Exploring the Influence of Science Writing Instruction on Fourth Graders' Writing Development

Jennifer L. Tilson, Jill Castek, and Megan Goss
Lawrence Hall of Science, University of California, Berkeley

This study examined fourth grade students' informational writing development in the context of an integrated science-literacy unit. The purpose of this study was to investigate the possibility that teaching writing as an authentic part of science can benefit students in their ability to organize and express scientific ideas. An experimental group received an integrated science-literacy unit that included scaffolded instruction on science writing, while a control group received science-only instruction on the same topic during the same time period. All students completed an open-ended writing prompt as a pre/post assessment. This assessment was analyzed using a specially designed rubric that included seven dimensions: science content, use of evidence, introduction, clarity, conclusion, vocabulary usage, and a count of the number of science-specific vocabulary words used. Writing from students in the experimental group were compared with the control group to allow the researchers to examine student outcomes to determine to what extent elements of informational writing were acquired by the students. Results indicated that students in the experimental condition outperformed students in the control condition on all but two of the dimensions of science writing (conclusion and vocabulary usage).

Nature of the Problem

In order to address the pressing global challenges that face us, educators need to cultivate students who are highly literate problem solvers. Academic and workplace achievement is increasingly dependent on the ability to write, read, and speak in sophisticated ways (Moje, 2007; Friedman, 2005; Alexander & Jetton, 2002; Bransford, Brown, & Cocking, 2004). However, current efforts to increase students' literacy achievement are falling short. This is particularly evident in content areas such as science that require a high degree of domain-specific literacy. For instance, one indicator, the Trends in International Mathematics and Science Study (TIMSS) reveals that students in the U.S. score below their international peers (Martin, Mullis, Gonzalez, & Chrostowski, 2004). Similar patterns of low achievement among U.S. students have been documented over the past several years (National Center for Education Statistics, 2006, 2007). According to recent research (Greenleaf & Hinchman, 2009; Carnegie Council on Advancing Adolescent Literacy [CCAAL], 2010) the way to address these concerns is to infuse literacy into the content area classroom.

Scientific literacy goes beyond familiarity with a corpus of information. Developing domain-specific knowledge requires a firm grasp of the skills needed to communicate about the ideas that are central to conceptual understanding. In science, paramount among these skills are the ability to reason, to interpret and evaluate information, and to make evidence-based explanations (Driver, Newton, & Osborne, 2000; Lemke, 1990). Writing is a key practice that scientists engage in in order to communicate their findings and explanations with the scientific community and to advance the field (Osborne, 2002). However, in science classes at all levels, writing is often practiced rather than explicitly taught. We contend that providing elementary students with explicit, scaffolded instruction in informational writing will increase their achievement in both science and literacy and move us closer to the goal of creating more scientifically literate citizens. Findings (e.g. Morrow, Pressley, Smith, & Smith, 1997) are beginning to suggest that teaching informational writing in elementary school is an essential part of jump-starting these efforts.

Examining students' writing is an especially important area for continued research for several reasons. First, recent studies have called attention to writing as the 'Neglected R'. Successive reports from the National Commission on Writing in America's Schools and Colleges (2003) asserts that writing does not get nearly as much money or attention as reading and math do, and that we pay for this neglect as a society, since much of the success in later schooling and in work is contingent upon solid writing skills. Furthermore, Bereiter & Scardamalia (2003) suggest that writing enhances content learning because it allows students to call on higher forms of thinking in order to process and express ideas: in other words, they have to think deeply about the content when they write about it. Moreover, writing may be a central mode through which students connect the dots in their knowledge and make sense of what they are learning (The National Commission on Writing, 2003). As a result, informational writing must become a central focus of content area instruction in classrooms (Cutler & Graham, 2008). Finally, despite the many effective approaches for teaching writing, the instruction and practice of writing in schools is increasingly shortchanged, especially in science. Recent research (e.g. Cervetti, Pearson, Bravo, & Barber, 2006) has shown that integrated science-literacy instruction increases students' proficiency with both content knowledge as well as literacy processes. Additional evidence suggests that students who learn to write well within one content area, such as science, may ultimately gain an understanding of the common characteristics of informational writing in general and will be better able to effectively communicate important ideas across content areas (Holliday, Yore, & Alvermann, 1994). Therefore, teaching writing in the context of science seems a promising way to help students gain the skills needed to communicate complex ideas in writing.

Theoretical Framework

Like many educators and educational researchers, (e.g. Draper & Siebert, 2009; Guthrie, Anderson, Alao, & Rinehart, 1999; Moje, Ciechanowski, Kramer, Ellis, Carrillo, & Collazo, (2004), Shanahan & Shanahan, 2008; Vitale & Romance, 2007) we see literacy as a tool for learning and recognize there is a reciprocal relationship between literacy development and content knowledge. Pearson (2009) points out that to make literacy instructional most purposeful students need to read, write, and talk about some content. For example, in a science unit on light, if a student acquires knowledge about light interactions, she can more capably express her thoughts in writing and speaking, and better understand what she reads. This stance - recognizing literacy as the vehicle by which students will accomplish learning about a particular domain - grounds our thinking about reading instruction and writing instruction. Recent work by Lee & Spratley (2010) and Schoenbach & Greenleaf (2009) adopts a similar orientation and recognizes the important role that content knowledge plays in literacy development.

Literature Review

Early work from the 1980s and 1990s examined the overlapping cognitive demands required to be successful in both science and literacy (Carin & Sund, 1985; Holiday, Yore & Alvermann, 1994). Since that time, educators in both fields have been acknowledging, on one hand, the critical role that language plays in science, and, on the other, the central importance of content-area knowledge in supporting the development of literacy. For instance, Palinscar and Magnussen (2001) developed and tested a curriculum in which texts were designed to support inquiry by providing data and experiences that were difficult to gain firsthand in the classroom. They found that this approach enhanced the students' understandings of difficult science concepts. Romance and Vitale (1992, 2001) as well as Guthrie and colleagues (1999, 2002) have found that integration of science and literacy offers significant gains for students in both domains.

