THE IMPACT OF A NEW INSTRUCTIONAL MODEL ON MIDDLE SCHOOL SCIENCE WRITING

Patrick J. Enderle
The Florida State University
patrick.enderleadi@gmail.com

Victor D. Sampson
The Florida State University

Hope Campbell
Florida State University Schools

A paper presented at the 2012 Annual Meeting for the American Educational Research Association in Vancouver, British Columbia, Canada
April 16, 2012

Objective

This study focuses on the development of middle school students' scientific writing ability after participating in a year of classroom laboratory investigations designed using the Argument-Driven Inquiry (ADI) instructional model (Sampson, Grooms, & Walker, 2011)**.** ADI engages students in several writing activities, including the development of a report of their investigation and blind peer review sessions for in which they critique others' writing.

Theoretical Framework

Scientific Argumentation and Argument-Driven Inquiry

Argumentation in science represents "a logical discourse whose goal is to tease out the relationship between ideas and evidence" (Duschl, Schweingruber, & Shouse, 2007, p. 33). Discursive activity characterized as scientific argumentation involves the construction of knowledge claims supported through genuine evidence drawn from authentic inquiry and justification for those claims and evidence through connection to ideas and models privileged and accepted by the broader scientific community. The process of argumentation encompasses interactions in which individuals propose, support, critique, and refine ideas for the purpose of understanding the natural world (Driver, Newton, & Osborne, 2000; Kuhn, 1993; Sampson & Clark, 2011). These types of interactions become fundamental to the creation and evaluation of scientific knowledge, practices which serve to uniquely distinguish science from other ways of making sense of the world (Duschl & Osborne, 2002).

As central as argumentation is to the scientific enterprise, students in science classrooms are rarely afforded the opportunity to engage in these aspects of scientific practice, much less learn the epistemological commitments and warrants that separate scientific argumentation apart from other forms of argumentation (Duschl et al., 2007; National Research Council, 2005, 2008). Students must engage in authentic scientific practices in order to learn science, both concepts and skills, from their experiences. Rather than participate in laboratory experiences where they are provided with a predetermined set of procedures and organizations of data followed by several short analysis questions, students need opportunity to participate in the discursive practices of science, including the coordination of evidence and theory to support knowledge claims (Reiser, Tabak, Sandoval, Smith, Steinmuller, & Leone, 2001). Students must understand the practices, such as investigation design and collection of informative data, valued in science by experiencing them first hand. These experiences help students understand the types of methods that are privileged in science and more productive for generating scientific knowledge (Sandoval & Reiser, 2004).

As researchers and educators have come to understand the importance of scientific argumentation in the learning of science, several new instructional approaches and curricula have been developed to provide students more opportunities to learn about and how to meaningfully participate in this discursive activity. One such instructional model is called *Argument-Driven Inquiry* (ADI) (Sampson, Grooms, & Walker, 2011). The ADI model involves eight stages of educative activity that reflect the practices of science embedded within contexts that teachers can use to teach scientific concepts to their students. These stages engage students in the designing unique investigations for the purpose of answering a guiding research question through the generation of scientific arguments that are shared among their peers. The stages also necessitate student participation in scientific discourse through writing expository and persuasive investigation reports and critiquing other students' writing and arguments through a blinded peer review process. Figure 1 provides a graphical representation of the ADI instructional model.

