

# **The Development of Science Proficiency Through Argument Focused Lab Instruction in High School Biology\***

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

Melanie Hester  
Florida State University Schools

Patrick Enderle  
Learning Systems Institute  
The Florida State University

Jonathon Grooms  
Learning Systems Institute  
The Florida State University

## ***Introduction***

This study explored how high school students' level of science proficiency changed after experiencing a set of lab activities that were developed using an instructional model called *Argument-Driven Inquiry* (Sampson, Grooms, & Walker, 2011). The Argument Driven Inquiry instructional model or ADI is designed to serve as a template or guide for teachers and science educators to use when they develop new or modify existing laboratory activities. The intent of this model is to give students more opportunities to develop science proficiency while they are in school by making laboratory activities more authentic and educative.

## ***Theoretical Framework***

Science proficiency, as defined by Duschl, Schweingruber, and Shouse (2007), is 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). From this perspective, individuals who are proficient in science 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.

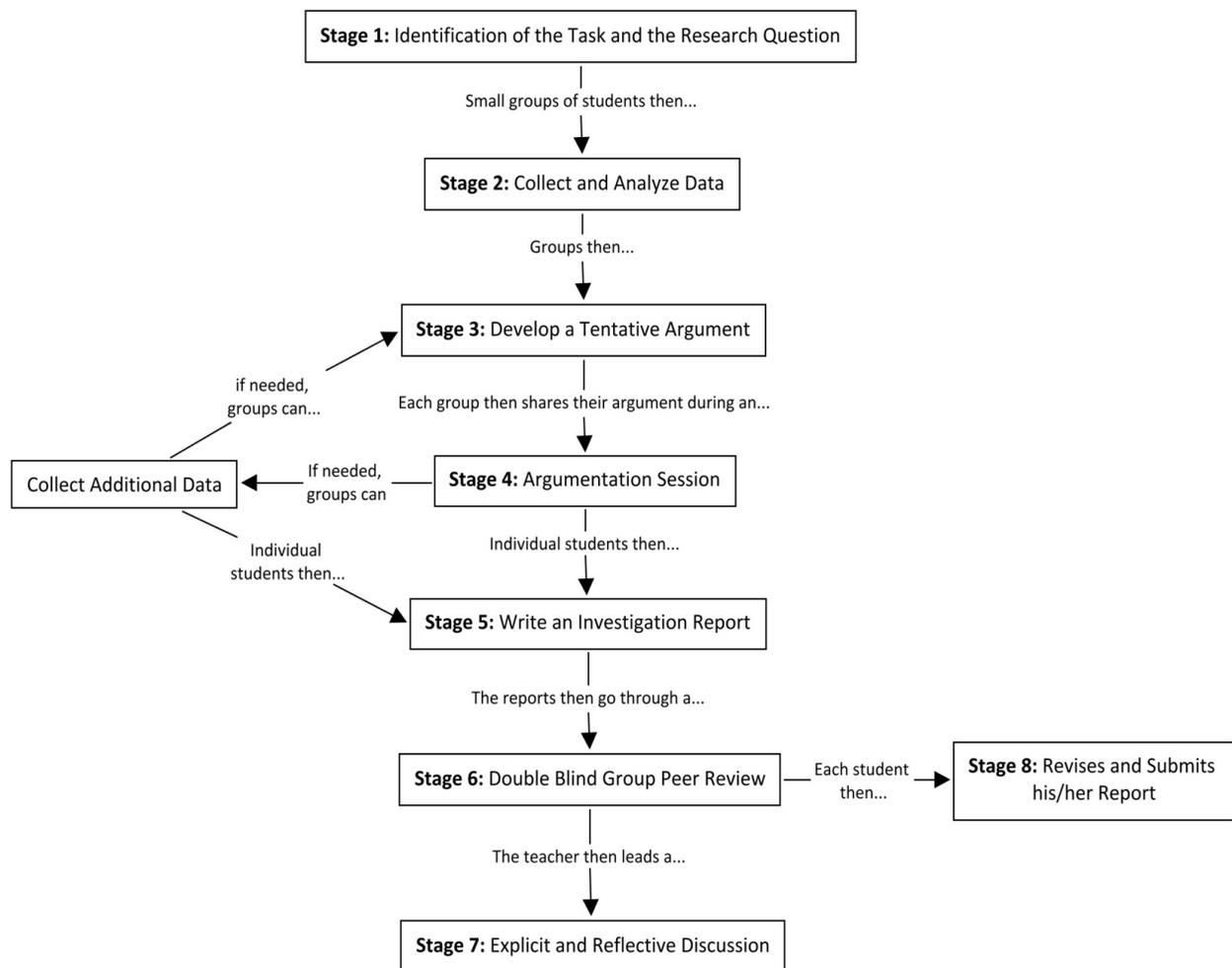
### ***Argument-Driven Inquiry as a way to promote the development of science proficiency***

