first present the results of our grade-level analysis. Next, we detail the results of our cluster analysis and detail how the intervention had different impacts on five unique groups of students.

Writing Achievement by Grade Level

We first ran one-way ANOVAs examining the students' scaled TOWL scores to confirm that no statistically significant differences existed between the middle and high school students' writing abilities prior to the intervention (Writing Conventions [$F(2, 50) = 1.152$, $p = .288$]; Story composition [$F(2, 50) = 0.389$, $p = .535$]). Next, we conducted an ANOVA to determine if differences existed between the middle and high school students' scores on the argumentative writing assignments. These results demonstrate no significant differences in writing achievement based upon grade (see supplemental online materials).

Because statistical significance is impacted by sample size (Thompson, 2006), we also examined effect sizes describing the growth from the first to the third argumentative writing tasks (see Table 4). These results demonstrated moderate to strong growth in scientific writing by all students. Convergent with previous literature, the high school students demonstrated more growth; however, this comparison does not mean that the intervention was ineffective for middle school students. Therefore, we conducted a cluster analysis to examine other characteristics that may explain for whom writing-to-learn in science class is most effective.

Table 4 about here

Cluster Analysis

We conducted a hierarchical cluster analysis to describe the similarities between the participants in the writing-to-learn intervention. This process is similar to creating a factor analysis, but participants are grouped by person based on the squared Euclidian distances between their variable scores. All variables were converted to z-scores to ensure one did not have undue influence due to scale (Meyers, Gamst, & Guarino, 2013). We used the Ward method and squared Euclidian distance because the Ward method has strong discriminating power (Hammett et al., 2003), and these methods have been used in other writing-to-learn intervention studies (see Glogger et al, 2012). The squared Euclidian distances between cases ranged from 3.42 to 93.61.

We examined four different models, respectively fitting the data into three, four, five, and six different profiles. For each model, we conducted ANOVAs to identify group differences in the outcome variables. The model of best fit contained five clusters and provided the most explanation for why some students achieved more than others.

Cluster Descriptions. General descriptive information about the participants in each cluster is detailed in Table 5.

Table 5 about here

Table 6 details the cluster z-scores on each variable. All group differences were significant at the .01 level. We have labeled each cluster based upon the groups' pre-intervention measure scores as well as how they distinguished themselves during the intervention.

Table 6 about here

Cluster Outcomes. We conducted one-way ANOVAs to examine differences between the group means on the argumentative writing assignments. Again, we examined the cluster's average scores on each factor of the *RSW* (i.e., Scientific Rhetoric, English Composition, and Overall Quality). No statistical differences existed between groups on any of these measures (See supplementary materials). However, because of the small sample sizes, we again calculated effect sizes for each groups' growth from the first to the third writing assignment (see Table 7).

Table 7 about here

Cluster Descriptions. In the following section, we summarize each cluster's pre-intervention skills, behavior during the intervention, and outcomes.

Cluster 1: Activities. Cluster 1 contained the most students (31%), and the majority of the participants in this cluster were middle school students. These students began the intervention with slightly higher than average scores on the SWAS and TOWL, indicating they had strong writing skills and were motivated writers. During the intervention, the majority of their writing included descriptions of activities, with fewer instances of vocabulary and facts than many of their peers. The teachers noticed this trend and encouraged students to write about what they *learned* rather than what they *did*. Despite this extra support, these students continued to describe classroom activities. Therefore, this is the *Activities* cluster.

These students' overall Scientific Rhetoric writing scores demonstrated a small, positive effect size (d = 0.24). By contrast, their English composition scores dropped (d = - 0.74). However, as we did not provide scaffolding for this aspect of writing, growth was not expected. Overall, their writing scores yielded an effect size of -0.10, suggesting the intervention had minimal impact on this group of students.

