Abstract Title Page

Not included in page count.

Title:

The Use of Argumentation in Science Education to Promote the Development of Science Proficiency: A comparative case study

Authors and Affiliations:

Patrick Enderle
Florida State University

Jonathon Grooms
Florida State University

Victor Sampson
Florida State University

Abstract Body

Limit 4 pages single-spaced.

Background / Context:
Description of prior research and its intellectual context.

Science proficiency, as defined by Duschl, Schweingruber, and Shouse (2007), refers to the knowledge and skills that individuals need to have in order 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). Individuals that are scientifically proficient can: (a) understand and use scientific explanations of the natural world; (b) understand the nature and development of scientific knowledge; (c) create and evaluate scientific explanations and arguments; and (d) productively participate in the practices and discourse of the scientific community.

In order to focus more on scientific proficiency, classroom instruction needs to shift from traditional, prescriptive activities to those that afford students the opportunity to engage in the practices of science such as argumentation (Duschl, Schweingruber, & Shouse, 2007; National Research Council, 2005, 2008). The Argument-Driven Inquiry (ADI) instructional model (Author, 2011) is one strategy designed to foster the development of the four key aspects of scientific proficiency. Classroom activities structured according to the ADI model engage students in designing data collection and analysis procedures, argument generation, group argumentation, scientific writing, and double blind peer review processes. Figure 1 describes the stages involved in the ADI model. 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.

[Insert Figure 1 Here]

The assessment of multifaceted constructs, such as science proficiency, requires the use of several different instruments. The ability to know and use scientific content knowledge to solve and explain problems, for example, is a key component of scientific proficiency and requires a unique assessment when compared to other aspects of scientific proficiency such as the ability to participate in the practices and discourse of a scientific discipline. Using one assessment to measure scientific proficiency would offer a biased view of students' abilities, as not all assessments are adequate for all learning outcomes.

Purpose / Objective / Research Question / Focus of Study:
Description of the focus of the research.

This study explores students' development of science proficiency over the course of an academic year in two different high schools. The students at both high schools were enrolled in the same Biology course; however, the nature of the laboratory instruction in these two contexts was rather different. The objective of this study, as a result, was to document what the students in these two contexts learned in order to examine how the nature of laboratory instruction can hinder or foster the development of science proficiency over time.

Setting:
Description of the research location.

This research setting involved the high school biology course at a university research school (School A) and a neighboring public high school in the same county (School B). The student demographics at each school are provided in Table 1. Two teachers from each school participated in this study.

[Insert Table 1 Here]

Population / Participants / Subjects:
Description of the participants in the study: who, how many, key features, or characteristics.

The initial subject of students included 265 students from 16 different sections of the course, taught by four teachers. However, due to consent form considerations and attendance on the assessment administration days, the analyzed samples vary for each assessment (see results).

Intervention / Program / Practice:
Description of the intervention, program, or practice, including details of administration and duration.

Each teacher enacted a variety of laboratory experiences over the course of the school year; Table 2 provides an overview of the nature of these experiences. These data were obtained through reflective interviews with each teacher at the end of the school year. The activities associated with the treatment group teachers are consistent with the ADI instructional model.

[Insert Table 2 Here]

The ADI instructional model places a heavy emphasis student participation in scientific practices, such as argumentation and writing, as a means for students to make sense of science concepts. This type of emphasis on student engagement in the practices of science is not necessarily common in traditional approaches to teaching laboratory; therefore, it is important to differentiate between the contexts with respect to opportunities to participate in the practices of science in each context. Table 3 provides information about how often student participated in each scientific practice in both contexts.

Research Design:
Description of the research design.

We conducted a comparative case study (Yin, 2003) to examine what students learned from their laboratory experiences at the two different schools. We administered four different assessments at the beginning and end of the school year in order to document how students' performance on each assessment changed over time. Our goal for examining how student performance on each assessment changed was not to determine which group of students learned more; rather it was to examine what these students learned in each instructional context. The assessments included:

- *Biology Content Knowledge Assessment* (ICC = .79): This assessment measures how well a student knows and can use scientific explanations of the natural world. The assessment is comprised of eight scenarios, each related to one of several "Big Topics" in biology. Each scenario includes an opening paragraph that provides a relevant context, followed by two questions; one which asks the student to *describe* the fundamental biology concept (*Know*) and the other asks the student to *apply* that concept to explain the scenario provided (*Use*).