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

---

\* Reference as: Sampson, V., Enderle, P., Hester, M., & Grooms, J. (2012, April). The development of science proficiency through argument focused lab instruction in high school biology. Paper presented at the annual international conference of the *American Educational Research Association* (AERA). Vancouver, Canada

Council, 2005, 2008). The Argument-Driven Inquiry (ADI) instructional model (Sampson, Grooms, & Walker, 2011) 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 scientific practices such as planning and carrying out investigations, collecting and analyzing data, the generation of arguments, group argumentation, scientific writing, and double blind peer review. Figure 1 describes the stages involved in the original iteration of 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.



**Figure 1:** The original stages of the Argument Driven Inquiry (ADI) instructional model

Teachers and science educators can use the ADI instructional model to transform the nature of laboratory experiences in science classrooms so that they are more authentic and educative for students. 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. We therefore predicted that the students who engaged in laboratory instruction designed using the ADI instructional model

throughout a school year would demonstrate improvement in several aspects of their science proficiency. Figure 2 offers a graphical representation of this hypothesis.



**Figure 2:** Hypothesis describing the potential impact of implementing ADI instruction

### ***Method***

The study described here occurred during year one of a three-year project aimed at assessing students' improvements in science proficiency as a result of experiencing ADI-based instruction (IES Grant #: R205A100909). The project is using an 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. This research setting involved the high school biology course at a university research K12 school. Each class engaged in 14 unique ADI investigations over the course of the 2010-2011 school year. These investigations, which were designed by the project teachers in collaboration with the researchers, focused on seven major big ideas in biology.

### ***Participants***

There were 229 students enrolled in the 9 sections of the biology course at the beginning of the school year. Only 128 of these students, however, agreed to participate in the study. In addition, many of the students who agreed to participate did not complete the full range of assessments because they were absent on one or more of the days when the assessments were given. The demographics of the students enrolled in biology mirrored that of the school (51% Female; 52% White, 29% Africa-American; 10% Hispanic; 11% Free and Reduced Lunch).

### ***Data Sources***

The assessment of a multifaceted construct, such as science proficiency, requires the use of several different instruments and each instrument must be well aligned with each aspect of the construct. The ability to know and use scientific content knowledge to solve and explain problems, for example, is a key aspect 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. Thus, assessments aimed at understanding the development of science proficiency must be crafted with these specific aspects in mind. We therefore developed a set of assessments that were aligned with each aspect of science proficiency. Table 1 identifies the

aspects of science proficiency and the accompanying assessment. In the paragraphs that follow, we describe each assessment in more detail.

**Table 1:** Aspects of Science Proficiency and Associated Assessment

Aspect	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

*Biology Content Knowledge Assessment:* This assessment measures how well a student knows and can use scientific explanations of the natural world. The assessment is comprised of eight free response questions, each related to one of several “Big Topics” in Biology, as determined by the teachers and researchers. Each question includes an opening paragraph that provides a relevant scenario or context, followed by two questions. One question asks the student to *describe* the fundamental science concept (*Know*) and the other asks the student to *apply* that concept to the scenario provided (*Use*). The rubric for this assessment was developed from answers provided for the questions by a practicing biologist. A student’s score on this assessment can range from a low of 0 to a high of 48.

