ARTICLE

DETERMINING PATHS TO SUCCESS: Preparing Students for Experimental Design Questions on Standardized Tests

MATTHEW J. TURNER JOSE M. RIOS

Recent education reform efforts are at the forefront of educators' minds across the nation, science teachers notwithstanding. At least 48 states have developed a mandated standardized test, the majority of which also publish an individual school proficiency report (Olsen, 2001). Washington State's new standardized science test is an example of such reforms efforts. The Washington Assessment of Student Learning in Science (Science WASL), which is administered at the fifth, eighth and tenth grades, specifically measures science content and process skills using questions from earth, physical, and life science courses (Partnership for Learning, 2003). In addition to content knowledge, the Science WASL requires that students think critically, solve problems, apply reasoning skills, and design novel laboratory investigations (OSPI, 2003). The Science WASL is one of several mandated tests that students will have to pass by the year 2010 to earn their Certificate of Academic Achievement in Washington State, (OSPI, 2005). Standardized science tests in other states have similar requirements, drawing both on content knowledge and the critical thinking skills necessary for designing experiments from given scenarios.

To best prepare students to take such an assessment, science educators must make careful decisions about teaching both science content and process skills. Many educational researchers contend that students learn skills best and gain better attitudes toward science through authentic, inquiry-based science instruction (Freedman, 1996; Udovic et. al., 2002; Gibson & Chase, 2002; Ad-Marbach & Sokolove, 2000). Inquiry science lessons, according to the National Research Council (1996), are:

> *Multifaceted activities that involve making observations; posing questions; examining books and other sources of information to see what is already known; planning investigations; reviewing what is already known in light of experimental evidence; using tools to gather, analyze, and interpret data; proposing answers, explanations, and predictions; and communicating results. Inquiry requires identification of assumptions, use of critical and logical thinking, and consideration of alternative explanations.* (p. 23)

Successful implementation of authentic inquiry, however, is no easy task for even the most experienced teachers. Though the National Research Council (1996) mandates teaching science through inquiry, science educators are often reluctant to deviate from more traditional methods (Yerrick et. al., 1997). Lack of funding, poorly selected curriculum materials, insufficient laboratory preparation time, and discomfort with scientific methodology can be difficult issues for science educators to overcome (Chinn & Malhotra, 2001).

An advantage of inquiry instruction quickly surfaces when one delves deeper into the available research. Students who received science instruction through inquiry may have an advantage on standardized tests over those who have been taught through more traditional methods (Schneider et. al., 2002). Student achievement on such measures has never been more important. Recent mandates such as the *No Child Left Behind Act* (2003) will directly affect how federal monies are allocated to schools that fail to meet standards. Ultimately, the reputation of the school and the jobs of administrators and teachers may hang in a precarious balance according to how well students perform on state mandated standardized tests. Given the high-stakes nature of such tests, and the research-based effectiveness of inquiry, it seems imperative that science educators implement more inquiry activities within their classrooms and examine their relationship to student performance.

Purpose

This classroom-based study focused on whether the inquiry-based activities would enhance students' ability to design laboratory investigations such as those presented on the Science WASL. This task was accomplished by completing an in-depth exploration of an inquiry-based instructional unit. The goal was to determine the effectiveness of unit assignments in increasing both the students' content knowledge of the scientific method and WASL laboratory design question skills.

MATTHEW J. TURNER is a science teacher at Spanaway Lake High School, Spanaway, WA 98387-6828; e-mail: mturner@bethelsd.org. JOSE M. RIOS, Ph.D., is Interim Co-Director and Associate Professor, Education Program, University of Washington, Tacoma, Tacoma, WA 98402: e-mail: jrios@u.washington.edu.

Figure 1. Names and Descriptions of Pre/Post-Assessments and Inquiry-Based Activities.

Biology Knowledge Pre-Assessment. Mr. Turner assigned this pre-assessment to collect baseline data on what general biological concepts students had covered in previous classes and what they could remember off the top of their heads about it. This piece essentially gave a baseline for students' degree of specific inquiry skills, namely observation, graphing, measuring, and experimental design ability.

Metric Measurement Lab. This lab reviewed measuring and weighing of various objects in metric units of meters, centimeters, and millimeters. Students worked in pairs to measure lab tables and test tubes, mass several objects, and found the volume of a test tube of water using a graduated cylinder. After completing this piece, students were asked to make simple measurements required in subsequent labs.

Graphing Skills. An oral review of effective graphing strategies preceded a graphing activity. Students were asked to analyze graphs, and then construct a line graph and a bar graph from two different data tables. This activity reviewed proper graph construction so that students would be ready to construct graphs from data they collected within subsequent inquiry labs.

Lab Skills Review. This curriculum piece focused on taking averages of given data sets, and also introduced data table construction. Both skills were needed to successfully complete subsequent inquiry lab activities.

Scientific Method Lecture Notes. This short lecture provided students the individual steps of the scientific method, which served as the base vocabulary for subsequent lab activities.

Tootsie Pop Lab. Students wrote the question they remembered from the original Tootsie Pop television commercial and made observations on the candy sucker. They defined a method to lick, defined a stopping point, and then proceeded to collect data by licking the sucker at home. They brought back the number of licks it took to get to the Tootsie Roll center, and participated in a follow-up discussion of the variables involved. Students wrote a paragraph detailing how they could eliminate or control all variables and arrive at a succinct answer to the question. The focus of this assignment was to explore the importance of recognizing and controlling variables. Students were required to do both within the subsequent inquiry lab activities in order to feasibly answer the questions they posed.