- *Science Specific Argumentative Writing Assessment* (ICC = .79): This assessment provides a small amount of background information and a related data table followed by a prompt. The prompt presents an argument by a scientist who provides an inaccurate explanation for the data. The students are directed to respond to the scientist's claim by generating an argument in support of a countering claim, which includes evidence and a justification based on the data and information provided in the question, being mindful of writing style and grammar.

- *Biology Performance Task Assessment* (ICC = .85): This assessment measures students' abilities to design an investigation that will allow them to generate an argument in response to a research question. Completing this assessment involves developing an original investigation and making decisions about what data to collect and evidence to use to generate an argument. The task includes areas for students to describe their investigation design, the data collected, and the final argument, along with justification for each of these sections.

- *Student Understanding of Science and Scientific Inquiry* (SUSSI) instrument (Liang, et al., 2006) was adapted to measure students' understanding of the development and nature of scientific information. The assessment was comprised of 44 statements about science with Likert-scale agreement responses offered. Statements representing accurate ideas about science and scientific inquiry were scored on a scale from 0 (strongly disagree) to 4 (strongly agree). Statements representing inaccurate ideas about science were scored in a reverse manner. The researchers condensed the original subscales into two groups to better align them with Aspect 2 of the science proficiency framework.

Table 4 provides an overview of how each assessment is aligned with the four aspects of science proficiency.

Data Collection and Analysis:
Description of the methods for collecting and analyzing data.

All of the assessments were administered at the beginning and the end of the year. All assessments were scored using rubrics developed by the research team. A triad of research team members scored at least 25% of the full set of each assessment, which had been blinded concerning student identity, teacher, and pre/post timing. The intra-class correlation coefficient (ICC), a measure of reliability similar to Cohen's Kappa and interpreted using the same scale, was determined for each team (two-way random effects, absolute agreement). An ICC above 0.6 is considered substantial agreement (Landis & Koch, 1977), and once this level of agreement was determined, the team members scored the remainder of their assessment sets individually. ICC values are included above with the assessment descriptions. Paired-samples t-tests were then used to assess growth over the school year and effect sizes were calculated to compare gains made by students in each context.

Findings / Results:
Description of the main findings with specific details.

Table 5 shows the results for the paired-samples t-tests conducted within each context for each assessment used in this study. The students in both contexts made statistically significant gains in terms of their content knowledge and their overall scores on the performance task. However, only the students who participated in ADI laboratories made significant gains with respect to their scientific writing abilities and their understanding of the development of and nature of scientific knowledge.

An alternative approach to evaluating the gains for students in each context is to compare the effect size of time on the students' scores with respect to each aspect of science proficiency. Figure 2 shows a comparison of the effect size for the significant gains of each group aligned to the four aspects of science proficiency. An effect size of zero is associated with non-significant gains in an aspect of science proficiency. When using Cohen's d as a measure of effect size, small, medium, and large effects are associated with values of 0.2, 0.5, and 0.8 respectively (Cohen, 1992).

[Insert Figure 2 Here]

The students in each context made significant gains in content knowledge over the course of the school year, but the improvement demonstrated by the students who participated in ADI laboratories are associated with much larger effect sizes. The effect sizes associated with the treatment students' gains in their abilities to *use scientific explanations* (1.50) and *generate scientific explanations and arguments* (0.98) are nearly twice as much as those for the comparison students (0.67 and 0.52, respectively). These differences are likely due to the strong emphasis on scientific argumentation found in the ADI instructional model.

The effect sizes associated with the students who participated in ADI laboratories for the aspects of science proficiency related to the development of (0.30) and nature of scientific knowledge (0.25) are rather small; however, considering the traditional approach to laboratory instruction generated no effect in these areas, these results are encouraging. Finally, the students who participated in ADI laboratories also demonstrated larger gains in their abilities to communicate in writing as evident by a 0.34 effect size compared to zero effect within the comparison group. Given the heavy emphasis on scientific writing within the ADI model, one might expect a larger effect; however, considering that no effect was observed for students who participated in traditional laboratory instruction, the magnitude of the effect of time is encouraging in terms of establishing the potential and promise of the ADI instructional model.

Conclusions:

Description of conclusions, recommendations, and limitations based on findings.