Cluster 2: Strong Scientific Knowledge. The second cluster of 11 students, including one ELL student, demonstrated the highest scores on the General Science Knowledge assessment and SWAS, indicating these students were motivated to write and had significant science background knowledge. This was especially evident during the intervention, when these students included the most vocabulary and facts. Together, these findings led this group to be named the *Strong Scientific Knowledge* cluster. This groups' overall growth yielded an effect size of d = 0.15, with modest but promising growth in scientific rhetoric (d = 0.22). These results indicate that the intervention supported the learning of all students, including high achievers.

Cluster 3: Visuals. Seven participants, including one ELL student, had near average scores on all pre-intervention measures. This group was composed of all high school students in the physics classes. While they wrote fewer words than their peers, they created the most visuals, nearly two standard deviations above the mean. Therefore, this cluster is called *Visuals.*

This group's English composition scores demonstrated moderate growth (d = 0.57) – a surprise considering the lack of instruction and the fact that these students produced some of the shortest writing samples. However, their scientific rhetoric scores revealed strong growth (d = 0.98), indicating they improved in their ability to convey scientific concepts in writing.

Cluster 4: Avoiders. The nine students in Cluster 4 began the intervention with lower writing motivation than most of their peers. Additionally, these students demonstrated some of the lowest scores on the TOWL, indicating that writing may be an area of difficulty. This group, made up of an equal mix of high school and middle school students, wrote the least number of words during the intervention, and used the fewest scientific vocabulary or facts in their writing. This group is thus called *Avoiders*.

Despite the concerning pre-intervention scores, this group of students made the most growth during the intervention, with their scientific argumentation score growth yielding an effect size of 1.47. Interestingly, this group also showed a moderate growth (d = 0.57) for English conventions. This finding indicates that the writing intervention may be especially promising for students who are struggling with writing.

Cluster 5: Low Motivation. The final cluster, consisting of three middle school students and four high school students, scored the lowest on the SWAS and TOWL, indicating they had the most negative feelings towards writing and likely found writing to be a difficult task. This group also demonstrated slightly lower than average scientific background knowledge, and their writing produced negative z-scores on all measured areas.

Although they wrote less than many of their peers during the metacognitive writing tasks, this group demonstrated growth in all aspects of writing. The strongest effect size was in English composition (d = 0.67). Scientific rhetoric also demonstrated moderate growth (d = 0.55), indicating that the intervention supported their ability to express scientific knowledge in writing.

Discussion

Previous findings have indicated that students in middle school were less likely than high school students to benefit from writing-to-learn interventions (Bangert-Drowns et al., 2004). However, few researchers have attempted to implement the same intervention in both middle and high school classes. Therefore, the purpose of the present study was to examine the effects of a minimally-intrusive intervention on students' scientific writing skills at both the middle and high school levels. Furthermore, through cluster analysis, we can gain a more nuanced understanding of why (or why not) adolescent learners respond to a writing-to-learn intervention.

How does the impact of a writing-to-learn intervention vary across grade levels?

Examining the results by grade level revealed that there were no statistically significant differences between the performance of the middle and high school students. This finding contrasts existing literature, which generally demonstrates that writing-to-learn interventions are less effective for middle grade students (Bangert-Drowns et al., 2004). Rather than grade level, in the present study, the impact of the writing-to-learn intervention varied depending upon students' pre-intervention skills and beliefs towards writing along with *what* they actually did during the intervention. Therefore, we turn our attention to the clusters that emerged to determine for whom writing-to-learn interventions may be most effective.

For whom are writing-to-learn interventions in science class most effective?

All five clusters of students demonstrated growth in the area of scientific rhetoric (Cohen's d range .22 to 1.47). This finding alone indicates that writing-to-learn activities have a positive impact on science literacy. This is especially noteworthy for the students in the *Avoiders* and *Low Motivation* clusters, groups that both began the intervention with lower than average writing skills and motivation. This finding indicates the intervention procedures supported all students' abilities to write and communicate scientifically, even those who may be difficult for teachers to engage in typical instructional practices. This finding aligns with previous research; writing can increase engagement, extend critical thinking, and enhance the meaning-making process (Knipper & Duggan, 2006).