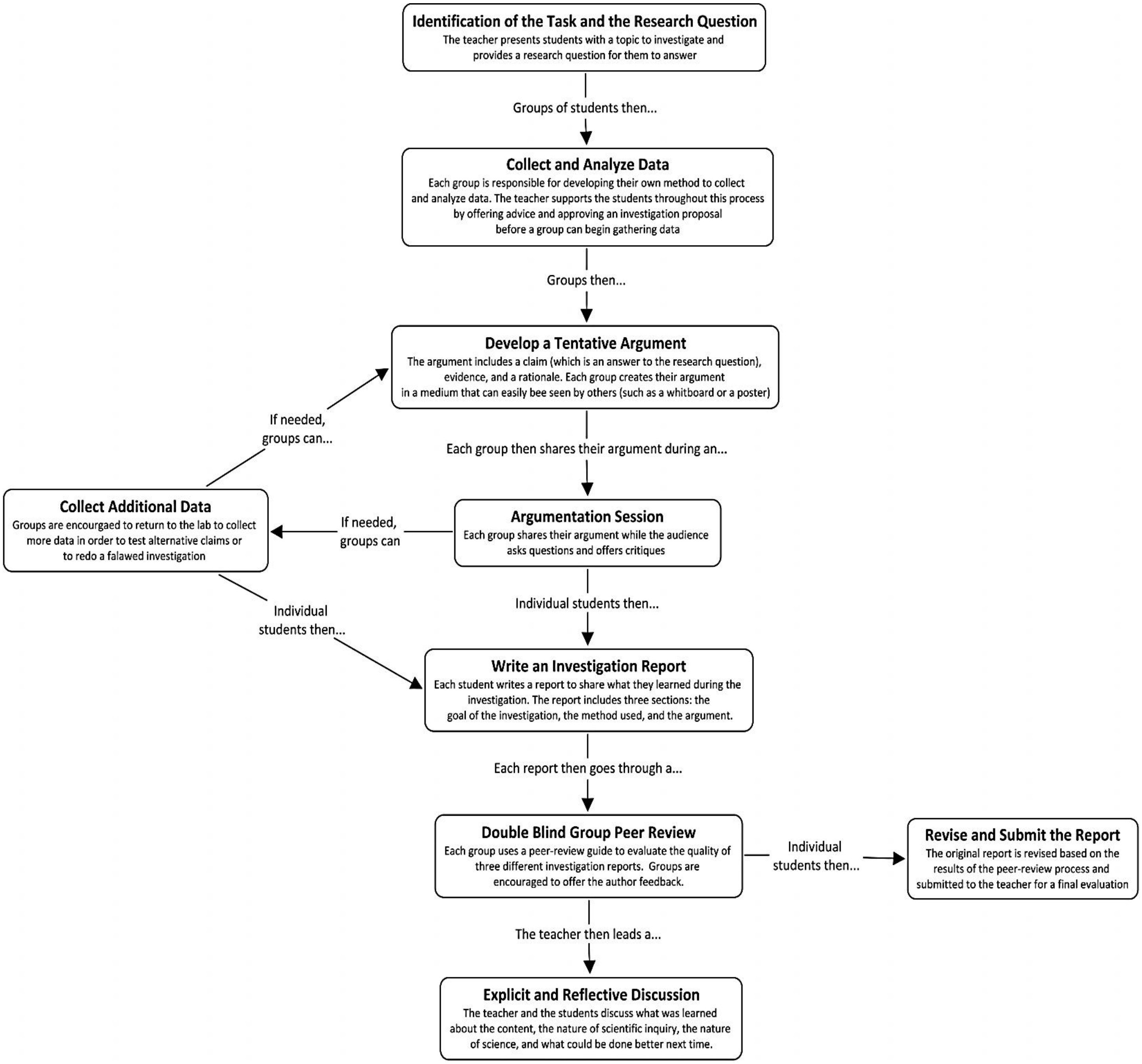

Figure 1: The Stages of the ADI Instructional Model.

