NEW INSTRUMENTS THAT CAN BE USED BY RESEARCHERS TO ASSESS THREE DIFFERENT ASPECTS OF SCIENCE PROFICIENCY

Victor Sampson, Jonathon Grooms, Patrick Enderle

Background/Context

Scientific proficiency, as described by (Duschl, Schweingruber, & Shouse, 2007), encompasses 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 et al., 2007; National Research 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 data collection and analysis, argument generation, group argumentation, scientific writing, and double blind peer review processes. The ADI instructional model therefore 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. However, the dilemma of assessing students' scientific proficiency or measuring changes in science proficiency over time still prevails. Typical paper/pencil, multiple choice type assessments are not adequate for assessing the many facets of scientific proficiency (Duschl et al., 2007; National Research Council, 2001, 2008). New assessments, therefore, are needed in order assess students' current levels of scientific proficiency and how they develop different aspects of science proficiency over time.

Purpose/Objective/Research Question/Focus of Study

This proposal describes three new and innovative assessments that are being used within a larger study to measure students' scientific proficiency. Multifaceted constructs such as science proficiency require a variety of tools to assess students' knowledge and abilities related to science. The ability to know and use scientific content knowledge to solve and explain problems is a key component of scientific proficiency, which 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. Using a single assessment or inappropriate assessments over estimate or even under estimate students' scientific proficiency depending on the nature of the assessment. The three assessments, which will be discussed in this proposal, were designed to measure students': (1) understanding of science content and their ability to use those understandings to explain; (2) understanding of the nature and development of scientific knowledge; and, (3) abilities to generate and evaluate scientific arguments.

Setting and Participants

The assessments described in this proposal are being pilot-tested by six teachers ranging from grade seven through grade twelve within a K-12 University Research School. This pilot testing is taking place 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 pilot testing of the model and the assessments are taking place in middle school life science, middle school physical science, high school biology, and high school chemistry courses. There are approximately 600 student participants involved in this research project, whose responses will be used to determine the validity and reliability of the assessments.

Intervention/Program/Practice

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 three assessments of science proficiency discussed in this proposal are being administered on three occasions during the 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 allows 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. The assessments used as part of this study, and their development, will be described in the following section.

Research Design

The assessments discussed in the proposal focus on three of the four key aspects of scientific proficiency. Students' abilities to participate in the practices and discourse of the scientific community in a productive manner, the fourth aspect of science proficiency, is inherent in each of the assessments; however, it is not easily assessed using a paper/pencil style assessment. This aspect of scientific proficiency is being assessed using the Assessment of Scientific Argumentation in the Classroom observation protocol, which has already been developed and validated (Enderle, Walker, Dorgan, & Sampson, 2010). Table 1 identifies the aspects of science proficiency and the accompanying assessment that will be addressed in this proposal. The three assessments were developed by the research team to both align with particular aspects of scientific proficiency as well as the "Big Ideas" of a particular domain of science. Therefore, four versions of each assessment were created to accommodate the Big Ideas of each subject area that is a part of this study: middle school life science, middle school physical science, high school biology, and high school chemistry. Furthermore, to account for potential testing effects, three slightly different versions of each assessment were created to use during the pre-, mid-, and post-intervention data collection periods. This proposal addresses the general features of each type of assessment and uses high school biology as a representative sample of all the assessments.

Table 1. *Alignment of assessment instruments with science proficiency constructs*

Aspect of Science Proficiency	Assessment Instrument
Students know, use, and can interpret scientific explanations of the natural world	Topic-specific Content assessment
Students can generate and evaluate scientific explanations and arguments	Topic-specific Writing assessment Topic-specific Performance Task
Students understand the nature and development of scientific knowledge	Topic-specific Performance Task

Topic-specific Content Assessment. Each content knowledge assessment is comprised of eight free response questions, with a question related to each Big Idea identified for the corresponding course. Each question begins with the presentation of a background scenario and/or information coupled with some type of graphical representation (ex – a graph of reaction data, an image of the organism discussed, a map of an area described, a diagram of a system, etc.). This information provides a specific context for the students to apply their understanding of the related Big Idea.

Following the opening paragraph in each question, students are presented with a set of two questions. The first question in each set asks the student to *describe* the science concept underlying the question related to the Big Idea, followed by the second question, which asks the student to *apply* their understanding of that science concept in order to explain the information provided in the introductory scenario. Structuring the questions in this way allows us to assess the *know* and *use* aspects of scientific explanations, which as noted earlier, are an indicator of science proficiency. Figure 1A provides a sample question from the high school biology content assessment.

