A Comparative Study of the Development of Science Proficiency in High School Chemistry

Jonathon Grooms
Learning Systems Institute
The Florida State University

Patrick Enderle
Learning Systems Institute
The Florida State University

Victor Sampson
School of Teacher Education and FSU-Teach
Florida State University

Subject/Problem

This study explores students' development in several aspects of science proficiency in the context of two instructional models for laboratory investigations implemented in high school chemistry. Students experienced either a traditional model of laboratory instruction that can be generally described as teacher-centered, while students at a different institution experienced the argument-driven inquiry instructional model (ADI) (Author, 2011).

Theoretical Framework

Science proficiency, as defined by Duschl, Schweingruber, and Shouse (2007), includes a variety of knowledge and skills that individuals need to develop 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 and discourse of science (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. The ADI instructional model is well aligned with the 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.

Additionally, 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.

Method

This study occurred during year two 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). In order to further investigate the merits of the ADI approach, a quasi-experimental study was conducted during the second year of the project to compare the development of students' scientific proficiency in the context of the ADI model and a more traditional approach to laboratory instruction. Changes in students' science proficiency were measured using a battery of pre/post-intervention assessments.

Instructional Contexts

This research setting involved the high school chemistry course at a university research school and a neighboring public high school in the same county. The student demographics at each school are provided in Table 1.

Table 1. *Demographic data for each participating school*

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

Two teachers from each school participated in this study. Each teacher enacted a variety of laboratory investigations over the course of the school year; through reflective interviews with each teacher at the conclusion of the school year, each investigation was categorized as an *activity*, *structured investigation*, *guided investigation*, or open investigation. The investigations for each of the treatment group teachers (A & B) were categorized as guided investigations (100% of 10 investigations), where as a majority of the investigations for the comparison group teachers (C & D) were categorized as *activities* (76% and 82% respectively). [More detailed data will be included in the full paper.]

The ADI instructional model places a heavy emphasis on writing as a means for students to make sense of science concepts and to communicate their understanding of concepts. Such a large emphasis on students' writing is not necessarily common in traditional approaches to teaching laboratory; therefore, it is important to differentiate between the comparison and treatment groups with respect to the type and amount of writing that occurred in each context. Based on reflective interview data, 100% of the investigations completed in the treatment group required a full laboratory report, where as only 15%, on average, in the comparison group. Additionally, about 73% of the investigations in the comparison group, on average, only required the students to fill out a handout. [More detailed data will be included in the full paper.]

Data Sources

Chemistry Content Knowledge Assessment (ICC = .935): 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 chemistry. Each scenario includes an opening paragraph that provides a relevant context, followed by two questions; one which asks the student to *describe* the fundamental chemistry concept (*Know*) and the other asks the student to *apply* that concept to explain the scenario provided (*Use*). The rubric for this assessment was developed from answers to the questions, which were provided by a practicing chemist.

Scientific Writing Assessment (ICC = .786): The assessment was developed to assess students' ability to communicate in science. 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. The rubric was divided into subscales with components addressing, *Argument Structure*, *Argument Content*, and *Writing Mechanics*.

Chemistry Performance Task Assessment (ICC = .939): This assessment was developed to understand and measure the progress in 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. The scoring rubric focuses on technical and theoretical elements present in each section that relate to the nature of scientific inquiry. There were two subscales for this rubric: *Investigation Design* focusing on the quality of the data collected and the justification of the student's methods (20); and *Quality of the Argument* focusing on the nature of the student's claim, sufficiency of the evidence and accompanying justification of his/her evidence (12).

SUSSI: The 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 2. *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	Chemistry Content Knowledge Assessment
Aspect 2	Students can generate and evaluate scientific explanations and arguments	Chemistry 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	Chemistry Performance Task - Investigation Design Section Scientific Writing Assessment

Data Collection and Analysis

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. The initial subject sample 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. Paired-samples t-tests were used to assess growth over the school year and effect sizes were used to evaluate gains across groups.

Results

Table 3 shows the results for the paired-samples t-tests conducted within each group for each assessment used in this study. Both the comparison and treatment groups made statistically significant gains in terms of their content knowledge. Only the treatment group students made significant gains with respect to their scientific writing abilities and the performance task, except the comparison group students did make significant gains in their argument portion. The comparison group students were the only ones to demonstrate a significant gain in understanding the development of scientific knowledge.

Table 3. *Results of paired-samples t-test analyses for each group on each of the four assessments used in this study*