Science Proficiencies as Learning Goals for Science Education

Scientific proficiency has emerged as an updated and broader concept representing the fundamental science learning desired for K12 students, stemming from the ideas of science literacy that have been central to reform efforts of the past two decades. The multiplicity of meanings developed for science literacy (Roberts, 2007) necessitated that a more comprehensive construct embodying a variety of knowledge and skills be developed. Duschl, Schweingruber, and Shouse (2007) describe science proficiency to encompass a variety of knowledge and skills required by an individual to be able to function effectively in an increasingly complex, information-driven society. The framework of scientific proficiency positions science as "both a body of knowledge and an evidence-based, model-building enterprise that continually extends, refines, and revises knowledge" (p. 2). In this view, individuals that are proficient in science: (a) know, use, and can interpret scientific explanations of the natural world; (b) can generate and evaluate scientific explanations and arguments; (c) understand the nature and development of scientific knowledge; and (d) can participate in the practices and discourse of the various scientific disciplines in a productive manner. The elements of science proficiency are also reflected in the language and substance of the emerging Common Core State Standards for Science (NRC, 2011), developed and supported by a majority of states, which move beyond a primary concern for content knowledge to encompass performance expectations and development of critical scientific practices.

By implementing instructional strategies that focus on scientific proficiency, classroom instruction shifts from traditional, prescriptive activities to those that afford students the opportunity to engage in the practices and discourse of science (Duschl, Schweingruber, & Shouse, 2007; National Research Council, 2005, 2008). The ADI instructional model is one strategy that is designed to foster the development of the four key aspects of scientific proficiency. Classroom activities structured according to the ADI model engage students in data collection and analysis, argument generation, group argumentation, scientific writing, and double blind peer review processes. The ADI instructional model is well aligned with various aspects of the scientific proficiency framework and provides a way for students to develop the knowledge and skills they need to be proficient in science while in school. For this study, the researchers chose to focus on the development of students' scientific writing abilities, a prominent discursive activity in scientific disciplines.

Writing serves as a sense-making process for students that can not only help them improve technically in their writing skills, but also provide metacognitive opportunities for learners that increase their content learning (Indrisano & Paratore, 2005; Wallace, Hand, & Prain, 2005). By measuring students' improvement in scientific writing, researchers and teachers gain insight into their development of proficiency in generating and evaluating scientific explanations and their ability to productively participate in the discursive practices of the scientific community. Multifaceted constructs such as science proficiency require a variety of tools to assess students' knowledge and abilities related to science. The researchers developed an argument focused writing assessment for this study in conjunction with several types of assessments targeting other aspects of science proficiency.

The ADI instructional model most specifically targets the enhancement of laboratory experiences in science classrooms. The working hypothesis for this study predicts that students who engage in laboratory instruction designed using the ADI instructional model throughout the course of a school year will improve in their proficiency with regard to scientific writing. In a broader sense, the design of the ADI instructional model is based on a hypothesis that efforts to improve science proficiency will require the development and continued use of laboratory experiences that are more authentic and educative. Figure 2 offers a graphical representation of this hypothesis.

Figure 2: Hypothesis Describing the Potential Impact of Implementing ADI Instruction

Methodology

The study described here occurred during year one of a larger, three-year project aimed at refining the ADI instructional model and assessing students' improvements in science proficiency as a result of experiencing ADI-based instruction (IES Grant #: R205A100909). The research presented in this study takes place in the middle school Life Science and middle school Physical Science courses at a university research K12 school in the Southeast US.

Research Context

The broader context of this research is aimed at refining the Argument-Driven Inquiry instructional model so that teachers can use it within the context of an existing middle or high school science curriculum to provide a high quality laboratory experience for their students. The project is using an iterative outcome-focused approach that is consistent with the major tenets of design-based research (Brown, 1992; Brown & Campione, 1996; The Design-Based Research Collective, 2003) to develop and refine the ADI instructional model through several iterative cycles of design, enactment, analysis, and redesign. As part of this project, the scientific writing assessment, along with several other project specific assessments, were administered on three occasions during the 2010-2011 school year; once at the beginning of the year, again at the mid-point of the school year, and finally at the conclusion of the school year. This pre-, mid-, post-assessment strategy allowed the researchers to track students' progress over the course of the school year and to measure how their levels of different aspects of science proficiency change over time. To limit potential testing effects, three slightly different versions of each assessment were created to use during the three data collection periods.