Much of the attention of earlier work on science and literacy integration focused on the connection between reading and scientific reasoning in particular, although students' writing development, as an adjunct outcome, was considered as well. Some researchers have also examined the use of student science notebooks as a means of developing science content knowledge (e.g. Palinscar and Magnusson, 2001). Findings from this study showed that student science notebooks were a key component that could be used to link students' experiences in reading and classroom science experiences. When writing was made into a central component linking reading and investigations, students gained greater access to science content. However, this study did not examine students' writing development over the course of the intervention, nor were written products scored on how well students expressed and organized their ideas. Some researchers have focused on writing as a both vehicle for promoting content understanding as well as a skill set that students can develop. Notably, Klentschy and Molina-De La Torre (2004) found that students' use of science notebooks as an authentic extension of inquiry improved their informational writing in a variety of important ways and also showed benefits for content learning.

A more focused area of research on science writing addresses the need for students to be able to understand the specific writing demands of a given content area as well as their need to be able to reproduce content-area writing models in their own writing (Cervetti, Pearson, Barber & Bravo, 2006; Daniels, Zemelman, & Steineke, 2007; Pappas and Varelas, 2004). Involving students in writing science texts can heighten their awareness of how such texts are created (Littlefair, 1992) and increase their understanding of why and when to use them (Moss, 2005). Over time, students develop the ability to read such texts as writers and begin to use them to search for evidence they can use in their own writing. In this way, they become empowered to create their own texts based upon what they have gathered from these models (Pappas, Kiefer, & Levstik, 1999). Krajcik & McNeill (2007) have focused on one important science writing genre – the explanation. These researchers found that a scaffolded approach to teaching students to write scientific explanations led to students' increased ability to make claims, cite relevant supporting evidence, and express their scientific understanding in written form.

However, this work was focused on middle school students, and centered, as others have, on the analysis of science content through students' writing products.

Although studies of science-literacy integration have shown promising results in terms of student learning outcomes (Cervetti, Pearson, Barber & Bravo, 2006; Guthrie et al 1999; Guthrie & Ozgungar, 2002; Palinscar & Magnusson, 2001; Romance and Vitale, 1992, 2001; Wang & Herman, 2005), essential questions remain about the effectiveness of various elements of integrated instruction. One important question in this area is whether students' writing development can improve when learning to write is a part of science instruction (Alvermann, Swafford, & Montero, 2004). Our work on an integrated science-literacy curriculum (e.g. Cervetti et al, 2006) seems to suggest that learning the written discourse of science is key in enabling students to understand and express scientific understandings. Thus, we aim in this study to answer the question: *Do students who are taught scientific writing in the context of an integrated science-literacy unit make greater gains in writing than students who are taught science separate from literacy?* To this end, the research reported in this paper examines in detail pre/post scores from an assessment of science writing administered to fourth grade students.

Methods

Participants

Fourth grade teachers were recruited to participate in the study. Ninety-four 4th grade classrooms in 48 schools volunteered to participate (n=2144). Each classroom was located in the same Southern state, and was comprised of near equal numbers of boys and girls. In our sample, 47 classrooms were in the experimental condition (n=215) that taught the integrated science and literacy unit, and 47 classrooms were in the control condition (n=241) that taught a content comparable unit. Demographic information about student participants was only available for 45% of the sample. Below, we include school demographics, available student demographics, and percentages of missing data in order to characterize the population involved in the study as well as possible.

Experimental group. At the school level, we examined socio-economic designations that were available from the Great Schools nation-wide database. A classroom was considered to come from a low-SES school if more than 50% of the students at that school were economically disadvantaged. Using that criteria, 24 of the classrooms in the experimental group came from low-SES schools and 17 of the classrooms did not. For six of the classrooms, information describing the level of economics was not available. Of the 47% if students whose demographics were available, 22% were White, 18% were African American and 4% were Hispanic. The remaining 3% of students were Asian (1.6%) and Multiracial (1.4%). Approximately 4% of students were English language learners.

Control group. At the school level, we examined socio-economic designations that were available from Great Schools nation-wide database. A classroom was considered to come from a low-SES school if more than 50% of the students at that school were economically disadvantaged. Using that criteria, 41 of the classrooms in the control group came from low-SES schools and 8 of the classrooms did not. For five of the classrooms, information describing the level of economics was not available. Of the 44% if students whose demographics were available, 18% were White, 18% were African American and 4% were Hispanic. The remaining 4% of students were Asian (2.7%), Multiracial (1%), and American Indian (.3%). About one percent of students were English language learners.

Equivalency of groups. A comparison the student demographics indicated that the two groups were composed of near-equal amounts of ethnically and linguistically diverse students, but that there were there were fewer numbers of low-SES students in the experimental group and fewer English language learners in the control group. To determine the extent to which these indicators may have been an indication of pre-existing differences among the groups, we examined two standardized student achievement measures. The first was a state-wide criterion-referenced test of English Language Arts (CRCT ELA) that was administered to all 4th graders in 2007, just prior to the beginning of the study. An Independent-samples T-test was run to compare the means scores for each group. The results indicated there were no statistical differences between the groups: t (400) =.86, $p < .39$. The second measure was a state-wide criterion-referenced test of general science content that was administered in all 4th graders in 2007. The results indicated that there was a statistically significant mean difference favoring the control group t (646) =2.37, $p < .02$. Table 1.1 includes the descriptive statistics for the criterion-referenced tests and Table 1.2 contains the mean comparison statistics.