Additionally, while all students demonstrated growth, the students for whom the intervention was most effective were in the *Visuals* and *Avoiders* clusters. Both of these clusters demonstrated particularly strong growth in scientific rhetoric (d = 0.98 and 1.46, respectively). As the *Visuals* students were all enrolled in the physics class, there may be an innate connection between writing, physics content, and visual representations of knowledge. During the intervention, the students studied waves, and their teacher encouraged them to create mental images of the waves during class discussions. Therefore, their visuals may be a clear reflection of the scientific thinking promoted by their teacher. Graphical representations of information are common in K-12 science texts (Authors, 2010) and research has demonstrated that older students are more likely to consider graphics when reading science materials (Authors, 2009). In that way, their writing is starting to reflect the disciplinary expectations of the field.

The *Avoiders* cluster began with some of the lowest pre-intervention scores and, overall, produced the least amount of writing. The majority of their writing represented descriptions of class activities, with some graphics to supplement the writing. However, these students made the most growth. Perhaps the requirement to produce written work prompts these students to engage more than they would with a receptive task. Additionally, even if students were simply recalling activities from class, writing provided an opportunity to reflect on those activities and possibly extend learning. As described earlier, writing product provides unambiguous evidence that a literacy task has been completed (Konopak, Martin, & Martin, 1987). Together, these findings suggest that the intervention may be most effective for students struggling with writing.

By contrast, not all groups made growth in the area of English Composition; in fact, one group (*Activities*) demonstrated a negative effect size in this area (Cohen's d = -0.74). However, this finding is not altogether surprising as there was no instruction in this area during the intervention. While research has demonstrated that students require multiple opportunities to write (Bangert-Drowns et al., 2004), students also require feedback and instruction in order for writing to improve (see Graham & Perin, 2007). Certain aspects of writing considered valuable in science may hold less value within English composition, due to disciplinary expectations (Shanahan & Shanahan, 2008). For instance, literature demands the use of imagery through language, whereas science demands the use of actual images.

In summary, our study indicates that a minimally-intrusive writing intervention in science can help students develop stronger writing skills and deeper scientific rhetoric. Specifically, we found that by incorporating visual representations into their writing, students improved in their scientific rhetoric. We also found that students who struggle with writing may benefit the most from multiple, varied opportunities to write across disciplines.

Considerations for Future Research and Practice

Writing-to-learn interventions provide one method to integrate literacy and higher-order thinking into science classes. Future research must examine whether this apparent effect on students' fundamental scientific literacy will extend to students' derived scientific literacy. That is, it is still not clear whether students' ability to engage in scientific argumentation (as demonstrated in their writing) correlates to their science content knowledge acquisition. As Graham and Perin's (2007) findings indicate, however, skilled writers rely on their subject knowledge; as a logical extension, when a learner's disciplinary writing skills increase, it may also reflect greater subject knowledge.

Beyond content knowledge, a second aspect of skillful disciplinary writing is writer's use of genre specific text conventions. These results provide evidence that some students gained genre-specific text conventions in this short intervention – most notably as related to the use of visuals. While reading, scientists tend to be particularly focused on the different representations of an idea (e.g., comparing prose to diagrams), and the reader processes these representations recursively as the pages turn (Shanahan & Shanahan, 2008). When considered simultaneously, these two modes of representation allow scientists to construct a mental model of a scientific phenomenon. However, visual literacy in science has put far more focus on the reading (rather than creation) of visuals (e.g., Authors, under review). Future research should investigate the relationship between students' use of graphics while reading science texts, the graphics they produce while writing scientifically, their understanding of science concepts, and the overall quality of their scientific writing.