*Scientific Writing Assessment:* The scientific writing assessment was developed to assess students’ ability to communicate in science. 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 who provides an inaccurate explanation for the available data. The students are directed to refute the scientist’s claim and then support a counter-claim with an evidence-based argument while being mindful of writing style and grammar. The rubric, with an overall possible score of 28 points, was divided into three subscales: *Structure* focusing on the inclusion of fundamental argument components including claims, evidence, and rationale (6); *Content* concerning the quality and relevance of the argument components (10); and *Mechanics* regarding the punctuation, grammar, and technical quality of the writing (12).

*Biology Performance Task:* 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. The students must develop an original investigation and make decisions about the appropriate data to collect and evidence to use to generate their argument. These assessments are done in groups of 3-4 students, and the group submits a final product for scoring. The final product includes areas for students to describe the investigation they designed, the data they collected, and the argument they created, along with justification for each of these sections. Initial group composition was maintained as much as possible during separate administrations, and if it was not, the resulting scores were not included in the analysis.

The rubric for this assessment followed the structure of the assessment packet and focused on technical and theoretical elements present in each section that related to the nature of scientific inquiry.

*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. Analysis of these answers assigned raw points to each response in relation to the nature of the item. 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 authors of this instrument originally separated the assessment into several subscales representing major NOS concepts; however, the researchers condensed these subscales into two groups to better align them with Aspect 2 of the science proficiency framework.

### ***Data Collection and Analysis***

All of the assessments were administered at the beginning and the end of the year. A pair of research team members was responsible for scoring each assessment. Each pair scored at least 25% of the full set of each assessment, which had been blinded concerning student identity and pre/post timing. The intra-class correlation coefficient (ICC), a measure of reliability, was determined for each team. 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. The ICC (two-way random effects, absolute agreement) for *Biology Content Knowledge* test was 0.89, the ICC for the *Scientific Writing Assessment* was 0.71, and the ICC for the *Biology Performance Task* was 0.79. Once the assessments were scored, we conducted a series of paired-samples t-tests to determine if students made significant gains on each aspect of science proficiency.

