The Impact of a New Instructional Model on High School Science Writing

Jonathon Grooms
The Florida State University
jgrooms@fsu.edu

Victor D. Sampson
The Florida State University

Peter Carafano
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/Purpose

This study focuses on the development of high school students' scientific writing ability after participating in a year of classroom laboratory investigations designed using the Argument-Driven Inquiry (ADI) instructional model. ADI engages students in several writing activities, including the development of a report of their investigation and blind peer review sessions for critiquing 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 et al., 2007). Discursive activity characterized as scientific argumentation involves the construction of knowledge claims supported through genuine evidence drawn from authentic inquiry and justification for the 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.

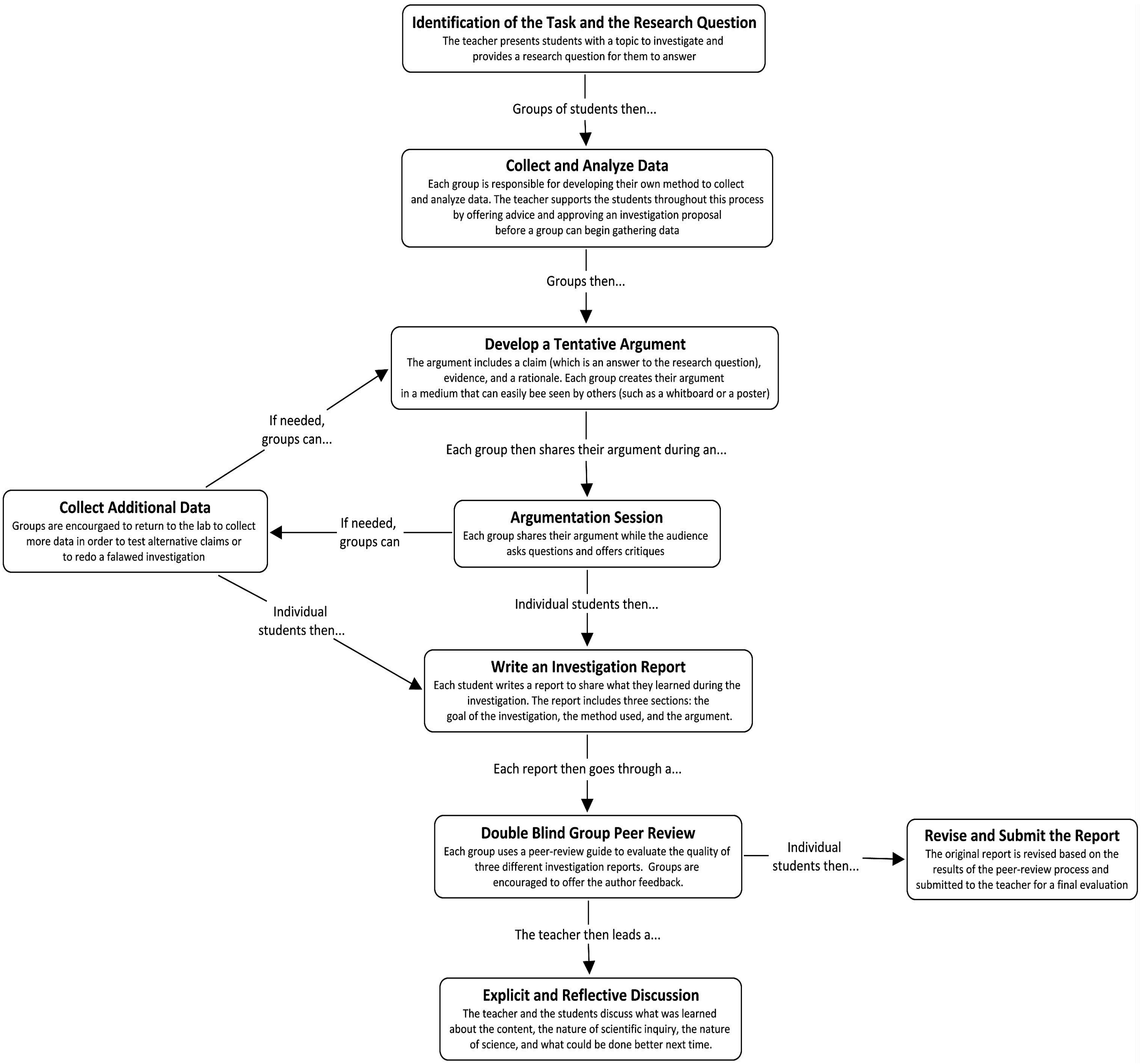

Figure 1. The stages of the ADI Instructional Model.