Limitations

The greatest limitation of this study is that we had no control group. As we sought to examine how the effectiveness of writing-to-learn varies by grade level, we purposefully recruited teachers who taught multiple grades. However, this meant our teachers did not have multiple sections of the same grade, which would have allowed for a comparison group. To isolate the effects of a writing intervention, future researchers should emphasize an experimental design. Second, we were unable to collect longitudinal data that would have allowed us to describe the connection between intervention growth and academic achievement over time. Furthermore, our sample was relatively small and from one geographic location. Expanding this work to larger samples (allowing for control and intervention groups) and longitudinal data collection will strengthen the generalizability of findings.

Conclusion

While many agree that writing can make a unique contribution to learning, it has been difficult to harness that contribution, particularly in the middles grades. We approached this challenge through a lens of pragmatism, meaning that we focused on outcomes and shared meaning-making (between teachers and researchers) for the creation of practical solution to this educational problem. Specifically, the purpose of this study was to measure the efficacy of a research-based, *feasible* writing-to-learn intervention in grade six through 11 science classes, for different types of adolescent learners. Research has identified a barriers preventing writing integration in content-area classes, so we prioritized intervention features that allow for ease of replication and adaptation by classroom teachers. Finally, we focused on middle school students because in previous research this group has been least responsive to writing interventions.

Our results indicate that through exposure to relatively brief writing tasks, students' ability to engage in scientific rhetoric improved. Most notably, students with low writing motivation and science knowledge showed the largest response to the intervention. These findings suggest that writing-to-learn strategies can be implemented strategically in middle and high school science classes to support scientific literacy development.

Furthermore, with minimal training the science teachers were able to take ownership of the intervention strategies and adapt them to fit their classroom practices. We described the principles behind the writing tasks (i.e., using evidence to form arguments, writing for authentic audiences, providing multiple opportunities to write, and using prompts that encourage metacognition), and therefore any modifications did not interfere with the overall effectiveness of the intervention. Unfortunately, many practitioner publications that make recommendations for supporting science literacy are largely atheoretical (Authors, 2015). However, these findings indicate that when content teachers are provided with theoretical foundations about literacy integration, they can make modifications that support the needs of their students while preserving the integrity of instructional practices. In a world with rapidly advancing science and technology, our schools need innovative and pragmatic approaches that foster collaboration between science and literacy, to educate a scientifically literate population who can critically analyze and communicate in science.

References

Authors (under review, 2009, 2010, 2015, 2016, 2017, 2018) *Information removed for blind review*

American Association for the Advancement of Science. (2011). *Vision and change in undergraduate biology education: A call to action*. Retrieved from http://visionandchange.org/files/2011/03/Revised-Vision-and-Change-Final-Report.pdf

Bangert-Drowns, R. L., Hurley, M. M., & Wilkinson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis. *Review of Educational Research, 74*(1), 29–58. doi: 10.3102/00346543074001029

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

Choi, A., Notebaert, A., Diaz, J., & Hand, B. (2010). Examining arguments generated by year 5, 7, and 10 students in science classrooms. *Research in Science Education, 40*(2), 149–169. doi: 10.1007/s11165-008-9105-x

Drew, S. V., Olinghouse, N., Faggella-Luby, M., & Welsh, M. (2017). Framework for disciplinary writing in science grades 6-12: A national survey. *Journal of Educational Psychology, 109*(7). 935-955. doi: 10.1037/edu0000186

Dunn, J. E. (2000). Reading in the content areas: Unlocking the secrets and making them work. *Journal of Adolescent & Adult Literacy, 44*(2), 168–170.

Eccles, J. S., & Wigfield, A. (2002). Motivational, beliefs, values, and goals. *Annual Review of Psychology, 53*(1), 109-132. doi: 10.1146/annurev.psych.53.100901.135153

Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis* (5th ed.). Chinchester, England: Wiley.

Glogger, I., Schwonke, R., Holzapfel, L., Nuckles, M., & Renkl, A. (2012). Learning strategies assessed by journal writing: Prediction of learning outcomes by quantity, quality, and combinations of learning strategies. *Journal of Educational Psychology, 104*(2), 452–468. doi: 10.1037/a0026683

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

Graham, S., & Perin, D. (2007). What we know, what we still need to know: Teaching adolescents to write. *Scientific Studies of Reading, 11*(4), 313–335. doi: 10.1080/10888430701530664

Haden, C. A., Reese, E., & Rivush, R. (1996). Mothers' extratextual comments during storybook reading: Stylistic differences over time and across texts. *Discourse Processes, 21*(2), 135-169.