Classroom Context

The Life Science teacher, who is also the third author of this paper, implemented 12 different ADI investigations and only one of these activities did not include all stages of the model. The Physical Science teacher implemented eight ADI investigations, six of which included every stage of the model. The samples for this study include 76 students from the Life Science course (7th grade) and 75 students from the Physical Science course (8th grade). The age range for the majority of these students is 12 to 14 years old. Both of these teachers were involved in the development of the ADI investigations that were implemented in their classrooms, including the selection of focal content topics and design of the activities.

The limited number of investigations in the Physical Science course was due to a contextual pressure on the study that could not be alleviated. The Physical Science course is taken in the students' 8th grade year, which is also the year where they must take the standardized state science exam involved in determining AYP for the school. The Physical Science teacher discontinued teaching his curriculum through ADI during the second semester to conduct a six week long review unit of all middle school science content.

Data Sources

Scientific Writing Assessment: The scientific writing assessment was developed to assess students' abilities to generate and evaluate scientific arguments. This assessment provides a student with a small amount of background information and a related data table followed by a prompt. The prompt presents an argument by a scientist/expert who provides an explanation involving the data provided but flawed in a noticeable manner. The students are then directed to refute the scientist's claim using information and data provided in the question and then provide and support a counter claim using evidence and a rationale. They are also expected to be mindful of writing style and grammar. Students complete their assessments in a class period, approximately 55 minutes.

During this assessment the students are provided with several pieces of lined paper to help organize their writing. The students are initially asked to engage in a pre-writing activity to outline their argument and then generate a rough draft. Students are expected to refine any initial drafts or pre-writing exercises to provide a final draft of their argumentative essay addressing the task identified for the assessment.

Data Collection and Analysis

A general rubric was developed by the research team for scoring the writing assessment. The rubric, with an overall possible score of 28 points, was divided into three subscales: *Argument Structure* focusing on the inclusion of fundamental argument components including claims, evidence, and rationale (6); *Argument Content* concerning the quality and relevance of the argument components with respect to scientific discourse (10); and *Mechanics* regarding the punctuation, grammar, and technical quality of the writing (12). The authors acknowledge the emphasis on mechanics in the rubric developed, a feature necessitated by the context of the middle schools in an era of accountability. Although the focus of the ADI instructional model involves learning through engagement in scientific argumentation, the scientific writing assessment focuses on measuring students' proficiency to productively engage in scientific discourse through writing. The writing assessment described also provides evidence for another aspect of science proficiency, the ability to generate and evaluate scientific arguments. However, other assessments in the collection administered to students provide additional evidence of learning in regards to this aspect. The scientific writing assessment rubric privileges technical aspects of students' writing as a measure of students' ability with this type of discourse involving a unique style of expository and persuasive writing.

Using this rubric, two teams of two researchers each scored the sets of writing assessments for Life Science and Physical Science, including the pre-, mid-, and post versions. The sets of assessments were randomly assembled and blinded concerning student identity and which timed version each was. Each team scored at least 25% of the set together, and these results were used to calculate inter-rater reliability. The intraclass correlation coefficient (ICC), a measure of reliability similar to Cohen's Kappa and interpreted using the same scale, was determined for each team. The ICC (two-way random effects, absolute agreement) for the Life Science team was 0.759, and for the Physical Science team, it was 0.738. These scores demonstrate substantial agreement between the raters (Landis & Koch, 1977). The remainder of the assessment sets were divided and scored individually. The subscale and overall scores were recorded and entered into a data spreadsheet for analysis using SPSS statistical software. The data were analyzed using repeated measures general linear modeling.

Findings/Results

The primary focus of this study was to understand the development of students' proficiency in scientific writing, as measured by the assessment described above. The middle school students who served as subjects for this study experienced laboratory instruction in their respective courses that was designed using the ADI instructional model (Sampson et al., 2010). The fundamental hypothesis for the overarching study posits that students' participation in more authentic scientific practices and discourses, specifically the writing and argumentation events embedded within the ADI model, will assist in the further development of their science proficiency. As such, this nature of this study is mainly quantitative, using the scores from the different administrations to analyze the development of writing proficiency. The results presented represent a broader analysis of the data set collected to note trends related to the hypothesis. A small amount of qualitative data is included to provide examples of changes seen in student writing as measured by the scientific writing assessment.