These results suggest that ADI laboratories may have a positive impact on the development of science proficiency in the context of high school biology. Significant changes from the beginning to end of the year were noted on all assessments with moderate to large effect sizes in many areas although the results suggest that students improved on some aspects of science proficiency more than others. Overall, the results of this study suggest that the ADI instructional model has promise and potential as a way to enhance students' science proficiency. The broad performance profile generated from this analysis also demonstrates the importance of using multiple assessments for gaining insight into complex science learning.

This study provides further evidence of the potential benefits of using argument-focused instruction in science classroom. The findings also illustrate potential targets for improvement and new directions for future research. As K12 science education shifts focus to developing students' science proficiency, this study contributes to the research base on ways of assessing aspects of a very broad and complex construct that serves as one approach to understanding the learning of critical thinking skills.

Appendix A. References

References are to be in APA version 6 format.

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

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.

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

Liang, L., Chen, X., Chen, S., Kaya, O., Adams, A., Macklin, M. & Ebenezer, J. (2006). Student Understanding of Science and Scientific Inquiry (SUSSI): Revision and Further Validation of an Assessment Instrument. *Paper presented at the 2006 Annual Conference of the National Association for Research in Science Teaching (NARST). San Francisco, CA, April 3-6*.

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.

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.

Appendix B. Tables and Figures

Not included in page count.