Hair, J. F., & Black, W. C. (2000). *Reading and understanding more multivariate statistics*. Washington, D.C.: American Psychological Association.

Hammett, L. A., van Kleeck, A., & Huberty, C. J. (2003). Patterns of parents' extratextual interactions during book sharing with preschool children: A cluster analysis study. *Reading Research Quarterly, 38*(4), 442-438. URL: https://www.jstor.org/stable/4151836

Hammill, D. D., & Larsen, S. C. (2009). *Test of Written Language – Fourth Edition* [Measurement Instrument].

Hand, B., Park, S., & Suh, J. K. (2018). Examining teachers' shifting epistemic orientations in improving students' scientific literacy through adoption of the Science Writing Heuristic approach. In K. S. Tang & K. Danielsson (eds) *Global developments in literacy research for science education* (pp. 339-355). New York, NY: Springer.

Hohenshell, L. M., & Hand, B. (2006). Writing-to-learn strategies in secondary school cell biology: A mixed methods study. *International Journal of Science Education, 28*(2–3), 261–289. doi: 10.1080/09500690500336965

IEA International Association for the Evaluation of Educational Achievement. (2013). *Trends in international mathematics and science study*. Boston, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

Klein, P. D. (1999). Reopening inquiry into cognitive processes in writing-to-learn. *Educational Psychology Review, 11*(3), 203–270. doi: 1040-726X/99/0900-0203$16.00/0

Klein, P. D. (2006). The challenges of scientific literacy: From the viewpoint of second-generation cognitive science. *International Journal of Science Education, 28*(2-3), 143-178. doi: 10.1080/09500690500336627

Knipper, K. J., & Duggan, T. J. (2006). Writing to learn across the curriculum: Tools for comprehension in content area classes. *The Reading Teacher, 59*(5), 462-470. doi: 10.1598/RT.59.5.5

Konopak, B. C., Martin, M. A., & Martin, S. H. (1987). Reading and writing: Aids to learning in the content areas. *Journal of Reading, 31*(2), 109–115. URL: https://www.jstor.org/stable/40029822

McDermott, M. A., & Hand, B. (2010). A secondary reanalysis of student perceptions of non-traditional writing tasks over a ten year period. *Journal of Research in Science Teaching, 47*(5), 518-539. doi: 10.1002/tea.20350

Miller, D. M., Scott, C. E., & McTigue, E. M. (2016). Writing in the secondary-level disciplines: A systematic review of context, cognition, and content. *Educational Psychology Review, 30*(1), 83-120. doi: 10.1007/s10648-016-9393-z

Meyers, L. S., Gamst, G., & Guarino, A. J. (2013). *Applied multivariate research: Design and interpretation.* Thousand Oaks, CA: Sage

National Assessment of Educational Progress. (2014). *General information about The Nation's Report Card: Sample questions, science*. National Center for Education Statistics. Retrieved from http://nces.ed.gov/nationsreportcard/subject/field_pubs/sqb/pdf/2104_sq_s_g4_g12.pdf

National Council of Teachers of English. (2008). *Writing now: A policy research brief produced by the National Council of Teachers of English*. Retrieved from http://www.ncte.org/library/NCTEFiles/Resources/PolicyResearch/WrtgResearchBrief.pdf

The National Institute of Education. (1988). *Writing and Reading in the Classroom* (J. Britton) (Tech. Rep. No. _425). Washington, DC: Center for the Study of Reading.