Table 1.1

Descriptive statistics summarizing pre-existing differences between groups on criterion reference tests

	Group	N	Mean	Std. Deviation	Std. Error of the Mean
CRCT ELA 2007	Experimental	180	820.86	25.68	1.91
	Control	222	823.06	25.68	1.72
CRCT Science 2007	Experimental	317	813.75	35.25	1.98
	Control	331	820.33	35.56	1.96

Table 1.2

Independent Samples (t-test) mean comparisons between groups on criterion reference tests

	t-test for Equality of Means						
	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
						Lower	Upper
CRCT ELA 2007	.855	400	.39	2.20	2.58	-2.86	7.27
CRCT Science 2007	2.37	646	*.02	6.58	2.78	1.12	12.05

* p < .05

Instruction in the Experimental Condition

Students in the experimental group were taught an integrated science-literacy curriculum composed of forty 60-minute sessions designed to build students' science knowledge about physical science concepts related to light and energy. This curriculum was in the form of a detailed teacher's guide, student books, and a kit of materials mailed to teachers participating in the study. Topics in the unit were the characteristics of light, how light interacts with materials (such as reflecting or refracting), light as a form of energy, and energy transformations. The unit included a balance of both science- and literacy-focused sessions: approximately 40% of instructional time was spent on literacy activities (reading, writing, and listening/speaking), 40% on hands-on, inquiry science (including discussions of data and findings from students' investigations), as well as about 20% reflection and assessment.

Instruction was designed so that text and experience were mutually supportive of student learning -- students investigated essential concepts through both text and hands-on inquiry experiences. For example, students investigated reflection by testing a set of materials such as foil, felt, and wood to see if they reflected light from a flashlight. They discussed their findings and made a table displaying their data. They then referred to a text, which provided additional data about the materials. Students used the data in the book to resolve disagreements about which materials reflect light, as well as to draw inferences about the amount of light reflected by classes of materials (such as fabric or metal). In this sequence of lessons, students used reading, writing, and oral language skills in support of learning both science content and the discourse of science. As they gathered more information about which materials reflect light, students debated and revised their findings, and finally wrote explanations about idea that all materials reflect light, using the evidence they had amassed. Throughout the integrated science-literacy unit, students were explicitly instructed on using evidence to support their explanations and arguments. This was approached in both oral and written form. Students were asked questions throughout the unit meant to prompt the use of supporting evidence (such as, "Why do you think that?"). Students learned that evidence can come from a variety of sources, including firsthand investigations as well as texts.

Writing instruction in the unit employed a gradual release of responsibility model, with the teacher introducing various aspects of scientific writing each time students were asked to write. At the beginning of the unit, the teacher modeled writing an entire piece with student input. As the unit progressed, there was less teacher modeling and more student independence. Scaffolds such as graphic organizers and partnered discussion before writing were utilized throughout the unit to help students focus on essential concepts, organize information, and generate ideas. In addition to several sessions devoted entirely to writing instruction, students had a number of opportunities to write for various purposes. Throughout the sessions, students wrote to record information and data from their investigations, to take notes, to respond to informational texts they read, and to prepare for small-group discussions.

Instruction in the Control Condition

Students in the control condition were taught similar science concepts as students in the experimental group, but in a non-integrated fashion. During the period of the study, teachers who had been assigned to the control condition taught students concepts about light and energy as expressed in the state science standards, but in ways that were familiar to them. Information about the instruction in these classrooms was collected through a survey asking teachers to describe the content they taught and the methods they used to teach the content. Survey results indicated that teachers in the control condition used various materials and approaches to teach about light and energy including textbooks, published inquiry-science units, and teacher-made units involving hands-on science experiences.

Assessments Administered in Experimental and Control Conditions

The writing assessment was administered as part of a battery of pre/post measures that accompanied the unit. All assessments were administered to both the experimental and the control classrooms. The prompt was as follows: *How does light interact with materials? Give three examples.* This prompt was focused on essential science content taught in both the experimental and control conditions. We hypothesized that students who received focused writing instruction in conjunction with science would be more proficient in developing their ideas in writing, including evidence to support them, organizing their writing, using the language of science with a higher degree of precision, and in expressing science content more accurately. In addition, we posited that students in the experimental group would use a greater amount of science vocabulary words accurately in their writing than students in the control condition.

Rubric Development

Development of the writing rubric involved several stages of collaborative work. The first stage involved reading 65 randomly selected student papers. This first reading was used to determine which aspects of writing were central to an expertly crafted response. Those characteristics were then classified into five dimensions including evidence, introduction, clarity, conclusion, and language use. Assessing a sixth dimension, a vocabulary count of unit specific science words was also discussed. As we read and discussed papers, aspects that made a response well-written were described in the highest score point of the rubric. Developing the other score point descriptions involved a systematic, iterative process of reading papers, detailing their characteristics, evaluating writing against the descriptions, and making adjustments. The use of evidence dimension measured how well the evidence the student included supported and explicated the main points in the piece. The introduction dimension examined the initial statement that prefaced the content that followed. The Conclusion dimension looked at final statement and how effectively it wrapped up the piece. The Clarity dimension looked at how well the ideas in the piece were expressed. Vocabulary use examined how words were used and defined in context. Vocabulary Count was an indicator of how many of the 32 unit vocabulary/inquiry words were used. Assessing this dimension consisted of examining papers for the use of any words that appeared in a corpus of 32 targeted for instruction during the unit. The dimensions on the rubric were scaled from 1-4 with the exception of the clarity dimension, which was scaled from 1-3. Vocabulary count was a frequency count ranging from 0 to 32. The Writing Rubric is contained in Appendix A.

Scaling the rubric. Scaling involved aligning a subset of papers to the characteristics described and modifying the descriptions until they accurately described the important aspects present in all of the papers examined. Rubric developers worked through many iterations of each dimension in order to match the description to the papers with a high degree of consistency. Once each of the score-point characteristics were detailed, a scale was created for each of five dimensions: evidence, introduction, clarity, conclusion, and language use.

Science content. In addition to scoring how well students expressed their ideas in writing, we were also interested in an accurate expression of science ideas in writing. Science curriculum developers crafted a science content rubric. This rubric had four levels that ranged from a minimal level of understanding (level 1) to clear understanding of the content (level 4). The Science Content Rubric is contained in Appendix B.

Scoring

Once the rubrics were complete, anchor papers for each dimension of the writing rubric were assembled. Once the scorer training process was complete, a random selection of 21 % (456 of 2144) of the writing tests were scored (215 experimental and

241 control). These tests consisted of about five matched pre- and post- writing assessments from each of the participating classrooms.