Figure 1. Stages of the Argument Driven Inquiry (ADI) instructional model

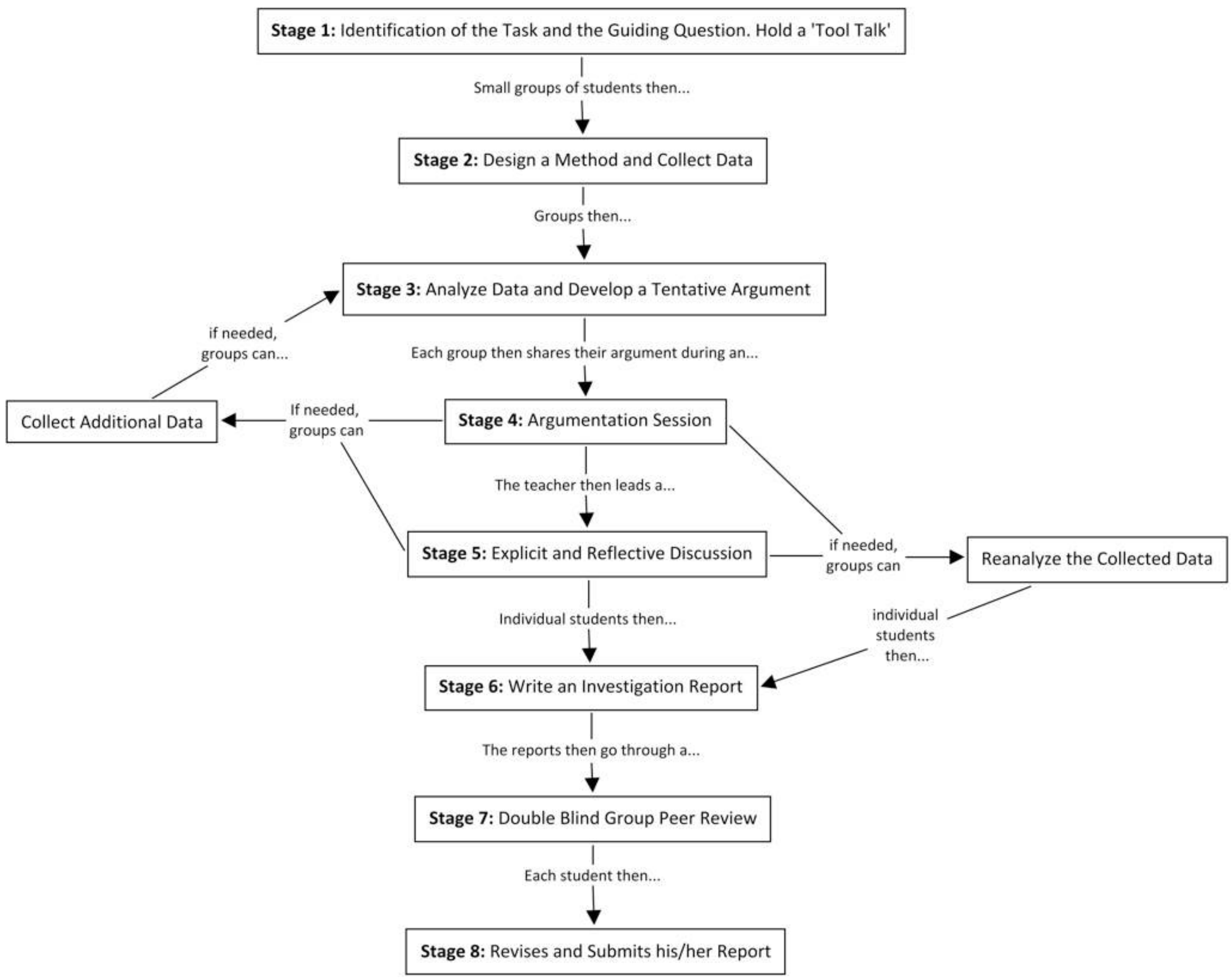

Figure 2. Effect size of time for each aspect of science proficiency in each context

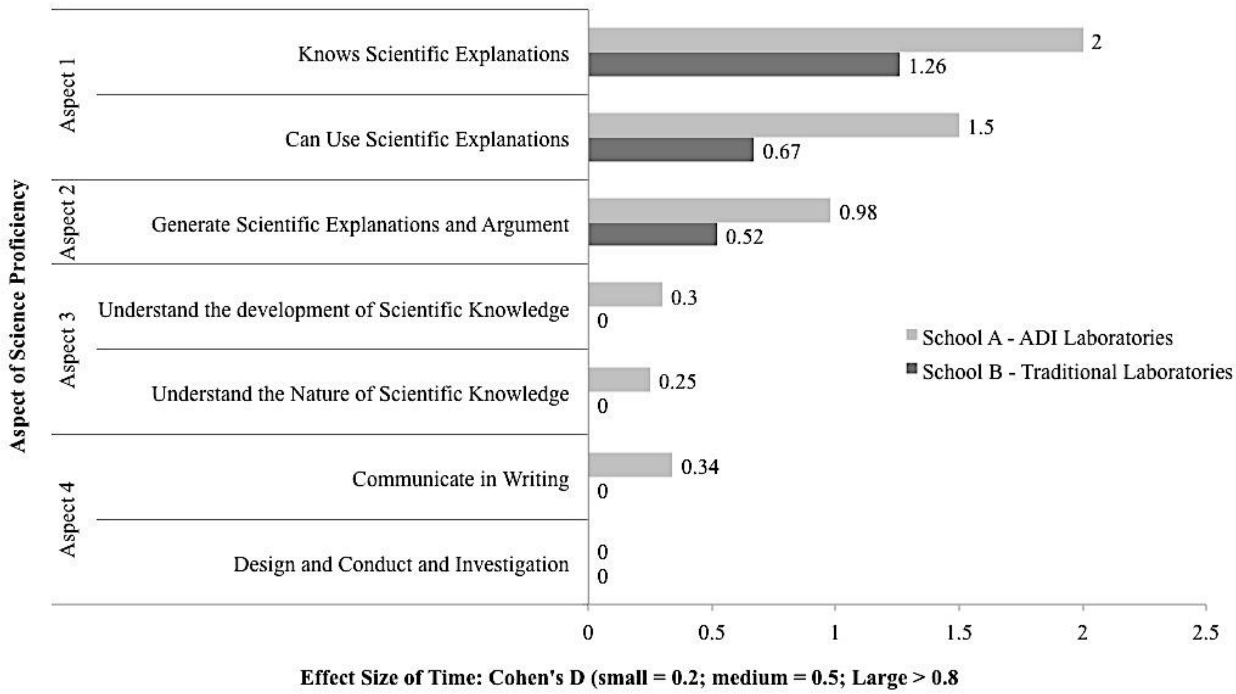

Table 1. *Demographic data for each participating school*

Context	Total Enrollment	Race (% of enrollment)					Free/Reduced Lunch Eligibility (%)	2010-11 School Grade
		White	Black	Hispanic	Asian	Other		
School A (ADI Laboratories)	659	50.8	31.4	9.7	2.0	6.1	28.3	A
School B (Traditional Laboratories)	1,877	60.6	30.2	3.2	3.2	2.7	17.8	B

Table 2. *The number and the nature of the laboratory experiences provided by the teachers*