Norris, S., & Phillips, L. M. (2002). How literacy in its fundamental sense is central to scientific literacy. *Science Education, 87*(2), 224–240. doi: 10.1002/sce.10066

Otfinowski, R., & Silva-Opps, M. (2015). Writing toward a scientific identity: Shifting from prescriptive to reflective writing in undergraduate biology. *Journal of College Science Teaching, 45*(2), 19-23. doi: 10.2505/4/jcst15_045_02_19

Poe, M., Lerner, N., & Craig, J. (2010). *Learning to communicate in science and engineering: Case studies from MIT*. Cambridge, MA: MIT Press.

Rainey, E., & Moje, E. B., (2012). Building insider knowledge: Teaching students to read, write, and think within ELA and across the disciplines. *English Education, 45*(1), 71-90. URL: https://www.jstor.org/stable/23365001

Ritchie, S., Rigano, D., & Duane, A. (2008). Writing an ecological mystery in class: Merging genres and learning science. *International Journal of Science Education, 30*(2), 143-166. doi: 10.1080/09500690601161783

Roid, G. H. (1994). Patterns of writing skills derived from cluster analysis of direct-writing assessments. *Applied Measurement in Education, 7*(2), 159-120. doi: 10.1207/s15324818ame0702_4

Shanahan, T., & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: Rethinking content-area literacy. *Harvard Educational Review, 78*(1), 40–59. doi: 10.17763/haer.78.1.v62444321p602101

Thompson, B. (2006). *Foundations of behavioral statistics: An insight-based approach.* New York: Guilford.

Vecchione, M., Alessandri, G., & Marsicano, G. (2014). Academic motivation predicts educational attainment: Does gender make a difference? *Learning and Individual Differences, 32*, 142-131. doi: 101016/jlindif.2014.01.003.

Vygotsky, L. S. (1980). *Mind in Society: The development of higher psychological processes.* Boston, MA: Harvard University Press.

Willingham, D. T. (2007). Critical thinking: Why is it so hard to teach? *American Educator*, *109*(4), 21–29. doi: 10.3200/AEPR.109.4.21-32

Yiğit, E., & Gökpinar, F. (2010). A simulation study on tests for one-way ANOVA under the unequal variance assumption. *Communication Science University of Ankara Series, 59*(20), 15-34.

Supplemental Online Materials

ANOVA Results for pre-intervention scaled writing scores (TOWL)

				Middle School (*n* = 26)		**High School (*n* = 28)**	
	df	*F*	*Sig.*	*M*	*SD*	*M*	*SD*
Writing Conventions	53	1.152	.288	80.37	23.16	74.08	18.65
Story Composition	53	0.389	.535	79.44	30.62	73.84	34.14

ANOVA Results for Long Writing Scores by Grade Level

Long Writing #1							
				Middle School (*n* = 25)		**High School (*n* = 26)**	
	df	*F*	*Sig.*	*M*	*SD*	*M*	*SD*
Scientific Rhetoric	50	1.096	.300	1.48	.71	1.68	.67
English Composition	50	.827	.368	1.59	.70	1.75	.52
Overall Score	50	1.088	.302	1.54	.68	1.72	.56
Long Writing #2							
				Middle School (*n* = 26)		**High School (*n* = 20)**	
	df	*F*	*Sig.*	*M*	*SD*	*M*	*SD*
Scientific Rhetoric	45	1.599	.213	2.18	.60	2.40	.56
English Composition	45	1.259	.268	1.99	.53	2.18	.56
Overall Score	45	1.581	.215	2.09	.54	2.29	.53
Long Writing #3							
				Middle School (*n* = 23)		**High School (*n* = 27)**	
	df	*F*	*Sig.*	*M*	*SD*	*M*	*SD*
Scientific Rhetoric	49	3.952	.053	1.80	.51	2.09	.51
English Composition	49	1.620	.209	1.66	.54	1.83	.41
Overall Score	49	2.971	.091	1.73	.50	1.96	.44