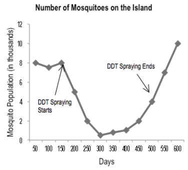
Once the content knowledge assessments were fully developed, an expert consultant reviewed the assessments suitable to their expertise. The consultants were asked to verify the appropriateness of the questions in relation to the associated Big Idea; and they were also asked to provide an "expert answer" to each of the questions. These answers were used to construct itemized rubrics that will be used to score student responses to each question.

Mechanism of Evolution

6. After World War II, the insecticide DDT was used liberally to control populations of various insect pests, including malaria-spreading mosquitoes. Scientists have since discovered that some mosquitoes are resistant to DDT. Mosquitoes that are resistant to DDT are able to make an enzyme called esterase, which breaks down DDT into a harmless chemical. Non-resistant mosquitoes either do not produce this enzyme or produce it in very low amounts.

   The graph below shows a population of mosquitoes on a tropical island. The island was sprayed with DDT over a period of several months. Measurements of the number of mosquitoes on the island every 50 days resulted in the information shown on the graph below.

Number of Mosquitoes on the Island



Biologists have used the theory of *Natural Selection* to explain similar observations in other contexts.

a. **Describe** the process of *Natural Selection.*

b. **Use** the process of *Natural Selection* to **explain** the change in the size of the population of Mosquitoes over time on this tropical island.

A

Introduction: Much of the food we eat is turned into *glucose* and is carried in the bloodstream for our bodies to use for energy. A hormone produced by the pancreas called *insulin* helps sugar in our blood move into the cells of our bodies. If your body doesn't make enough insulin or if the insulin doesn't work the way it should, blood sugar can't get into your cells. Instead, the glucose stays in your blood, raising your blood sugar level, and decreasing the amount of energy available to your cells.

A person is considered to have diabetes if their blood sugar level (after fasting) is over 120 mg/dl. Examine the following data table. It provides medical information about six different fifth graders.

Student	Mass (lb)	Height (ft)	Ratio of Mass to Height (Lb/in)	Diet	Activity Level	Family History of Diabetes?	Blood Glucose Level (after fasting (mg/dl)
A	175	5'3"	2.8	High Calorie	Low	Yes	150
B	150	5'11"	2.11	Moderate Calorie	High	Yes	120
C	190	5'5"	2.9	High Calorie	Moderate	Yes	140
D	140	4'10"	2.4	Low Calorie	Low	No	160
E	160	5'7"	2.4	High Calorie	High	No	118
F	180	5'5"	2.8	High Calorie	Low	Yes	165

A local doctor thinks that a child's weight is what causes a child to develop type II diabetes. This is his/her argument:

> The weight of a child is the most important factor in causing type II diabetes. Three of the four heaviest children in this group had high blood glucose levels (Over 120 mg/dl). Therefore a significant weight gain causes diabetes.

Directions: Examine the data table above and think about what you know about the factors influencing Type II diabetes. Then write an essay to convince the doctor that the ratio of height and weight is more likely the cause of type II diabetes. As you write your essay, remember to:

- Discuss the doctor's argument and then provide evidence and/or other reasons to demonstrate why the claim is inaccurate;
- Clearly state your claim and then provide adequate evidence and a sufficient rationale to support it;
- Present your ideas in a clear and logical order, including an introduction, body, and conclusion;
- Use a variety of words and well-constructed sentences to create tone and voice; and,
- Use proper punctuation, capitalization, sentence formation and spelling.

You will have one class period (but no more than 60 minutes if your class is longer) to plan and write your essay.

B

The Task: Comparative anatomy is the study of different organisms' bodies. During this kind of work, scientists will look at the structure and shape of different organisms' bodies, including such areas as hands/paws, skulls, chest size, or leg placement on the body. These investigations include looking at these parts and making detailed observations about which features are more or less similar in different organisms. Organisms with similar body features often live in similar habitats, eat like foods, or behave in similar patterns. Understanding the similarities and differences among organisms helps scientists understand how these organisms live together in ecosystems and how they are related to each other.

In this investigation, you will be given several anatomical specimens from five different animals. Several of these specimens have been labeled with the organism that they come from, and one specimen comes from an unknown animal. You and your group will have one class period (but no more than 60 minutes if your class is longer) to plan and conduct your investigation.