Scorers were undergraduate students who held part-time positions in research and evaluation at a large public university. They were trained monitored by an assessment specialist and worked independently from the rubric development team. During the training, scorers were introduced to six writing dimensions (excluding vocabulary word count). Annotated anchor papers were used to illustrate the high and low end of each dimension. The vocabulary word count was completed by a separate set of scorers who were put through a separate training session. On tests where there was a disagreement between the two scorers the assessment specialist scored the discrepant papers. A research specialist ran inter-rater reliability checks for each of the dimensions on the rubric.

Twenty papers were then chosen for scorer collaboration. An assessment specialist, reading specialists and the science curriculum developers each scored the twenty papers. Then, the interdisciplinary research team met to moderate the papers. The moderation consisted of iterative sessions where the team discussed their scores and the rational for their scores. During these sessions, additional changes to the rubrics were recorded and communicated to all parties. The rubric was then revised the scorer training was repeated and scored the remaining subset of papers were scored. In the end, each scorer achieved a 90% or higher reliability score with the assessment specialist on the six writing dimensions.

Data Analysis Procedures

Prior to the statistical analyses, scores on all items were examined for accuracy of data entry, missing values, and normality of distributions on the target variables. This initial data review was conducted in preparation for running independent t-tests. Independent t-tests were chosen as the best fitting analyses because this statistical procedure is commonly used to assess whether the means of two independent groups are statistically different from each other (Green & Salkind, 2007).

Assumptions for the Use Of the T-test

Group independence. The group independence assumption was adequately met within the design of the study. Control and experimental teachers were instructed not to work together or to share any information about the study (such as materials or approaches). Teachers in the study collected the same assessment data during same period of time regardless of whether they were in the experimental or control group.

Analysis of continuous variables. The continuous variables assumption was met with regard to all student scores that were collected. These included rubric based assessments in the following areas: use of evidence, introduction, clarity, conclusion, language usage, and science concepts. Each of the scores collected had a finite interval and each of the points awarded were equidistant from each other on the same scale. An additional variable, vocabulary count was also continuous with a finite interval of 32 points that could be awarded (one point for each unit-vocabulary word).

Determining whether differences between experimental and control groups were pre-existing. In order to compare the performance between the experimental and control group, the extent to which these groups were statistically different at the outset of the study was examined. Three comparisons of the experimental and control groups were conducted on assessments administered during the state-wide testing period immediately prior to our intervention. These comparisons included an examination of differences between achievement on: a) a state wide criterion referenced test of English Language Arts (CRCT ELA 2007), b) a state wide criterion referenced test of Science Concepts (CRCT Science 2007), and d) an examination of pre-test scores on the dimensions measured on the writing rubric including science content, use of evidence, introduction, conclusion, clarity, vocabulary usage, and vocabulary count.

A Bonferroni adjustment was made to control for Type I error rate due to the fact that multiple comparisons were made. The analysis indicated that there was no statistical differences between the experimental and control groups on CRCT ELA 2007 $t\ (400) = .86, p < .39$. In addition, there were no statistical differences between the experimental and control groups on Science Content $t\ (454) = .55, p < .59$, Use of Evidence $t\ (454) = .77, p < .44$, Introduction $t\ (454) = -.28, p < .78$, Conclusion $t\ (454) = -.80, p < .42$, Clarity $t\ (454) = -.53, p < .60$, Vocabulary Usage $t\ (454) = 1.15, p < .25$, and Vocabulary Count $t\ (454) = .50, p < .62$. However, the groups did differ on CRCT Science 2007 scores $t\ (646) = 2.37, p < .02$ with the control group having a higher mean score. Table 1.3 includes the descriptive statistics for these measures and Table 1.4 contains a summary of the mean comparisons.

Table 1.3

Descriptive statistics summarizing pre-existing differences between groups on writing pre-test

	Group	N	Mean	Std. Deviation	Std. Error of the Mean
Science Content (range 0-4)	Experimental	215	1.71	.74	.05
	Control	241	1.75	.70	.05
Use of Evidence (range 0-4)	Experimental	215	1.53	.85	.06
	Control	241	1.59	.84	.05
Introduction (range 0-4)	Experimental	215	2.04	.84	.06
	Control	241	2.02	.74	.05
Conclusion (range 0-4)	Experimental	215	1.92	.75	.05
	Control	241	1.87	.59	.04
Clarity (range 0-3)	Experimental	215	1.73	.79	.05
	Control	241	1.69	.74	.05
Vocabulary Usage (range 0-4)	Experimental	215	1.43	.90	.06
	Control	241	1.53	.92	.06
Vocabulary Count (range 0-32)	Experimental	215	2.27	1.32	.09
	Control	241	2.33	1.34	.09

Table 1.4

Independent Samples (t-test) mean comparisons between groups on writing pre-test

	t-test for Equality of Means						
	t	df	Sig. (2-tailed)	Mean Differenc e	Std. Error Differenc e	95% Confidence Interval of the Difference	
						Lower	Upper
Science content	.55	454	.59	.04	.07	-.10	.17
Use of Evidence	.77	454	.44	.06	.08	-.09	.22
Introduction	-.28	454	.78	-.02	.07	-.17	.13
Conclusion	-.80	454	.42	-.05	.06	-.17	.07

Clarity	-.53	454	.60	-.04	.07	-.18	.10
Vocabulary Usage	1.15	454	.25	.10	.09	-.07	.27
Vocabulary Count	.50	454	.62	.06	.13	-.18	.31

Results

To address the research question, *Do students who are taught scientific writing in the context of an integrated science/literacy unit make greater gains in writing than students who are taught science separate from literacy?* post-test scores on the writing prompt were analyzed. Before interpreting the t-test statistics, Levene's test for equality of variance was examined and determined to be significant in three cases, indicating that there were variance differences among the two groups on Science Content, Use of Evidence, and Introduction. According to Moser, Stevens, & Watts (1989) and Ruxton (2006) the equal variance not assumed statistic makes conservative adjustments for unequal group variances and is therefore an interpretable statistic. The unequal variance statistic for vocabulary usage was therefore examined, interpreted, and reported in these three cases.