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 K-12 students, stemming from the ideas of science literacy that have been central to reform efforts of the past two decades. The varied 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.

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 (Sampson et al., 2011) is designed to highlight key aspects of scientific proficiency and provide a means by which to structure classroom activities such that teachers may foster the development of these proficiencies in students. Laboratory activities aligned with the ADI model engage students in data collection and analysis, argument generation, group argumentation, scientific writing, and double blind peer review processes.

One aspect central to the ADI model is scientific writing. 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 students 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. Students’ abilities to generate written scientific arguments which, are valid, coherent, and focused, is only one aspect of science proficiency. Thus, the writing assessments generated for this research serve as one facet of a multifaceted approach to assessing students’ knowledge and abilities related to science.

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 is taking place in the high school chemistry and high school biology courses in a university research K-12 school. The samples for this study include 122 students from the chemistry course (10^{th} & 11^{th} grade) and 128 students from the biology course (9^{th} & 10^{th} grade).

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 pre-, mid-, and post-intervention data collection periods.

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 inaccurate explanation for the data. The students are directed to generate their own argument in response to the scientist's claim by arguing in support of a countering claim. They are reminded to include evidence and a rationale as part of their argument based on the data and information provided in the question, all-the-while being mindful of writing style and grammar. Students completed the assessments in one class period, approximately 55 minutes.

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

The research team developed a general rubric for scoring the writing assessment. The rubric, with a maximum score of 28 points, was divided into three subscales: *Argument Structure, which* focuses on the inclusion of fundamental argument components (i.e. claims, evidence, and rationale) (6 points); *Argument Content*, which concerns the quality and relevance of the argument components with respect to the specific task/prompt (10 points); and *Mechanics*, which addresses the punctuation, grammar, and technical quality of the writing (12 points). The authors acknowledge the emphasis on mechanics in the rubric developed. 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' abilities with this type of discourse, which involves a unique style of expository and persuasive writing.

Using this rubric, two teams of two researchers each scored the sets of writing assessments for Chemistry and Biology, including the pre-, mid-, and post-intervention versions. The sets of assessments were randomly assembled and blinded to student identity and time. Each team scored at least 10% of the set together in order 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 Biology team was 0.71, and for the Chemistry team, it was 0.62. These scores demonstrate substantial agreement between the raters (Landis & Koch, 1977); however, the ICC for Chemistry was lower than desired, therefore the scoring team scored each assessment individually and generated a consensus score for each student, the consensus scores were used in all analyses. The remainder of the assessment set for biology was divided among the raters and scored individually. These data were then analyzed using repeated measures general linear modeling.

Results

The main focus of this study was to measure the changes in students' science writing abilities over the span of a single school year after experiencing a course that implemented the ADI instructional model. The hypothesis stated above, suggests that students engaged in laboratory activities aligned with the ADI model will develop science proficiency, which encompasses the ability to write and communicate ideas in a scientific manner; therefore, it is anticipated that students' performance on the writing assessment described above will improve over time as measured in this study. Figure 3 below shows the trends in students scores on the writing assessment as measured at the beginning, middle, and end of the school year. In both the biology and chemistry courses students made positive gains in their writing scores from the beginning of the year to the end of the year. However, it appears that the students in the biology course made more consistent gains across the year, whereas the students in the chemistry course made gains over the first semester and then leveled-off during the second semester.