ANOVA Results for Long Writing Scores by Cluster

				Cluster 1: Activities		**Cluster 2: Strong Scientific Knowledge**		**Cluster 3: Visuals**		**Cluster 4: Avoiders**		**Cluster 5: Low Motivation**	
Long Writing #1													
				(*n* = 13)		(*n* = 11)		(*n* = 7)		(*n* = 7)		(*n* = 7)	
	df	*F*	*Sig.*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Scientific Rhetoric	44	1.811	.146	1.60	.45	1.94	.85	1.47	.56	1.09	.60	1.55	.77
English Comp.*	44	2.021	.110	1.82	.23	1.93	.79	1.74	.44	1.29	.50	1.39	.73
Overall Score*	44	1.544	.208	1.68	.27	1.93	.81	.161	.46	1.27	.55	1.47	.74
Long Writing #2													
				(*n* = 15)		(*n* = 10)		(*n* = 4)		(*n* = 6)		(*n* = 6)	
	df	*F*	*Sig.*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Scientific Rhetoric	40	1.089	.377	2.23	.67	2.53	.47	2.29	.76	2.03	.59	1.97	.49
English Comp.	40	1.483	.228	2.05	.62	2.37	.44	2.04	.59	1.80	.41	1.80	.51
Overall Score	40	1.418	.248	2.14	.62	2.45	.38	2.17	.67	1.92	.48	1.89	.49
Long Writing #3													
				(*n* = 13)		(*n* = 11)		(*n* = 7)		(*n* = 8)		(*n* = 7)	
	df	*F*	*Sig.*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Scientific Rhetoric	45	.808	.527	1.73	.53	2.08	.45	2.06	.64	1.88	.46	1.94	.63
English Comp.	45	1.682	.173	1.55	.48	1.97	.44	1.71	.40	1.55	.38	1.83	.57
Overall Score	45	1.184	.332	1.64	.49	2.03	.44	1.89	.49	1.72	.34	1.89	.59

**Homogeneity of variance assumption not met (Levene’s statistic p = .002). Thus, we conducted a Welch’s ANOVA (Yiğit & Gökpinar, 2010).*

Tables

Table 1.

Participant Demographic Information

Class	**Teacher**	**Content**	***N***	**% Female**	**% ELL**
6th Grade	Mrs. James	General Science	18	33	0
7th Grade	Mrs. James	General Science	8	62	0
10th Grade	Mrs. James	Chemistry	14	50	28
11th Grade	Mr. Devin	Physics	14	21	7

Table 2

Variables Included in Cluster Analysis

Variable	**Rationale for Inclusion**
Average SWAS Motivation Score	Student motivation for writing will impact how he or she approaches writing tasks.
Average TOWL	Students' overall writing ability would impact performance during the intervention.
General Science Knowledge Scores (Pre-test)	It can be assumed that students with more background knowledge would feel more confident in scientific writing.
Number of activities*	While all students were provided similar writing tasks, how they responded to those tasks differed greatly. These measures help describe what students did during the intervention.
Number of scientific vocabulary*	
Number of scientific facts*	
Number of words*	
Percent of writing samples that include a visual	

* These items represent average counts across the 10 metacognitive writing samples

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Reading & Writing Quarterly*, published by Routledge. Copyright restrictions may apply. https://doi.org/10.1080/10573569.2018.1541769.