Table 1 provides the results for the repeated measures analysis of the overall and subscale scores on the Scientific Writing Assessment for both middle school Life Science and middle school Physical Science student samples. The sphericity assumption was met for these tests. The results indicate that middle school students in both Life Science and Physical science courses demonstrated notable improvement on the Scientific Writing Assessment during the course of the year. In Life Science, students made significant improvement over time on their total score ($F(2) = 16.92$, $p < 0.001$, partial eta = 0.18), the *Structure* subscale ($F(2) =17.36$, $p < 0.001$, partial eta=0.19), the *Content* subscale ($F(2) = 6.31$, p = 0.002, partial eta = 0.08), and the *Mechanics* subscale ($F(2) = 6.63$, p = 0.002, partial eta=0.08). The overall score increased over the course of the whole year by a mean of 3.81 points, and the subscales improved as well (Structure = 1.24; Content = 1.13; Mechanics = 1.18). Similarly, students in Physical Science demonstrated significant improvement over time on their total writing score ($F(2) = 23.01$, $p < 0.001$, partial eta=0.24), the *Structure* subscale ($F(2) = 9.64$, $p < 0.001$, partial eta = 0.12), the *Content* subscale ($F(2) = 16.81$, $p < 0.001$, partial eta = 0.19), and the *Mechanics* subscale ($F(2) = 12.59$, $p < 0.001$, partial eta = 0.15). This pool of students increased their overall score by a mean of 3.9 points, improving on each subscale (Structure = 0.97; Content = 1.64; Mechanics = 1.29).

Table 1: SPSS Output for Repeated Measures General Linear Modeling

Course	Measure	Pre		Mid		Post		F	Sig.	Partial Eta Squared
		M	SD	M	SD	M	SD			
Life Science	Total	11.47	4.93	13.51	5.37	15.28	5.13	16.92	**.000**	**.18**
	Structure	1.29	1.45	2.21	1.47	2.53	1.44	17.36	**.000**	**.19**
	Content	2.89	2.12	3.28	2.53	4.03	2.58	6.31	**.002**	**.08**
	Mechanics	7.53	2.54	7.89	2.54	8.71	2.12	6.63	**.002**	**.08**
Physical Science	Total	13.63	4.45	16.61	4.42	17.53	4.52	23.01	**.000**	**.24**
	Structure	2.35	1.41	2.97	1.66	3.32	1.46	9.64	**.000**	**.12**
	Content	3.65	1.97	4.92	2.21	5.29	2.10	16.81	**.000**	**.19**
	Mechanics	7.63	2.42	8.87	1.78	8.92	2.27	12.56	**.000**	**.15**

Table 2 presents the pairwise comparisons of the writing assessment data for both science courses among the different time points in the year. The Life Science total scores continually improved from the beginning of the school year to the midpoint ($p = 0.002$) and from the midpoint to the end of the school year ($p = 0.006$). The Life Science subscale scores improved significantly over the course of the entire year, yet the timing of those improvements varied throughout the year. The *Structure* subscale improved significantly from the beginning to midpoint ($p < 0.001$) but not so from midpoint to the end ($p = 0.154$). The *Content* subscale did not improve significantly in the beginning half of the year ($p = 0.238$), but did so in the latter half ($p = 0.021$). The *Mechanics* subscale was similar in pattern with no significant improvement in the first semester ($p = 0.324$) and marked improvement in the second semester ($p = 0.010$).