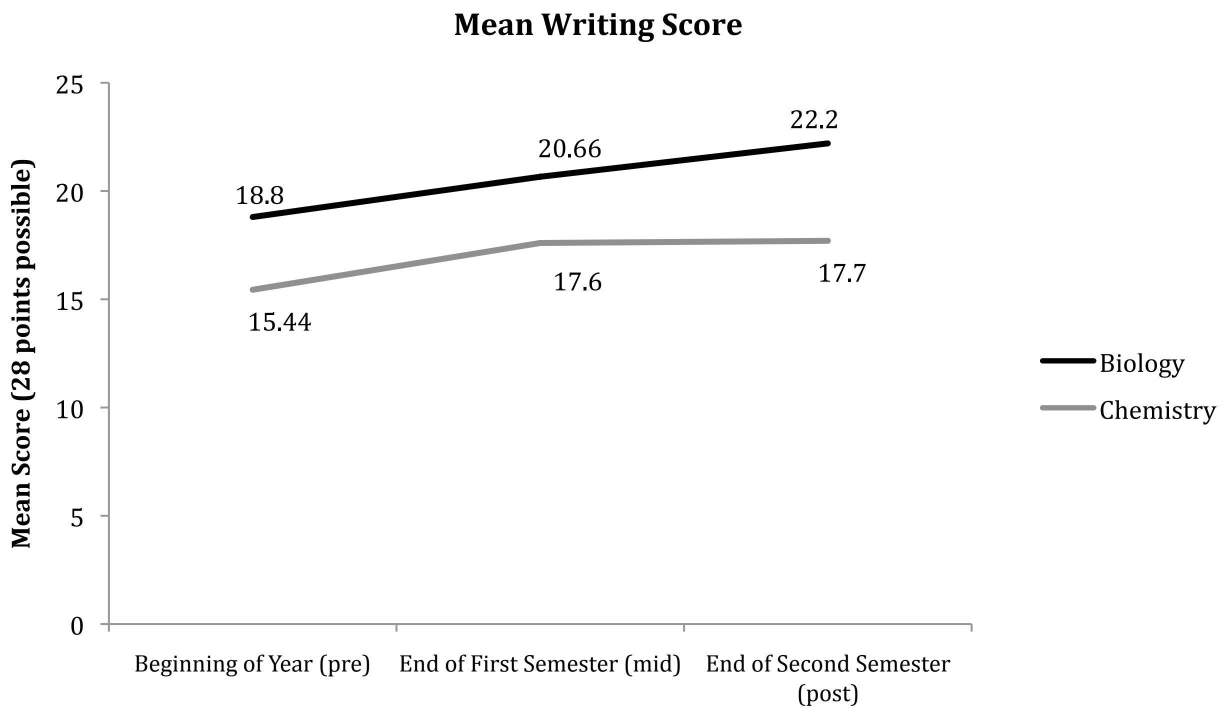

Figure 3. Mean writing scores for the biology and chemistry courses as measured a the pre-, mid-, and post-year time points.

Repeated measures analysis of variance (ANOVA) tests were conducted to determine if the changes in overall and subscale scores on the Scientific Writing Assessment for the biology and chemistry student samples were actually significant, Table 1 summarizes the results of those tests. The results indicate that high school students in both the Biology and Chemistry courses demonstrated notable improvement on the Scientific Writing Assessment over the span of the school year. In both courses the students made significant improvements in their *Total* score, *Argument Structure* score, and *Argument Content* score. However, only the Biology students made significant improvements in their *Writing Mechanics* scores. Furthermore, the partial eta squared values (η_p^2, a measure of effect size) included in Table 1 indicate that in addition to statistically significant differences in scores, the differences observed have practical significance due to the large effect size accompanying the improvements within each course. According to Cohen's (1988) scale, partial eta squared values corresponding to 0.01, 0.06, and 0.14 should be interpreted as small, medium, and large respectively. Thus, the mean differences demonstrated for the students' total writing scores, argument structure sub-scores, and argument content sub-scores are considered medium to large. However, the gains in writing mechanics are not as substantial. Even though the biology students made significant gains over the school year, $F(1.91) = 6.69$, $p < .01$, $\eta_p^2 = .05$, the effect of this gain is considered small. The chemistry students did not demonstrate significant gains in their writing mechanics subscale scores.

Table 1
Writing assessment repeated measures ANOVA results for total score and each subscale, by science course

Course	Measure (pts. poss.)	Mean Score (SD)			F(2)	*p*	Partial Eta Squared (η_p^2)
		Pre	Mid	Post			
Biology (N = 128)	Total (28)	18.80 (3.90)	20.66 (3.86)	22.20 (4.07)	45.31	< .001	.26
	Structure (6)	3.16 (1.29)	3.61 (1.27)	4.13 (1.43)	28.76	< .001	.19
	Content (10)	4.86 (2.24)	6.11 (2.38)	6.95 (2.17)	44.92	< .001	.26
	Mechanics (12)	10.68 (1.36)	10.98 (1.05)	11.12 (1.23)	6.69*	< .01	.05
Chemistry (N = 122)	Total (28)	15.43 (4.45)	17.68 (3.56)	17.70 (4.39)	20.32	< .001	.14
	Structure (6)	2.64 (1.49)	3.77 (1.91)	3.87 (1.47)	38.34	< .001	.24
	Content (10)	3.80 (2.05)	4.80 (1.87)	4.82 (2.09)	16.14	< .001	.12
	Mechanics (12)	9.04 (1.94)	9.10 (1.70)	8.89 (1.88)	1.34	.57	-