### ***Results***

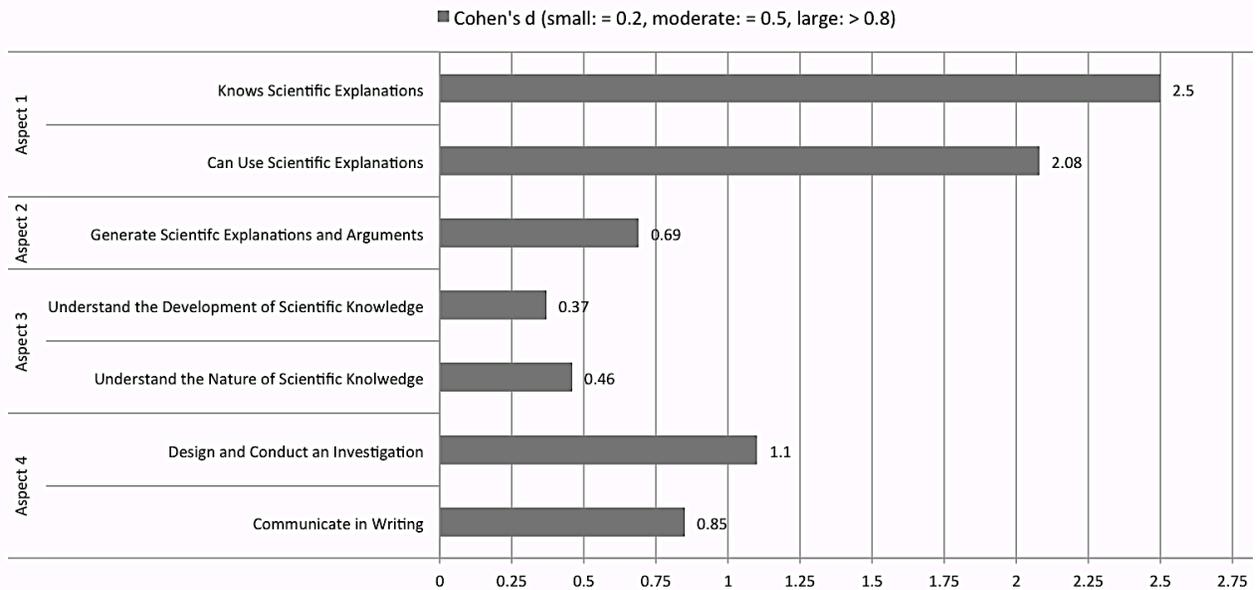
Table 2 presents the pre and post-intervention scores for each assessment. It also includes the result of each paired sampled t-test. The results of these tests indicate that all the observed gains on all four assessments were statically significant. We therefore decided to examine scores on each assessment by subscale as well. For the *Biology Content Knowledge* assessment, significant gains were made on both the *Know* and *Use* subscales. For the *Scientific Writing* assessments significant gains were observed on the *Structure, Content, and Mechanics* subscales. The *Biology Performance Task* was split into subscales concerning *Argument Development* and *Investigation Design* and the students made significant gains on each one. Finally, the SUSSI was split with regard to questions about the *Development* and the *Nature* of science knowledge and the students made significant gains on both subscale. Effect sizes for each of the pre/post comparisons were determined by calculating Cohen's *d*, a metric where the categorical range is: small = 0.2, moderate = 0.5, and large = 0.8 or higher (Cohen, 1988).

One way to examine the impact of this type of instruction on the development of science proficiency is to compare the intervention effect size for each aspect. The effect size is used to examine the magnitude of an observed difference in scores in standard deviation units. Thus, a significant difference in mean scores that has a large effect size indicates a meaningful change in the students' knowledge or skills. A graphical representation of the effect size of the intervention

on each aspect of science proficiency is therefore provided in Figure 3. This type of instruction had a rather large effect ( $d \geq 0.8$ ) on aspects 1 and 4 of science proficiency. The intervention, however, only resulted in a moderate effect ( $0.5 \leq d < 0.8$ ) for aspect 2 (ability to generate scientific explanations and arguments) and small to moderate ( $0.2 \leq d < 0.5$ ) effect on aspect 3 (understands the development and the nature of scientific knowledge).

**Table 2:** Results for Paired Samples t Tests of Overall and Subscale Assessment Data

Assessment	Pre		Post		<i>t</i>	<i>df</i>	<i>p</i>	<i>d</i>
	M	SD	M	SD				
<b>Biology Content Knowledge</b>								
Overall	4.45	3.74	17.89	6.51	21.78	110	< .001	2.51
<i>Know</i>	1.57	1.64	8.93	3.83	20.77	110	< .001	2.50
<i>Use</i>	2.90	2.54	8.96	3.24	20.81	110	< .001	2.08
<b>Scientific Writing Assessment</b>								
Overall	18.80	3.90	22.20	4.07	8.74	127	< .001	0.85
<i>Structure</i>	3.16	1.28	4.13	1.43	7.42	127	< .001	0.71
<i>Content</i>	4.86	2.23	6.95	2.17	8.67	127	< .001	0.95
<i>Mechanics</i>	10.68	1.36	11.12	1.23	3.23	127	.002	0.34
<b>Biology Performance Task</b>								
Overall	14.11	5.23	19.64	4.79	5.5	27	< .001	1.1
<i>Argument Development</i>	4.57	2.04	5.93	1.92	3.08	27	.004	0.69
<i>Investigation Design</i>	9.54	3.91	13.71	3.98	5.0	27	< .001	1.10
<b>SUSI</b>								
Overall	85.77	10.91	91.33	12.00	4.47	118	< .001	0.48
<i>Development of Science Knowledge</i>	52.79	8.82	56.15	9.19	3.64	118	< .001	0.37
<i>Nature of Science Knowledge</i>	32.98	4.53	35.04	4.35	3.78	118	< .001	0.46