A Bonferroni adjustment was made to control for Type I error rate due to the fact that multiple comparisons were made. Statistically significant differences between these groups were found on five dimensions of the writing rubric including Science Content t (414.02) = -7.17, $p < .00$, Use of Evidence t (412.66) = -3.61, $p < .00$, Introduction t (437.34) = -4.86, $p < .00$, Clarity t (454) = -4.63, $p < .00$, and Vocabulary Count t (454) = -9.39, $p < .00$. An examination of mean differences indicated that the experimental group outperformed the control condition.

Two additional dimensions measured on the rubric, Language Usage t (454) = -1.74, $p < .08$ and Conclusion t (454) = -1.47, $p < .14$ did not yield a statistically significant mean difference. Table 1.5 includes the descriptive statistics associated with the analysis of post-test differences between groups and Table 1.6 contains the means comparisons between groups.

To determine the strength of these statistical differences, a composite score was calculated. The composite score was total score made up of sub-scores on each of the seven dimensions that were used to assess students' writing. The composite score on the post-test was analyzed using an independent samples t-test. The overall Cohen's *d* effect size was calculated at .69. Cohen's rules of thumb for characterizing the strength of the effect sizes (Cohen, 1988) indicate a medium-size effect.

Table 1.5
Descriptive statistics summarizing differences between groups on writing post-test

	Group	N	Mean	Std. Dev.	Std. Error of the Mean
Science Content (range 0-4)	Experimental	215	2.7	.97	.07
	Control	241	2.1	.79	.05
Use of Evidence (range 0-4)	Experimental	215	2.22	1.24	.09
	Control	241	1.83	1.01	.07
Introduction (range 0-4)	Experimental	215	2.73	.93	.06
	Control	241	2.32	.86	.06
Conclusion (range 0-4)	Experimental	215	2.05	.56	.04
	Control	241	1.97	.60	.04
Clarity (range 0-3)	Experimental	215	2.14	.76	.05
	Control	241	1.82	.75	.05
Vocabulary Usage (range 0-4)	Experimental	215	2.21	1.10	.07
	Control	241	2.03	1.13	.07
Vocabulary Count (range 0-32)	Experimental	215	4.20	1.79	.12
	Control	241	2.71	1.63	.11

Table 1.6
Independent Samples (t-test) mean comparisons between groups on writing post-test

	t-test for Equality of Means						
	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
						Lower	Upper
Science content	-7.17	414.02	**.00	-.60	.08	-.76	-.43
Use of Evidence	-3.61	412.66	**.00	-.39	.11	-.59	-.018
Introduction	-4.86	437.34	**.00	-.41	.80	-.58	-.24
Conclusion	-1.47	454	.14	-.08	.06	-.19	.03
Clarity	-4.63	454	**.00	-.33	.07	-.47	-.19

Language Usage	-1.74	454	.08	-.18	.10	-.38	-.02
Vocabulary Count	-9.39	454	**.00	-1.50	.16	-1.82	-1.19

** $p<.001$

Discussion

The performance of the students in the experimental condition demonstrates that the intervention was generally successful in improving students' science writing. The findings suggest that when informational writing is taught in an integrated science-literacy curriculum, students make gains in the areas of accuracy of science content, use of evidence, providing an introduction, writing clarity, and using scientific vocabulary in their writing. However, the inclusion of additional information about vocabulary terms as well as ending written pieces with an appropriate conclusion appears to be challenging for students in this context.

In interpreting the results of this study, it is worth noting that, even though students were not prompted to organize their writing in any particular way during the assessment situation, they still made gains in producing pieces of scientific writing that were organized in a manner that conveyed ideas clearly. This is additional evidence for the supposition that the integrated curriculum improved students' writing organization and expression of science ideas. Another interesting finding concerns students' use of vocabulary in their writing. The intent of these two dimensions of the rubric (vocabulary count and vocabulary usage) was to measure students' productive use of science vocabulary in written form. While the count of science vocabulary terms used in student writing was significant, the dimension of vocabulary usage did not differ significantly between the experimental and the control groups. This could be an issue with the vocabulary usage dimension of the rubric, it could be a shortcoming of the curriculum, or it could be that, in the absence of a genuine audience for the final piece of writing, students did not feel the need to explain the terms that they used.

Implications

The findings of this study may help to inform further research and curriculum development pertaining to intermediate students' informational writing. The authenticity and purposefulness of integration of science with literacy, along with scaffolded instruction in writing, seems to allow students to learn and retain important aspects of science writing. However, we remain curious about which specific elements of the instruction were most influential. Further research might isolate variables such as time spent writing, teacher modeling of writing strategies, the impact of specific scaffolds offered, or a number of other factors that might have led to the positive outcomes for the students in the experimental group. The more we learn about how these factors work - in combination or in isolation - to help students become stronger writers, the better curricula or curricular models we can offer.

A model for improving student writing in the context of an integrated curriculum unit, where content knowledge and literacy work hand-in-hand, offers other potential implications. As teachers and students face increasing demands for improved student achievement, especially in reading and math, quality time for the teaching of writing becomes rare. Yet literacy, and in many ways writing in particular, is seen by many as the crucial component for success in the upper grades, high school, college and beyond (Moje, 2007; Friedman, 2005; Alexander & Jetton, 2002; Bransford, Brown, & Cocking , 2000). In fact, students with limited literacy abilities struggle with learning in school, especially in content areas such as science, and are more likely to drop out (Hammond, Linton, Smink, & Drew, 2007), thus limiting their ability to become informed citizens who are able to pursue the careers that are defining the twenty-first century, where jobs requiring specialized skills, training, and often, an advanced degree will dominate the new marketplace (U.S. Department of Labor, 2008; Levine, 2004). It is vital that educators have means to teach students sophisticated literacy skills even within the compacted school day that is the reality of today's classrooms. An integrated science and literacy curriculum like the intervention described in this paper may offer a solution for addressing student needs for learning informational writing skills well, and in a way that is efficient for teachers, since the integrated nature of this kind of intervention provides a curricular economy that is appealing in the kind of impacted learning environment students and teachers face today.