For Physical Science, the pairwise comparisons of the data from different time points support the significant results for the entire year discussed earlier. The total score improved significantly in the first semester ($p < 0.001$), but did not do so in the second semester ($p = 0.128$). The *Structure* subscale scores improved significantly from beginning to midpoint ($p = 0.008$), but not so from midpoint to the end of the year ($p = 0.125$). The *Content* subscale score improvement was significant in the first half ($p < 0.001$), and not significant in the latter half of the year ($p = 0.224$). The *Mechanics* subscale scores showed improvement in the first part of the year ($p < 0.001$) but not in the second half ($p = 0.853$).

Table 2: SPSS Output for Repeated Measures Pairwise Comparisons. Time designations represent different assessment points throughout the school year.

Course	Measure	Comparison	Mean Difference	Standard Error	Sig.	Cohen's d
Life Science	Total	Pre – Mid	2.039	.640	.002	.40
		Mid – Post	1.763	.625	.006	.34
		Pre – Post	3.803	.696	.000	.76
	Structure	Pre – Mid	.921	.204	.000	.63
		Mid – Post	.316	.219	.154	-
		Pre – Post	1.237	.230	.000	.86
	Content	Pre – Mid	.382	.321	.238	-
		Mid – Post	.750	.317	.021	.29
		Pre – Post	1.132	.335	.001	.48
	Mechanics	Pre – Mid	.368	.371	.324	-
		Mid – Post	.816	.307	.010	.35
		Pre – Post	1.184	.317	.000	.50
Physical Science	Total	Pre – Mid	2.987	.612	.000	.67
		Mid – Post	.920	.597	.128	-
		Pre – Post	3.907	.598	.000	.87
	Structure	Pre – Mid	.627	.228	.008	.40
		Mid – Post	.347	.223	.125	-
		Pre – Post	.973	.222	.000	.68
	Content	Pre – Mid	1.267	.303	.000	.61
		Mid – Post	.373	.304	.224	-
		Pre – Post	1.640	.282	.000	.81
	Mechanics	Pre – Mid	1.240	.280	.000	.58
		Mid – Post	.053	.286	.853	-
		Pre – Post	1.293	.308	.000	.55

Although the mean scores on the scientific writing assessment did improve over the course of a year, the quality and magnitude of improvement varied among individual students. The students were explicitly taught concepts related to the prompts during the course of the school year. However, these topics were not covered during times close to the administration of the assessment. The assessments were designed with the provision of enough information and data so that content knowledge was not a critical factor in being able to develop the argument requested. Table 3 provides student specific examples of responses generated for the scientific writing assessment at the beginning and end of year administrations. All responses were transcribed verbatim. These qualitative pieces of data serve only as examples of the kinds of improvement in writing that can be observed in the students in this study. The amount of change between the pre and post responses involves the explicit use of relevant data as evidence, the inclusion of critique of the expert argument, and proper application of scientific concepts.

Table 3: Examples of Pre and Post Student Responses on the Scientific Writing Assessment.