**Figure 3:** Effect Size of the Intervention for each aspect of Science Proficiency

### ***Conclusion and Discussion***

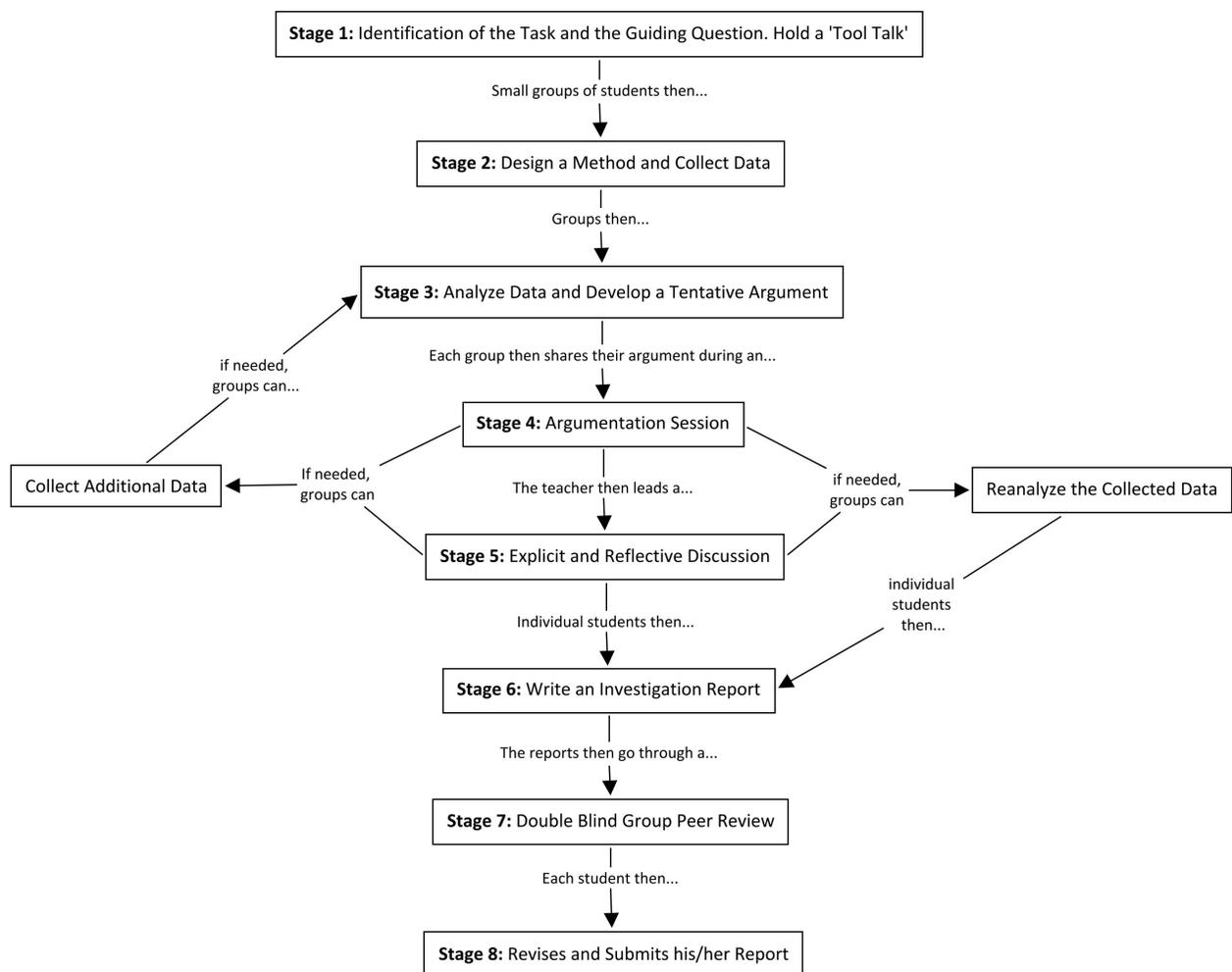
These results provide some evidence of the potential and promise of the ADI instructional model as a way to help students develop science proficiency in the context of high school biology. Significant changes from the beginning to end of the year were noted on all assessments with moderate to large effect sizes except for the SUSSI. It is important to note, however, that the nature of the research design used in this study only provides support for a correlational, not causal, relationship between ADI instruction and the development of science proficiency. Yet, this does not mean that the results reported here should be dismissed outright; the results simply suggest that ADI *might* be able to provide teachers with a way to help students develop science proficiency and that future research that examines the efficacy of the model in comparison to other approaches is warranted. This type of finding is also important, we argue, given the substantial amount of literature which indicates that most students do not become proficient in science as a result of their high school education (NRC, 2005; 2008; 2011).