School	Teacher	Number of Sections	Number of Laboratory Experiences	Nature of the Laboratory Experiences			
				Activity (Level 0)	Structured Investigation (Level 1)	Guided Investigation (Level 2)	Open Investigation (Level 3)
School A (ADI Laboratories)	A	2	12	0 (0%)	0 (0%)	12 (100%)	0 (0%)
	B	5	9	0 (0%)	0 (0%)	9 (100%)	0 (0%)
School B (Traditional Laboratories)	C	4	30	19 (63%)	8 (27%)	2 (7%)	1 (3%)
	D	5	15	13 (87%)	1 (7%)	1 (7%)	0 (0%)

Table 3. *Number of opportunities students had to participate in a practice of science during a experience*

School	Teacher	Number of Laboratory Experiences	Scientific Practices: Design an Investigation	Analyze and Interpret Data	Construct an Explanation	Argue from Evidence	Give an Oral Presentation	Evaluate Information	Communicate Ideas through Writing
School A (ADI Laboratories)	A	12	12 (100%)	12 (100%)	12 (100%)	12 (100%)	12 (100%)	12 (100%)	10 (83%)
	B	9	9 (100%)	9 (100%)	9 (100%)	9 (100%)	9 (100%)	9 (100%)	7 (78%)
School B (Traditional Laboratories)	C	30	5 (17%)	27 (90%)	11 (37%)	6 (20%)	3 (10%)	3 (10%)	12 (40%)
	D	15	0 (0%)	10 (71%)	1 (7%)	1 (7%)	1 (7%)	0 (0%)	1 (7%)

Table 4. *Aspects of science proficiency and associated assessment*

Aspect of Science Proficiency	Description	Assessment Instrument
Aspect 1	Students know, use, and can interpret scientific explanations of the natural world	Biology Content Knowledge Assessment
Aspect 2	Students can generate and evaluate scientific explanations and arguments	Biology Performance Task - Argument Generation Section
Aspect 3	Students understand the nature and development of scientific knowledge	SUSSI
Aspect 4	Students productively participate in the practices and discourse of the scientific community	Biology Performance Task - Investigation Design Section Scientific Writing Assessment

Table 5. *Results of paired-samples t-test analyses for each group on each of the four assessments used in this study*

Assessment	School	Scale	Pre		Post		t	df	p	d
			Mean	SD	Mean	SD				
Content Knowledge	A - ADI (N=99)	Total	3.86	2.80	16.60	7.55	18.55	98	<.001*	1.86
		Know	1.32	1.14	7.03	3.11	19.97	98	<.001*	2.00
		Use	2.54	2.11	9.57	5.05	14.95	98	<.001*	1.50
	B - Traditional (N=31)	Total	2.81	3.19	6.90	4.90	6.12	30	<.001*	1.10
		Know	1.00	1.06	3.19	2.04	7.02	30	<.001*	1.26
		Use	1.81	2.37	3.71	3.40	3.74	30	.001*	0.67
Writing	A - ADI (N=103)	Total	14.05	4.85	15.76	4.48	3.42	102	.001*	0.34
	B - Traditional (N=49)	Total	12.61	5.47	13.31	4.91	.805	48	.425	-
Performance Task	A - ADI (N=88)	Total	15.70	5.85	19.03	5.03	5.47	87	<.001*	0.58
		Investigate[1]	10.07	3.77	10.28	3.53	.499	87	.619	-
		Argument[2]	5.64	3.21	8.75	2.27	9.16	87	<.001*	0.98
	B - Traditional (N=30)	Total	11.00	3.81	12.87	5.78	2.18	29	.037*	0.40
		Investigate[1]	7.33	2.96	7.57	3.78	1.66	29	.740	-
		Argument[2]	3.67	2.92	5.30	2.83	2.87	29	.008*	0.52
SUSSI	A - ADI (N=106)	Total	85.27	14.55	89.66	15.94	3.55	105	.001*	0.35
		NoSK[3]	33.62	5.48	35.02	5.57	2.57	105	.012*	0.25
		DoSK[4]	51.65	11.45	54.64	12.11	3.04	105	.003*	0.30
	B - Traditional (N=43)	Total	81.79	9.29	82.37	9.22	.355	42	.724	-
		NoSK[3]	50.81	7.89	50.51	7.95	1.38	42	.174	-
		DoSK[4]	30.98	3.86	31.86	2.79	-.209	42	.836	-

*Significant at the p<.05 level
[1]Design and conduct and investigation subscale
[2]Generate an argument subscale
[3]Nature of scientific knowledge subscale
[4]Development of scientific knowledge subscale