*Greenhouse-Geisser correction was used due to a violation of the sphericity assumption

Table 2 provides additional data in the form of pair wise comparisons for the mean differences of the writing assessment data for Biology and Chemistry. These comparisons demonstrate the changes in students' scores from pre- to mid-year, then mid- to post-year and overall from pre- to post-year with respect to their *total score*, *argument structure score*, *argument content score*, and *writing mechanics score*. Also included is the effect size for each of the differences observed, in the form of Cohen's d statistic. Cohen's d values of 0.20, 0.50, and 0.80 are interpreted as small, medium, and large, respectively (Cohen, 1992).

The pair wise comparisons echo the line graphs displayed in Figure 3. For the students in the biology course the mean score on each subscale and the writing assessment as a whole show gradual improvement across the three time periods (pre, mid, and post). These results indicate that the students were continuing to improve their writing abilities over the course of the whole school year. Furthermore the results for the biology course show that the students made similar gains over the first semester (pre- to mid-year comparison) and the second semester (mid- to post-year comparison), which further supports the notion of gradual, continuous improvement in their writing abilities. The effect sizes for the gains in biology scores are only moderate when observed by semester, and as expected, are much larger when viewed from a pre- to post-year perspective. These data support the notion that large improvements in students' writing abilities take time (i.e. more than a semester) and prolonged experiencing with writing-intensive instructional strategies.

Table 2
Biology and Chemistry mean differences between pre-, mid-, and post-intervention writing assessments for total score and each of the three assessment subscales

Course	Measure	Comparison	Mean Difference	t	*p*	Cohen's d
Biology	Total	Pre – Mid	1.86	5.38	< .001	.48
(N = 128)		Mid – Post	1.53	4.58	< .001	.41
		Pre – Post	3.39	4.16	< .001	.77
	Structure	Pre – Mid	.45	.70	< .01	.31
		Mid – Post	.52	.76	< .001	.37
		Pre – Post	.96	1.22	< .001	.66
	Content	Pre – Mid	1.25	5.64	< .001	.50
		Mid – Post	.84	4.19	< .001	.37
		Pre – Post	2.09	8.67	< .001	.77
	Mechanics	Pre – Mid	.31	2.67	< .01	.24
		Mid – Post	.13	1.13	.26	-
		Pre – Post	.44	3.23	<.01	.29
Chemistry	Total	Pre – Mid	2.25	5.88	< .001	.53
(N = 122)		Mid – Post	.03	.06	.95	-
		Pre – Post	2.28	5.09	< .001	.46
	Structure	Pre – Mid	1.13	7.88	< .001	.72
		Mid – Post	.10	.62	.54	-
		Pre – Post	1.23	7.46	< .001	.68
	Content	Pre – Mid	1.00	4.91	< .001	.45
		Mid – Post	.03	.12	.90	-
		Pre – Post	1.03	4.88	< .001	.44
	Mechanics	Pre – Mid	.06	.31	.76	-
		Mid – Post	-.21	-1.08	.28	-
		Pre – Post	-.15	-.68	.50	-

The second half of Table 2 shows the gains in writing scores within the chemistry course. Referring back to the line graph in Figure 3, the mean chemistry score did not change from the mid- to post-year assessment. The pair wise comparisons in Table 2 also support that observation for total writing score and for each subscale. The chemistry students did make statistically significant gains across the year as a whole (pre- to post-year) and from the pre- to mid-year assessment on total score, argument structure, and argument content, however, those gains did not carry-on during the second semester.

Conclusion/Discussion

Improved Science Proficiency

The results above demonstrate that high school students that experience science instruction through the ADI model did improve in their ability to generate scientific arguments and productively communicate them through writing. Students in both courses made significant gains over the span of the school year. Additionally the students made significant gains in the structure and content of their arguments. The ADI model places argument generation and evaluation at the center of laboratory instruction, therefore improvements in students' abilities to produce quality arguments is encouraging and consistent with the nature of classroom instruction the students experienced. These gains also represent an increase in students overall science proficiency, specifically in the two aspects related to their ability to generate and evaluate scientific explanations and arguments and their ability to participate in the practices and discourse of the various scientific disciplines in a productive manner.