Our findings also support the validity of the hypothesis underlying this study and the overall research project because the outcomes we observed were consistent with our initial predictions. The biology students, in other words, exhibited substantial growth in their knowledge of and ability to use scientific explanations, their ability to design and conduct investigations, and their ability to write in a scientific manner when they participated in a series of lab activities that were designed to be more authentic and educative. The students also showed significant improvement in their ability to generate scientific explanations and arguments and their understanding of the nature and development of scientific knowledge but the magnitude of growth for these two aspects was not as large. This observation suggests that the ADI instructional model did not have an equal impact on all four aspects of science proficiency. In the paragraphs that follow, we will provide some potential explanations for the smaller effect sizes and outline our plans to address it as part of our future research.

In order to score high on the argument generation component of the performance task, the students needed to be able to develop a claim that answers the research question and the claim needed to be accurate. The students also needed to be able to support their claim with evidence and provide a rationale that explains why the evidence was important. Our analysis of the arguments generated by the students at the beginning of the year indicated that students were able to provide an adequate claim in response to the research question and support it with weak evidence. Most of the claims, however, were inaccurate. The students also did not explain why they included the evidence in their arguments. At the end of the year, the students were able to provide an argument with an accurate claim and strong evidence but were still not including a rationale. This issue, however, is not surprising given the available literature, which suggests that students often do not think that it is necessary to explain their choice of evidence to others in their arguments (Berland & McNeill, 2010; Berland & Reiser, 2009; McNeill & Krajcik, 2008). The teachers also commented that students often did not understand what they wanted when they asked the students for their rationale during the various stages of the model. To address this issue, we plan to change the term “rationale” in our argument framework and all instructional materials to “justification of evidence” in order to help students develop a better understanding of this important component of a scientific argument.

The small but significant growth in SUSSI scores observed during this study is not surprising in light of findings from other research that has established that developing an

understanding of the nature and development of scientific knowledge, which are rather abstract concepts, is often challenging for students (Hodson, 2009; McComas, 1998). We were also not surprised by the amount of growth given the limited amount of time the teachers devoted to nature or development of scientific knowledge during the reflect and explicit discussion stage of the model (stage 7, see Figure 1). The NOS/NOSI research literature has clearly demonstrated that students do not make substantial gains in their understanding of these abstract concepts when their teachers do not engage them in an explicit and reflective discussion about the concepts during a lesson (Abd-El-Khalick, Bell, & Lederman, 1998; Akerson, Abd-El-Khalick, & Lederman, 2000). In response to this observation, the teachers claimed that placing the explicit and reflective discussion stage towards the end of the model was too late for meaningful learning and discussion to occur. In response, this stage of activity has been purposefully moved to Stage 5 (see Figure 4), directly following the argumentation session and preceding the initial drafting of the investigation report. The project teachers felt this was a more sensible placement since they observed that these kinds of discussions more naturally emerged at this point in the model. The second year of data collection will provide insight as to the impact of this change.