Next steps

This study offers a small but important step in studying effective ways to teach science writing. Further studies will help answer questions about which particular elements of the instruction generated positive student outcomes. This research would allow us to offer more specific pedagogical information for teachers so that they might further improve their informational writing instruction. Another critical question for further research is whether or not gains in writing transfer to other contexts. For example, if students are taught the crucial elements necessary for successful creation of one type of scientific writing, will this help them as they create other types of scientific writing? Much of the writing in science is focused on leveraging evidence and creating logical explanations and arguments. These attributes may collectively create a frame for students to use as they engage in a variety of different kinds of science writing. And, going further, could this kind of instruction have positive affects on the informational writing that students engage in for other content areas as well? The rubric we used contained measures of several general characteristics of informational writing; it is therefore conceivable that effective writing instruction within the domain of science could allow for the transfer of ideas into other writing, as students create informational texts for other content areas in school.

Finally, we are also interested in taking a closer look at gains made by particular groups of students. For example, how effective is this instruction for English language learners, gifted students or students with learning disabilities? Further investigation through dis-aggregating the data could help us to see how this kind of instruction affects different groups of students, and how we might adjust instruction to better address their varying needs.

The results of this study offer promising insights into a pedagogical approach that can help teachers address a vital area of educational concern: teaching today's students critical elements of informational writing that can used to write more sophisticated texts. As new research suggests (Shanahan and Shanahan, 2008; Greenleaf and Hinchman 2009; CCAAL, 2010) we need to teach students to become literate within content-specific domains. Through informational writing in science, students are learning both critical elements of writing, as well as some that are more specific to science: how to leverage evidence in support of an argument, use specialized language, and explain scientific concepts in a coherent and organized way. Although there is more work to be done to ensure that we understand why and how this approach offered positive results in this study, we believe that this model of integrating science and purposeful science content writing is a promising one that can be drawn upon to create more innovative curriculum in the future.

References

Alvermann, D., Swafford, J., & Montero, M.K. (2004). *Content area literacy instruction for the elementary grades.* Boston: Allyn & Bacon.

Alexander, P.A. & Jetton, T.L. (2002). Learning from text: A multidimensional and developmental perspective. In M.L. Kamil, P. Mosenthal, P.D. Pearson, & R. Barr (Eds*). Handbook of Reading Research,* Vol 3, pp. 285-310 Mahwah, NJ.

Bereiter, C., & Scardamalia, M. (2003). Learning to work creatively with knowledge. In E. De Corte, L. Verschaffel, N. Entwistle, & J. van Merrienboer (Eds.), Powerful learning environments: Unravelling basic components and dimensions. Oxford, UK: Elsevier Science.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2004). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Research Council.

Carin, A.A., & Sund, R.B. (1985). *Teaching modern science* (4th ed). Columbus, OH: Merrill.

Carnegie Council on Advancing Adolescent Literacy [CCAAL]. (2010). *Time to act: An agenda for advancing adolescent literacy for college and career success*. New York, NY: Carnegie Corporation of New York. Retrieved on October 5th, 2009 from: http://www.carnegie.org/literacy/tta/pdf/tta_Main.pdf

Cervetti, G., Pearson, P. D., Bravo, M. A., & Barber, J. (2006). Reading and writing in the service of inquiry-based science. In R. Douglas, M. Klentschy, and K. Worth (Eds.). *Linking science and literacy in the K-8 classroom.* Arlington, VA: National Science Teachers Association.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Erlbaum.

Cutler, L., & Graham, S. (2008). Primary grade writing instruction: A national survey. *Journal of Educational Psychology*, 100, 907-919.

Daniels, H., Zemelman, S. Steineke, N. (2007). Content area writing. Heinemann Press:Portsmouth, NH.

Draper, R., & Siebert, D. (2009). Content area literacy in mathematics and science classrooms. In S.R. Parris, D. Fisher, & K. Headley (Eds.), *Adolescent literacy, field tested* (pp. 105-116). Newark, DE: International Reading Association.

Driver, R., Newton, P. & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education,* 84, 287–312.

Friedman, T. 2005 *The World is Flat*. Farrar Straus & Girox, New York, NY.

Green, S. B. & Salkind, N. J. (2005). *Using SPSS for Windows and Macintosh: Analyzing and Understanding Data*, (Fourth Ed.). Upper Saddle River, NJ: Pearson

Greenleaf, C.& Hinchman, K. (2009). Reimagining our inexperienced adolescent readers: From struggling, striving, marginalized, and reluctant to thriving. *Journal of Adolescent & Adult Literacy.* (53)1, 4-13.

Guthrie, J. T., Anderson, E., Alao, S., & Rinehart, J. (1999). Influences of concept-oriented reading instruction on strategy use and conceptual learning from text. *Elementary School Journal, 99*(4), 341-366.

Guthrie, J. T., & Ozgungor, S. (2002). Instructional contexts for reading engagement. In *Comprehension instruction: Research-based best practices*, eds. C. C. Block and M. Pressley. New York: Guilford Press.

Hammond, C., Linton, D., Smink, J. & Drew, S. (2007). *Dropout risk factors and exemplary programs: A technical report.* Clemson University: National Dropout Prevention Center at Clemson University and Communities In Schools, Inc.

Holliday, W.G., Yore, L.D., and Alvermann, D.E. (1994). The reading-science learning-writing connections: Breakthroughs, barriers and promises. *Journal of Research in Science Teaching,* 31 (9), 877-893.

Klentschy, M. & Molina-De La Torre, E. (2004). Students' science notebooks and the inquiry process. In E. Saul (Ed.), *Crossing Borders in Literacy and Science Instructio*n (pp. 161-189). Newark, DE: International Reading Association.

Krajcik, J.S., & McNeill, K.L (2007). Assessing middle school students' content knowledge and scientific reasoning through written explanations. Workshop presented at the Center for Assessment and Evaluation of Student Learning (CAESL) Conference. San Jose, CA.

Lee, C.D., Spratley, A. (2010). *Reading in the disciplines: The challenges of adolescent literacy*. New York, NY: Carnegie Corporation of New York.

Lemke, J.L. (1990). *Talking science: language, learning, and values.* Norwood, NJ: Ablex Publishing Corporation.