Assessment	Group	Scale	Pre		Post		t statistic	p	d
			Mean	SD	Mean	SD			
Content Knowledge	Treatment (N=76)	Total	2.30	2.04	11.17	5.48	T(75) = 16.94	<.001*	1.94
		Know	1.43	1.45	5.21	2.83	T(75) = 13.48	<.001*	1.55
		Use	0.87	0.94	5.96	3.08	T(75) = 15.85	<.001*	1.82
	Comparison (N=39)	Total	2.67	2.31	7.00	3.40	T(38) = 7.74	<.001*	1.24
		Know	1.90	1.85	3.95	2.43	T(38) = 4.56	<.001*	0.73
		Use	0.77	0.81	3.05	1.65	T(38) = 8.46	<.001*	1.35
Writing	Treatment (N=93)	Total	17.81	3.62	18.97	3.94	T(92) = 2.40	.018*	0.25
	Comparison (N=60)	Total	17.73	3.34	16.83	4.40	T(59) = -1.47	.148	-
Performance Task	Treatment (N=80)	Total	12.09	4.03	15.30	3.82	T(79) = 5.33	<.001*	0.60
		Investigate[1]	7.30	2.56	8.95	2.54	T(79) = 4.76	<.001*	0.53
		Argument[2]	4.79	2.48	6.35	2.23	T(79) = 4.08	<.001*	0.46
	Comparison (N=46)	Total	12.35	4.66	12.54	5.17	T(45) = .254	.801	-
		Investigate[1]	8.50	2.72	7.67	3.27	T(45) = -1.64	.109	-
		Argument[2]	3.85	2.69	4.87	2.72	T(45) = 2.12	.039*	0.31
SUSSI	Treatment (N=85)	Total	95.56	13.27	96.24	15.27	T(84) = .761	.449	-
		NoSK[3]	35.65	4.47	35.94	4.98	T(84) = .704	.483	-
		DoSK[4]	59.92	10.63	60.29	11.71	T(84) = .490	.625	-
	Comparison (N=55)	Total	86.78	10.86	90.13	12.10	T(54) = 2.81	.007*	0.38
		NoSK[3]	32.55	4.57	33.76	4.94	T(54) = 1.93	.060	-
		DoSK[4]	54.24	8.93	56.36	8.88	T(54) = 2.21	.032*	0.30

*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

An alternative approach to evaluating the gains for each group is through a comparison of the effect size of the students' gains with respect to each aspect of science proficiency. Figure 1 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). Both groups experienced significant learning gains in many areas, however, comparing the relative effect sizes of the learning gains demonstrates the far greater impact the ADI model had on students' development of science proficiency than the traditional model of instruction. For example the ADI model had nearly twice the effect of the traditional model in terms of students' knowledge of scientific explanations and nearly 35% more effect on their ability to use scientific explanations to explain phenomena. Furthermore the treatment group demonstrated small to moderate effects in their ability to communicate in writing and design investigations, but the traditional instruction engendered no effect in those areas.

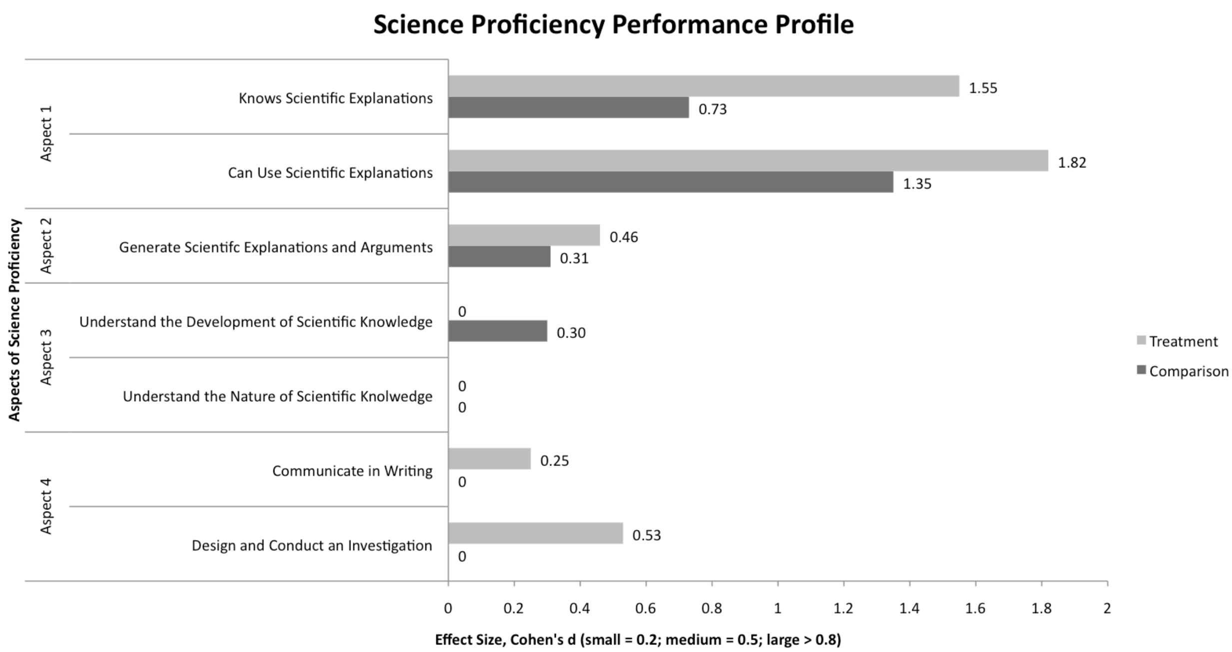

Figure 1. Effect Size of the Intervention for each aspect of Science Proficiency

Conclusions and Potential Implications/Contributions

These results provide evidence for the positive impact of ADI-based instruction on the development of science proficiency in the context of high school chemistry. Significant changes from the beginning to end of the year were noted on all assessments with moderate to large effect sizes in many areas. Aligning the scores from the assessments to related aspects of science proficiency, the findings demonstrate that ADI instruction enhanced the development of some elements of high school chemistry students' science proficiency more than others. Furthermore, this evidence suggests that the ADI instructional model is more effective at enhancing students' science proficiency than more traditional approaches to laboratory instruction. The science proficiency performance profile also demonstrates the importance of using multiple assessments for gaining insight into complex science learning, as opposed to relying on a single measure.

General Interest

The study described in this proposal provides further evidence of the benefits of argument focused science instruction in general. The findings further the research base on the impact of specific argument focused curricula and potential targets for improvement. 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.

Selected References

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.

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.