Argument generation and evaluation are essentially to the ADI model and the writing process is an integral strategy to help students refine their arguments and make sense of the science content addressed in ADI investigations. One aspect of the ADI model that is intended to help students with their scientific writing is preparation of the individual investigation report and the subsequent peer review session. These sessions are intended to provide feedback to help students improve their reports. In light of such efforts it may seem troubling that larger gains were not experienced in terms of students writing mechanics in the context of the writing assessment. The chemistry students demonstrated no significant gains in their mean writing mechanics subscale scores, whereas the biology students did demonstrate a small significant increase in writing mechanics scores, however, that change was accompanied by a small effect size (($\eta_p^2 = .05$).

These small gains in writing mechanics scores are not as troubling as they may initially appear when considered with respect to where the students began the year in terms of their mechanics scores. Them mean mechanics score for the biology course was approximately 83% at the beginning of the year and the mean mechanics score for the chemistry course was approximately 75%. It is entirely possible that there was a ceiling effect that influenced the students' abilities to demonstrate substantial growth in this area. When compared to the other writing assessment subscales, the students in either course demonstrated a mean no higher than 52% at the beginning of the year. Thus, the students had more room for growth in the areas related to their argument structure and the content of their argument, which is consistent with the results presented above.

Trends Over Time

Perhaps the most notable pattern within the data from this study is the differential trend in students' writing assessment scores over the span of the school year with respect to each course. Figure 3 depicts the steady, continuous improvement within the biology course from the beginning, to the end of the school year; however, the mean writing assessment score for chemistry does not follow the same trajectory. The scores in chemistry experienced marked improvement from the beginning to the middle of the year and then leveled off, showing no improvement, over the last half of the school year.

We posit that the leveling-off of scores within the chemistry course is due to a major contextual factor disproportionately influenced the chemistry courses at the research site. Within the state where this study was conducted, the grade level corresponding with the chemistry course is also the grade level in which the state standardized assessment, which contributes to school grades and AYP, is administered.

During the second half of the school year the chemistry instructors took time out of their class periods in order to review for the science portion of the state assessment, specifically topics related to physics, as the school does not offer a physics course. During the review time period, and standardized testing the chemistry instructors continued to complete ADI investigations, however, they strayed from the actual ADI model and truncated the ADI investigations after the argumentation session. By altering the instructional model in this fashion, the teachers removed the individual report writing stages and subsequent peer review stage, essentially eliminating the writing components of the ADI model. Additionally, due to further time constraints other ADI investigations completed toward the end of the school year also suffered from a lack of writing components.

The trend in chemistry scores over the span of the school year compared to the trend in biology scores over the same time period helps to bolster that the lack of writing components in the chemistry ADI investigations lead to diminished returns in chemistry scores over the second semester. The biology scores increased in a similar fashion over each of the two semesters, whereby the biology instructors completed similar numbers of investigations during both semesters. Referring back to Table 3, the growth in biology scores was consistent, with each semester accounting for approximately 50% of the total increase in scores. The mean chemistry score experienced a similar magnitude of increase as the biology scores during the first semester, when both courses again completed similar numbers of investigations. Thus, it stands to reason that had the chemistry courses continued to emphasize students writing, as the ADI model call for, during the second semester, it is likely that the mean writing assessment score would have continued to increase in a similar fashion to the biology mean score.

Implications

The different patterns of results between the Biology and Chemistry courses speak to the need for continuous involvement in activities that help students develop writing skills. During the school year, the Chemistry course was influenced more heavily by the annual state standardized assessment. The chemistry teachers took time out of their curriculum to help prepare students for the science portion of the test, which factors into AYP measures for the school. This resulted in the chemistry teachers, due to time constraints, removing the requirement for written investigation reports for each ADI activity they conducted during the second semester. Removing the report writing and peer-review phases of the ADI model seems to have impacted student improvements during the second semester.

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.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159.

Driver, R., Newton, P. and Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education* 83(3): 287-312.

Duschl, R. and Osborne, J. (2002). Supporting and promoting argumentation discouse in science education. *Studies in Science Education* 38: 39-72.

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.

Kuhn, D. (1993). Science as argument: Implications for teaching and learning scientific thinking. Science Education, 77(3), 319-337.

Landis, J. R. and Koch, G. G. (1977). A one-way components of variance model for categorical data. *Biometrics* 33: 671-679.

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.

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., & Clark, D. (2011). A comparison of the collaborative scientific argumentation practices of two hIgh and two low performing groups. *Research in Science Education 41(1),* 63-97.

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.