Table 3

Metacognitive Writing Coding

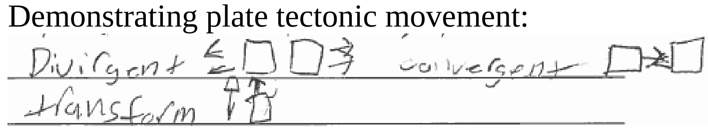
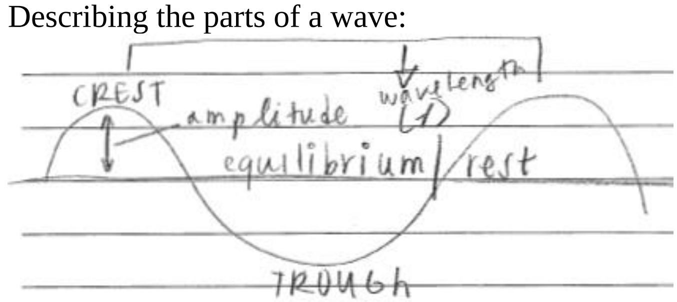
Code	**Description**	**Student Examples**
Activity	Student summarizes an activity done in class. Generally includes an active verb.	*Today in science we did a project on Chapter 4* *We were paired into groups to start a lab* *Today we went over covalent bonds.*
Scientific Vocabulary	Words or short phrases used to describe the course content.	*I learned that density is the amount of matter per unit of volume.* *Natural selection is a big part of an organism's life.* *We learned about ionic compounds, their formulas, and how their names changed when ionic bonding occurs.*
Scientific Facts	Course content paraphrased in students' words. Individual facts were identified if provided sufficient information to create a typical test question.	*Gregor Mendal used cross-pollination in plants and learned about dominant and recessive traits.* *The water cycle is also called the hydro-cycle, because hydro means water.* *With a pendulum, it doesn't matter how high you hold the string before you let it swing, because it will swing at the same speed.*
Visuals	Tables, charts, graphics, or formulas used to represent course content. These were coded as either included (1) or not included (0).	Demonstrating plate tectonic movement:  Describing the parts of a wave: 

Table 4

Cohen's d Effect Sizes for Argumentative Writing Scores (First to third writing assignments)

	Middle School	**High School**
Scientific Rhetoric	.52	.68
English Composition	.11	.17
Overall Score	.32	.48

Table 5

Participant Descriptives by Clusters

	Cluster 1	**Cluster 2**	**Cluster 3**	**Cluster 4**	**Cluster 5**
N	15	11	7	8	7
% Female	60	55	29	0	14
% High School	13.3	63.6	100	50	57.1
Mean Grade (SD)	6.73 (1.38)	9.00 (2.05)	10.71 (0.76)	8.38 (2.56)	8.35 (2.18)
% ELL	0	9	14	0	14

Note: ELL (English Language Learner) refers to students learning English as a new language

Table 6

Cluster Analysis In-put Variables (z-scores)

	Cluster 1: Activities	**Cluster 2: Strong Scientific Knowledge**	**Cluster 3: Visuals**	**Cluster 4: Avoiders**	**Cluster 5: Low Motivation**
SWAS	.52 (.76)	.59 (.68)	-.03 (.63)	-.29 (1.12)	-1.18 (.61)
TOWL	.69 (.62)	.36 (.97)	.09 (.28)	-.58 (.58)	-1.31 (.56)
GSK Score	-.51 (1.14)	.64 (.62)	-.07 (.78)	.55 (.54)	-.41 (1.28)
Activities	.54 (1.01)	-.17 (.82)	-.79 (.85)	.30 (.57)	-.59 (1.16)
Scientific Vocabulary	-.22 (.69)	1.21 (.77)	-.53 (.51)	-.93 (.44)	-.10 (.79)
Scientific Facts	-.09 (.44)	1.08 (.69)	-.88 (.43)	-1.01 (.59)	-.31 (.68)
Words	.07 (.60)	1.11 (.85)	-.83 (.50)	-.90 (.53)	-.14 (.98)
Visuals	-.56 (.22)	.30 (1.05)	1.91 (.52)	-.35 (.48)	-.40 (.40)

Note. Standard deviations are displayed next to means in parenthesis

Table 7

Cohen's d Effect sizes for Argumentative Writing Scores by Cluster (First to third writing assignments)

	Cluster 1: Activities	**Cluster 2: Strong Scientific Knowledge**	**Cluster 3: Visuals**	**Cluster 4: Avoiders**	**Cluster 5: Low Motivation**
Scientific Rhetoric	0.24	0.22	0.98	1.47	0.55
English Composition	-0.74	.06	0.57	0.58	0.67
Overall Score	-0.10	0.15	0.59	0.98	0.61