Pre-Assessment Argument Section	**Post-Assessment Argument Section**
Life Science	
"Expert" Argument: "Photosynthesis occurs in all parts of the plant because there is a change in carbon dioxide in each container after two days." Requested Student Argument: "Write an essay to convince the scientist that photosynthesis occurs in the leaves of the plant."	"Expert" Argument: "Photosynthesis is not affected by temperature because each container showed change in the amount of carbon dioxide present after two days despite exposure to differing temperatures." Requested Student Argument: "Write an essay to convince the scientist that photosynthesis is affected by temperature."
Student A: "Yes, all parts of a plants body gets photosynthesis but some of the parts get more then others. Like if you scale the plants trunk usually gets more photosynthesis and the leaf gets less. Thats because the trunk is one of the mane parts, So it holds all of the nutrients and sends some to the other parts of the plant."	Student A: "My claim is Photosynthesis is affected by temperature because the number of CO_2 wasn't consistent through the whole process. If it was then the CO_2 would be the same through all of the experiment. Photosynthesis occurs in the leaves because when the temperature is 15 the CO_2 is 100 which the CO_2 is higher than when the temp is 27 the CO_2 is 50. Also the leaves had a lower temp and CO_2 than any other part of the plant."
Physical Science	
"Expert" Argument: "Sample A is gold but sample B is not. The volume of displacement and the color of Sample A match the volume of displacement and the color of gold. The volume of displacement and color of sample B are relatively close to that of gold but not a perfect match. Therefore, since some of the properties of sample A match the properties of gold, sample A must be gold." Requested Student Argument: "Write an argument to convince the scientist that sample A is not gold."	"Expert" Argument: "Coin A and Coin B are real copper. The mass, solubility, and color of the two coins are the same as copper. Therefore, since some of the properties of coin A and coin B match the properties of copper, the coins must be made from pure copper." Requested Student Argument: "Write an argument to convince the metallurgist that the coins are not made of pure copper."
Student B: "I do not think that the scientist was correct because in sample B everything is about the same but in Sample A it is much to greater."	Student B: "Coin A and B are not pure copper because of the different volumes and melting points. The metallurgist suggest that both coin A & B are pure copper because of Physical similarities mass, but if you look at the volume and melting point you would probably think otherwise. The melting point of copper is about 1010°C and the volume is about 4.75 ml. The melting point and volume for material A & B were either much to high or much to low. "

Conclusions

Improvement of Scientific Writing Proficiency

The results provided above demonstrate that middle school students that experience science laboratory instruction through the ADI model did improve in their ability to generate scientific arguments and productively communicate them through writing. In both the Life Science and Physical Science courses, the repeated measures analysis determined significant increases in the mean score for the total assessment as well as for all three of the subscales identified in the scoring rubric. The partial eta squared values for these tests show large effect sizes for most of these results, with the Life Science *Content* and Physical Science *Structure* subscales having medium effect sizes (based on Cohen (1998): partial eta squared – small ~ .01; medium ~ .06; large ~ .14). These metrics signify the magnitude of changes in the assessment scores represent meaningful improvements.

The analysis of writing assessment scores provides supporting evidence for this study's working hypothesis that students who experienced classroom science laboratory instruction designed using the ADI instructional model did improve their proficiency in communicating through scientific writing. The evidence described also supports the more fundamental hypothesis underlying the ADI model, that is the need for providing more authentic and educative experiences in the science classroom in order to improve students' multiple science proficiencies. The authors acknowledge that the supporting evidence is limited in describing a correlational relationship between ADI instruction and the development of scientific writing proficiency, rather than causational, due to the lack of a comparison group of students who were similarly assessed but did not receive ADI instruction. However, the relationship should not be dismissed lightly, particularly in light of broader discussions that have established and maintain that the majority of science instruction in current classrooms does not provide authentic and educative experiences and does little to improve students' science proficiency (NRC, 2005, 2008).

The post hoc analysis between different time intervals of assessment administration with respect to total and subscale mean scores highlights other interesting trends in the development of scientific writing proficiency. The effect sizes for these comparisons varied, yet the majority of tests that demonstrated significant differences between means related to time periods also produced at least a medium effect (based on Cohen (1998): Cohen's d – small = .20; medium = .50; large = .80). These trends suggest that the magnitudes of improvement were notable for students' ability in those areas when they were significant. The stratified nature of the improvements with regard to timing and different subscale elements suggests potential learning progressions for developing science writing proficiency.

Trends Related to Time

In the Life Science course, students demonstrated significant improvement in their ability to understand and create structured arguments during the first semester, although they did not have similar improvements during the second semester. In contrast, during the second semester, these same students evinced significant improvement in the content quality of their arguments and the mechanics of their writing, which was not evident in the first semester comparison. The trends emerging from these comparisons suggest that students, particularly younger ones such as these 7th graders, may need to work on understanding the fundamental structures and design of scientific arguments before they can appropriately incorporate relevant science content knowledge. The improvement in the mechanics of their writing could also be dependent upon understanding the dynamics of structuring arguments. Yet, it is equally plausible that the development on this specific aspect of scientific writing could be attributed to more exposure and opportunity to engage in this type of writing as they experienced even more ADI investigation activities in the second semester.