Levine, L. (2004). Offshoring (a.k.a. Offshore Outsourcing) and Job Insecurity Among U.S. Workers. CRS Report for Congress. Washington, DC: Congressional Research Service (CRS).

Littlefair, A. (1992). Reading and writing across the curriculum. In C. Harrison & M. Coles (Eds.), *Reading for real handbook* (pp. 83–104). London: Routledge.

Martin, M.O., Mullis, I.V.S., & Chrostowski, S.J. (Eds.)(2004).TIMSS & PIRLS International Study Center, Chestnut Hill, MA: Boston College.

Moje, E.B. (2007). Developing socially just subject-matter instruction: A review of the literature on disciplinary literacy teaching. *Review of Research in Education*. 31, 1-44.

Moje, E.B., Ciechanowski, K.M., Kramer, K., Ellis, L. Carrillo, R., Collazo, T. (2004). *Working Toward Third Space in content area literacy: An examination of everyday funds of knowledge and Discourse. Reading Research Quarterly.* 39(1).

Morrow, L.M., Pressley, M., Smith, J.K., & Smith, M. (1997). The effect of a literature-based program integrated into literacy and science instruction with children from diverse backgrounds. *Reading Research Quarterly*, 32(1), 54-76.

Moss, B. (2005). Making a case and a place for effective content area literacy instruction in the elementary grades. *The Reading Teacher*, *59*, 46–55.

Moser, B.K., Stevens G.R., & Watts C.L. (1989). The two-sample *t*-test versus Satterwaite's approximate *F* test. *Communication Statistics Theory and Methodology*. 18 (39). 63–75.

National Center for Education Statistics (2007). The nation's report card: Reading 2007 (NCES 2007-496). Washington D.C.: Institute of Education Sciences, U.S. Department of Education.

National Center for Education Statistics (2006). *The nation's report card: Science 2005 (NCES 2006-466)*. Washington, D.C.: Institute of EducationSciences, U.S. Department of Education.

National Center for Education Statistics (2007). *The nation's report card: Reading 2007 (NCES 2007–496)*. Washington, D.C.: Institute of EducationSciences, U.S. Department of Education.

National Commission on Writing in America's Schools and Colleges. (2003). The neglected "R": The need for a writing revolution (No. 997548). New York: College Board.

Palinscar, A.S., & Magnusson, S.J. (2001). The interplay of firsthand and text-based investigations to model and support the development of scientific knowledge and reasoning. In S. Carver & D. Klahr (Eds.) *Cognition and Instruction: Twenty-five years of Progress*. (151-194) Mahway, NJ: Lawrence Erlbaum.

Pappas, C.C., Kiefer, B.Z., & Levstik, L.S. (1999). *An integrated language perspective in the elementary school: An action approach* (3rd ed.). New York: Longman.

Pappas, C.C., & Varelas, M. (2004). Promoting dialogic inquiry in information book read-alouds: Young urban children's ways of making sense in science. In E. Saul (Ed.), *Crossing Borders in Literacy and Science Instructio*n (pp. 161-189). Newark, DE: International Reading Association.

Pearson, P.D. (2009, Dec.) From an instructional perspective. In S. Israel & G. Duffy (Chairs) *Research on comprehension and comprehension instruction: Where to go from here?* An invited session presented at the National Reading Conference. Albuquerque, NM.

Pearson, P. D., and G. N. Cervetti (2005). Reading and writing in the service of acquiring scientific knowledge and dispositions: In search of synergies. Paper presented at the Reading Research 2005 preconvention of the International Reading Association, San Antonio, TX.

Romance, N.R., & Vitale, M.R. (1992). A curriculum strategy that expands time for in-depth elementary science instruction by using science-based reading strategies: Effects of a year-long study in grade four. *Journal of Research in Sience Teaching, 29(6*), 545-554.

Romance, N.R., & Vitale, M.R. (2001). Implementing an in-depth expanded science model in elementary schools: Multi-year findings, research issues, and policy implications. *International Journal of Science Education, 23(4),* 373-404.

Ruxton, G. D. (2006). The unequal variance t-test is an underused alternative to student's t-test and the Mann–Whitney U test. *Behavioral Ecology.* 688-690.

Shanahan, T & Shanahan, C. (2009). Teaching disciplinary literacy to adolescents: Rethinking content-area literacy. *Harvard Educational Review.* 78(1).

Schoenbach, R., & Greenleaf, C.L. (2009). Fostering adolescents' engaged in academic literacy. In L. Christenbury, R. Bomer, and P. Smagorinksy (Eds.), *Handbook on Adolescent Literacy Research*. New York, NY: Guilford Press.

U.S. Department of Labor (2008). Occupational outlook handbook 2008-2009 edition. Retrieved Nov. 3, 2009 from http://www.bls.gov/oco/ocos063.htm#

Vitale, M. R., & Romance, N. R. (2007). A knowledge-based framework for unifying content-area reading comprehension and reading comprehension strategies. In D. McNamara (Ed.) *Reading comprehension strategies: Theory, interventions, and technologies*. 73-104. NY: Erlbaum.

Wang, J., & Herman, J. (2005). *Evaluation of Seeds of Science/Roots of Reading project: Shoreline Science and Terrarium Investigations.* CRESST, Los Angeles, CA.