The guiding question of this investigation is: *What kind of animal did the unknown specimen come from?*

Part 1: Generate the data you need to develop a scientific argument that answers this research question. You can use any of the following materials:

- 1 footprint from an unknown animal (A)
- 2 footprints from wild cats (B and C)
- 2 footprints from wild dogs (D and E)
- 1 Ruler
- 1 Protractor
- 1 Electronic balance
- Tracing Paper
- 1 Magnifying Glass
- Stereoscope
- A highlighter
- Calipers

1. Record your observations and/or measurements in the space below.
2. Why did your group decide to make these observations and/or measurements?

Part 2: Answer the following questions.

1. Describe the method your group used to answer the research question. Be specific
2. Why did your group decide to design your investigation in this manner? Please provide a justification for the method your group decided to use.
3. Identify one strength of the method your group used and explain why it is a strength.
4. Identify one limitation of the method your group used and explain why it is a limitation.
5. What is your argument? Provide your group's answer to the research question and then support it with genuine evidence and a sufficient rationale. Once your group has completed your argument, use a highlighter to identify your evidence.

C

Figure 1. (A) A sample item from the high school biology content assessment, (B) a sample prompt from the high school biology scientific writing assessment, and (C) a sample task from the high school biology performance task assessment

Topic-specific 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/doctor who provides an inaccurate explanation for the data. The students are then directed to respond to the scientist's claim by arguing in support of a countering claim, being sure 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.

During this assessment the students are provided with several pieces of lined paper to help organize their writing. The students are initially asked to engage in a pre-writing activity to outline their argument and then generate a rough draft. By the conclusion of the assessment period the students are expected to have refined any initial drafts or pre-writing exercises in order to provide a final draft of their argumentative essay addressing the task indentified for the assessment. A sample prompt from a topic-specific writing assessment is provided in Figure 1B.

Topic-specific Performance Task. The performance task assessment was developed to understand and measure the progress in students' abilities to develop an investigation in response to a research question, making decisions about the appropriate data to collect and evidence to use in creating an argument that answers the research question. By completing a task like this students are able to demonstrate, in an authentic context, their understanding of the nature and development of scientific knowledge, as well as their ability to generate scientific arguments and explanations. For each course, the performance tasks were developed such that they were not specifically related to any of the Big Ideas or curricular standards, but dealt with material from the discipline, thus providing a novel context for the assessment. These assessments are done in groups of 3-4 students, and the group submits a final product for scoring. Groups are provided with a short paragraph describing some of the science concepts underlying the activity (ex – harmonic motion, comparative anatomy). Following this, a research question is given along with a list of materials that are available at the physical station where the group is performing this assessment. The assessment packet then has a set of open-ended questions, which ask the students to record and explain their decision-making processes and the results of them. Once students have completed the task and developed their groups' argument, they are asked to highlight particular elements of their argument for scoring clarity. An example of a topic-specific performance task is provided in Figure 1C.

Data Collection and Analysis

At the time of this proposal the data collection process is currently in progress, with the post-intervention data to be collected in late May 2011. Once the data set is complete the research team will use the rubrics that have been developed to assess the validity and reliability of the three assessments. During that process inter-rater reliabilities and item analyses will be determined in addition to the construct validity that has been generated by the content expert reviewers. It is anticipated that the complete data set will represent approximately 600 students, each having completed the three assessments above. The students' scores on these assessments will provide insight as to their levels of scientific proficiency with regard to the specific characteristics discussed above as well as providing measures of how their scientific proficiency has changed as a result of the ADI intervention.

Findings/Results

We will be able to report on the validity and reliability of the new instruments once data collection and analysis is complete in June 2011.

Conclusions/Potential Implications

We anticipate that the three new assessments will provide valuable data to help measure students' scientific proficiency, which is an important outcome of science instruction. The use of assessments such as these in other research project will provide a better measure of what students know and are able to do than traditional multiple-choice tests. Given the anticipated findings based on these assessments, we hope to show that similar assessments will be useful in measuring complex constructs that are valued in large-scale research projects.

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.

Enderle, P., Walker, J., Dorgan, C., & Sampson, V. (2010). *Assessment of argumentation: An observation protocol.* Paper presented at the 2010 International Conference of the National Association for Research in Science Teaching (NARST).

National Research Council. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, D.C.: National Academy 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.

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.