For the Physical Science course, a similar pattern of improvement in scores during the semesters, and thus more evidence for a learning progression relating scientific arguments and writing, was not observed. However, the lack of a similar pattern can be easily attributed to contextual factors that arose in this specific classroom. The students in Physical Science at this school take the state-level annual standardized science test used for calculating AYP scores across the state and other accountability measures. As this multiple choice test is comprehensive of three years' worth of science content knowledge, the teacher for this course dedicated over half of the second semester to conducting a thorough review session and test-taking practice. During this time of review, these students did not engage in any ADI instruction or any significant writing instruction. The lack of any significant improvement in students' writing scores across this same time frame makes sense in light of the lack of their engagement in authentic and educative writing practices.

Implications

These trends related to time, both in Physical Science and Life Science, provide support for the impact of prolonged and consistent implementation of curricular innovations and instruction, such as ADI-based laboratories. To achieve the learning gains and development of proficiency, which exist as common goals for most curricular change efforts, the implementation of such models must be more extensive than one lesson, one unit, or one module. The development of broader and more complex constructs like science proficiency necessitate longer and more complete immersions in learning activities, including more authentic classroom laboratory experiences like those engendered through the ADI instructional model.

Acknowledgements

The research reported here was supported by the Institute of Education Science, U.S. Department of Education through grant #R305A100909 to Florida State University. The opinions expressed are those of the authors and do not represent the views of the Institute or the U.S. Department of Education.

References

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences, 2*, 141-178.

Brown, A. L., & Campione, J. C. (1996). Psychological theory and the design of innovative learning environments: On procedures, principles, and systems. In L. Schauble & R. Glaser (Eds.), *Innovations in learning: New environments for education*. Mahwah, NJ: Lawrence Erlbaum Associates.

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

Duschl, R., Schweingruber, H., & Shouse, A. (Eds.). (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: National Academies Press.

Indrisano, R., & Paratore, J. (Eds.). (2005). *Learning to write and writing to learn: Theory and research in practice*. Newark, DE: International Reading Association.

National Research Council. (2005). *America's Lab Report: Investigations in High School Science*. Washington D.C.: National Academy Press.

National Research Council. (2008). *Ready, Set, Science: Putting Research to Work in K-8 Science Classrooms*. Washington, D.C.: National Academy Press.

National Research Council. (2011). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, D.C.: National Academies Press.

Reiser, B. J., Tabak, I., Sandoval, W. A., Smith, B. K., Steinmuller, F., & Leone, A.J. (2001). BGuILE: Strategic and conceptual scaffolds for scientific inquiry in biology classrooms. In S. M. Carver & D. Klahr (Eds.), *Cognition and instruction: Twenty-five years of progress* (pp. 263-305). Mahwah, NJ: Erlbaum.

Roberts, D.. (2007). Scientific Literacy/Science Literacy. In S. Abell & N. Lederman (Eds.), *Handbook of research in science learning* (p. 729-780). New Jersey: Lawrence Erlbaum.

Sandoval, W. A., & Reiser, B. J. (2004). Explanation-driven inquiry: Integrating conceptual and epistemic scaffolds for scientific inquiry. *Science Education, 88*(3), 345-372.

Sampson, V., Grooms, J., & Walker, J. (2011). Argument-Driven Inquiry as a way to help students learn how to participate in scientific argumentation and craft written arguments: An exploratory study. *Science Education, 95*(2), 217-257.

The Design-Based Research Collective. (2003). Design-Based Research: An emerging paradigm for educational inquiry. *Educational Researcher, 32*(1), 5-8.

Wallace, C., Hand, B., & Prain, V. (Eds.). (2004). *Writing and Learning in the Science Classroom*. Boston, MA: Kluwer Academic Publishers.