Appendix A: Writing Rubric

USING EVIDENCE FOR EXPLANATIONS: *This dimension looks at how well the evidence provided supports and explicates the main points in the piece. This dimension is NOT concerned with scientific accuracy or source of the evidence provided. The light interactions mentioned need to be specifically paired with an example. For the pairing to count, the material needs to be named – glass, plastic, wood, etc. Use of the word "something" is not specific enough. The interaction needs to be named either with a vocabulary word (transmit, block, reflect, refract, absorb) or with a meaning centered phrase (bounce off, go through). Heat up and warming up count, as elements of absorption. Translucent and transparent are counted only if they further explain how light is transmitted and are paired with a specifically named material. The mention of shadows must be related to blocking and paired with a specifically named material. The use of "someone" or "object" is a specific enough example but "object" or "someone" must be linked to an interaction "blocks light". Opaque must be used in conjunction with blocking light. Colors can be mentioned as things (e.g. the color black absorbs all other colors but reflects black).*

	1	**2**	**3**	**4**
Evidence	▪ None of the interactions are supported with examples/evidence.	▪ One interaction is supported with examples/evidence.	▪ Two interactions are supported with examples/evidence	▪ Three interactions are supported with examples/evidence. ▪

WRITING ORGANIZATION: *This dimension is not concerned with paragraph form, spelling, or readability of handwriting. This dimension looks at how well the writer conveys ideas.*

	1	**2**	**3**	**4**
Introduction	▪ The piece is too brief to have an introduction – it is only one or two thoughts. OR ▪ The beginning of the piece is difficult to understand. The reader is unable to follow what the student is intending to discuss.	▪ An interaction is NOT mentioned at the outset. ▪ Launches right into examples and details without setting up what will be discussed.	▪ One interaction is mentioned in the first sentence in the piece. (transmit, block, reflect, refract, absorb)	• The first sentence states that light can interact with materials in multiple ways. AND • Interactions are all further explained in some detail

<table>
<tr><td colspan="4">WRITING ORGANIZATION: This dimension is not concerned with paragraph form, spelling, or readability of handwriting. This dimension looks at <u>how well</u> the writer conveys ideas.</td></tr>
<tr><td></td><td>1</td><td>2</td><td>3</td></tr>
<tr><td>Clarity</td><td>▪ It is <u>difficult to grasp</u> the meaning.
▪ Very few ideas can be understood.
Clarity may be impeded due to one or more of the following:
▪ The relationships presented are difficult to understand.
▪ The piece is written as just a list of words.
▪ Syntax and word order interfere with the reader being able to understand ideas.
▪ Pronoun references are unclear.
▪ Several words used incorrectly, even very common words.
▪ Ideas are jumbled and many incomplete thoughts are strung together.
▪ Pieces are missing that cause the reader confusion.</td><td>▪ Ideas <u>are understandable but not always clearly communicated</u>.
▪ Requires some inferences to be made in order to grasp meaning (Some ideas make sense while others do not).
Clarity may be impeded due to one or more of the following:
▪ The context in which ideas are presented is lacking, making the flow of ideas hard to follow.
▪ Precise vocabulary is lacking.
▪ No distinction between main ideas and supporting details.
▪ Ideas are choppy and jump around.</td><td>▪ Ideas in the piece are <u>clearly communicated</u>.
▪ All ideas can be understood.
Clarity may be increased because:
▪ Transition words or phrases (for example, one idea is, next, finally, etc.)
▪ Ideas flow smoothly.
▪ The piece is well organized.
▪ Ideas are coherent.</td></tr>
</table>

WRITING ORGANIZATION: *This dimension is not concerned with paragraph form, spelling, or readability of handwriting. This dimension looks at how well the writer conveys ideas.*				
	1	**2**	**3**	**4**
Conclusion	▪ The piece is too brief to have a conclusion – it is only one or two thoughts. OR ▪ Piece is presented as just a list of words.	▪ Piece is more developed than just one or two thoughts. ▪ The presentation of ideas just ends abruptly.	▪ The reader knows that the last example is coming because a signal word or phrase is used toward the end of the piece (the last example is, finally, last but not least, the last thing is)	▪ Ends with a general statement that brings the writing to a close (e.g. And that's what light does.)

VOCABULARY: *This dimension looks at the extent to which the writer explained the interactions using the scientific words transmit, absorb, reflect, refract, and block. Additional science words used may be opaque, transparent, and translucent. For the purposes of this dimension, do not consider terms provided in the prompt (interact, light, material).*				
	1	**2**	**3**	**4**
Language Use	▪ No science word(s) are used OR ▪ Uses science word(s) BUT no definitions or examples were included.	▪ Science word(s) are used AND ▪ Word(s) are explicitly defined. ▪ NO examples are included.	▪ Science word(s) are used AND ▪ Examples are used that show an understanding of the meaning of the word(s). ▪ NO definitions are included.	▪ Science word(s) are used AND ▪ At least one of the terms is explicitly defined AND ▪ Examples are used that show an understanding of the meaning of the word(s).

Vocabulary Count
This dimension is a count of the number of words students used in their writing. 32 words are the maximum possible. The target vocabulary taught in the unit are as follows:

> **absorb, block, characteristic, emit, energy, interact, interaction, lens, light, material, ray, reflect, refract, shadow, source, transmit, transform, transformation, travel, analyze, claim, data, diagram, evidence, explanation, investigate investigation, observation observe, predict, prediction, record, scientific community**

Appendix B: Science Content Rubric

4- Above grade level	The response shows a clear understanding and describes at least three main types of interactions light can have with materials. *Names three interactions of light with a material and describes all three of them by explaining what happens in that interaction and/or giving an example with a specific material* *Uses science terms: absorb, reflect, transmit, refract, block* *No misconceptions.*
3- Grade Level	The response shows a good understanding of the main types of interactions light can have with materials. *Merely names three interactions, using science terms for some.* *Or names only two interactions (using science terms) and describes one or both at the level of Level 4* *Or describes three interactions without using science terms for more than one (light makes things hot, light makes shadows, light goes through glass)* *Misconceptions OK (e.g., black felt doesn't transmit or reflect any light).*
2- Below Grade Level	The response shows a beginning understanding that light can interact with materials in one or more ways. **Describes one or two *interactions* of light with material, at a basic level, or describes how you can see (or not see) through a material.** *Possible interactions: transmit (or go through) absorb (or block), reflect (or bounce off) (or block), refract* *Does not need to use science terms.* *May describe transmission including translucent, transparent and opaque* *Just naming one or two interactions (using science terms)* *Misconceptions OK.*
1- Inaccurate	The response does not show understanding that light can have different interactions with materials. **Describes everyday *experiences* with light.** *May include incorrect use of terms like refract* *May say generic things like "light interacts with metal"* *May include other information about light NOT related to interactions with materials* *Light travels in straight lines, light comes from a source, we see with it, it helps plants grow, it keeps us warm, you can get sunburned from light.*