**Figure 4.** Current version of the ADI instruction model

### ***Implications***

The study provides further evidence of the benefits of argument focused science instruction. The findings further the research base on the impact of a specific argument focused instructional model and potential targets for improvement. The use of argumentation inside the classroom and instructional models that promote and support these types of interactions can be useful in teaching not only science content, but also enhancing students' abilities in other areas important to learning and understanding science. By engaging students in authentic science practices, such as designing and conducting investigations, defending and critiquing scientific arguments, and coordinating claims, evidence, and justifications into scientific arguments communicated through writing, and providing them with educative feedback about their performance throughout a school year, students will not only enhance their potential to be scientists, but also improve their status as a global citizen with proficiency in understanding and evaluating that vast amount of information and public opinion increasingly shaped by science.

With the development of students' science proficiency as the current goal of science education, this study also contributes to the research base on ways of assessing the various aspects of such a broad and complex construct. The performance profile we used in this study to examine student learning also demonstrates the importance of using multiple assessments for gaining insight into what students know and can do, as opposed to common standardized measures that tend to focus on knowledge of facts and concepts. The development of a science proficiency performance profile, such as the one used in this study, can serve as one approach to understanding the learning of the concepts, skills, and habits of mind needed to be proficient in science in response to an intervention. This type of focus will be important in the years to come because the assessments that shape the educational system must come to reflect the current goals of the educational system and be able to provide a meaningful measure of student learning.

This study also highlights how teachers will implement an intervention in ways that may be inconsistent with the original intent of the researchers. Educational researchers, as a result, will need to acknowledge that teachers, and not researchers, are the ones that are ultimately responsible for the implementation of all educational interventions. All interventions will therefore be modified and adapted as it is implemented by a teacher based on how that teacher prefers to teach, what that teacher values, and the unique nature of the district, school or classroom where the intervention is being implemented. Researchers will need to be mindful of these issues when attempting to study the impact of any educational intervention on student learning and account for this issue in their research designs.

### ***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***

Abd-El-Khalick, F., & Lederman, N. G. (2000). Improving science teachers' conceptions of nature of science: A critical review of the literature. *International Journal of Science Education*, 22, 665-701.

- Akerson, V., Abd-El-Khalick, F., & Lederman, N. (2000). Influence of a reflective explicit activity-based approach on elementary teachers' conception of nature of science. *Journal of Research in Science Teaching*, 37(4), 295-317.
- Berland, L., & McNeill, K. (2010). A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts. *Science Education*, 94(5), 765-793.
- Berland, L., & Reiser, B. (2009). Making sense of argumentation and explanation. *Science Education*, 93(1), 26-55.
- 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.
- Hodson, D. (2008). *Towards science literacy: A teachers' guide to the history, philosophy, and sociology of science*. Rotterdam, The Netherlands: Sense Publishers.
- Landis, J. R. and Koch, G. G. (1977). A one-way components of variance model for categorical data. *Biometrics* 33: 671-679.
- Lederman, N., G. (1992). Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of Research in Science Teaching*, 29(4), 331-359.
- Liang, L., Chen, S., Chen, X., Kaya, O., Adams, A., Macklin, M., et al. (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).
- McComas, W. (1998). The principle elements of the nature of science: Dispelling the myths. *The nature of science in science education* (pp. 53-70). The Netherlands: Kluwer Academic Publishers.
- McNeill, K., & Krajcik, J. (2008). Scientific explanations: Characterizing and evaluating the effects of teachers' instructional practices on student learning. *Journal of Research in Science Teaching*, 45(1), 53-78.
- 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.
- 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.