Scientific Method Pop Quiz. To determine if students understood the parts of the scientific method, Mr. Turner gave a five-minute pop quiz. Students were asked to list the steps of the scientific method in the order in which scientists typically carry them out.

Quicker-Picker-Upper Lab. The first guided-inquiry activity, this lab involved a study of different brands of paper towels. Mr. Turner helped students define a fair test of ten different paper towels by combining water absorption and cost data. Students wrote a title, a purpose statement, and documented observations. Furthermore, they wrote hypothesis statements, made data tables and graphs, and then wrote an analysis section explaining which towel was the best value. Essentially, all of the targeted inquiry skills and the parts of the scientific method were explored using inquiry strategies.

Organism Observation Lab. Mr. Turner's classroom pets were the focus of the next inquiry lab. This piece was a visual prelude that helped students develop an original investigation of their choice. Students were given a choice of 12 different organisms and were required to sketch four, write 40 observations, generate 20 questions, and formulate 12 potentially testable hypotheses. Hypotheses written in the form of an "If. . . .then" statement were reviewed and stressed, as this is how students were required to write all hypotheses throughout the course of this unit.

Organism Investigation Lab. The Organism Lab was an independent research project on the same organism from the Organism Observation Lab. Students started from scratch to write titles, purpose statements, and further their observations. Following these activities, they wrote a testable hypothesis, a reproducible method, made data tables and graphs, and wrote analyses, all with very little help from Mr. Turner.

Video: The Learning Channel Mavericks of Science. Mr. Turner reviewed the scientific method with students before they viewed this video. During the video, students noted what part of the movie either showed or discussed scientists' observations, hypothesis, data collection method, analysis method, data reporting and experiment redesign.

Post-Assessment of Scientific Method, Inquiry Skills, and Lab Design. The post-assessment piece mirrored that of the pre-assessment, save for the fact that the author did not ask for information about units not yet covered. Students were again asked to list and describe as much as they could about the scientific method, make graphs, measure lines, and then design an investigation for the same scenario as that found on the pre-assessment. Coupled with the pre-assessment, this piece allowed a quantitative comparison of student skill level before and after the inquiry unit had been carried out. Moreover, this piece allowed us to gain a clear picture of the problem areas within the unit as well as determine the degree of improvement in students' inquiry skills following the intervention.

Methods

Participants

One hundred one sophomore students from a large suburban high school participated in this study. The student body was diverse: 1.5% Native American, 15.0% Asian American, 15.5% African American, 5.0% Hispanic, and a 63% Caucasian population. The socioeconomic status of students in the building varied widely, with 50% of students eligible for free or reduced priced meals. The largest employer in this community of 21,500 people was the school district itself, though two large military installations were in 10 miles of the campus. Nearly 30% of students had one or both parents employed in the military. As a result, students withdrew or enrolled in accordance with their parents' duty status and duration of assignment on post.

Mr. Turner taught the four biology classes, collected assignments, and analyzed student work as data sources. He facilitated student learning, which included lecturing, leading small group discussions, providing demonstrations, and supervising laboratory activities. Throughout the course of this study, he helped students construct knowledge as they experienced concepts and gained inquiry skills through a wide variety of hands-on lessons and laboratory activities.

Evaluation Tools

The materials used in this study were part of the district-adopted curricula for biology classes. Some of the activities found within the lab manuals were modified to ensure that specific inquiry skills were presented and practiced (see Figure 1). Once completed by students, these assignments and both a pre- and post-assessment were analyzed as the main sources of data. The pre- and post-assessment tool can be found in Figure 2.

Figure 2. The pre-test and post-test for the inquiry-based instructional unit.

CONCEPTS IN BIOLOGY PRE-ASSESSMENT

Completing the following questions to the best of your ability will allow me to assess your prior knowledge of some of the major biological concepts we will study. Answer each question as thoroughly as you can, based on what you remember from previous classes. You will receive full credit on this piece if and only if you complete every question — in other words do not leave questions unanswered. Please do your best as this information helps me to help you!

Please tell me what you know about the following:

1. Classification of Organisms/Scientific Naming

2. Cells

3. Photosynthesis

4. Respiration

5. Atom Structure/Chemistry

6. Ecology/Environmental Problems

7. Genetics/Heredity

8. Evolution

9. The Scientific Method

10. How long is this line ________________________________ in cm?
    In mm? In inches?
    Centimeters (cm) =
    Millimeters (mm) =
    Inches (in) =

11. Over the summer, Mr. Turner caught 12 perch, 23 trout and 15 bass on his fly rod. Construct a complete, vertical bar graph below from this information.

12. Construct a complete line graph below from the information given on the growth of three different plants over several days.

	Day 1	Day 2	Day 3	Day 4	Day 5
Plant 1 Growth (cm)	2	5	7	9	11
Plant 2 Growth (cm)	1	2	3	4	7
Plant 3 Growth (cm)	4	7	10	11	12

13. Dr. Ratatuey is one of the research scientists at Acme Pharmaceutical Corporation. He wants to see if female mice given Drug X can run through a maze more quickly than male mice given Drug X. Write a description of how the good doctor should design his investigation. Be sure to include:
    - Hypothesis
    - Control for at least two variables
    - Procedure
    - How data will be collected

    Use words, labeled pictures, and/or labeled diagrams in your description.

Procedure

Two sophomore honors biology classes (Honors A & B) and two general biology classes (General C & D) participated in the study. The unit activities are listed in the order that they were assigned to students (see Figure 1). The targeted inquiry skills in this unit were observation, hypothesis writing, procedure writing, graphing, measurement, data collection, and experimental design. Each unit activity focused on specific inquiry skills.

The order of delivery allowed students to learn and practice the inquiry skills needed for future unit activities. We felt that using a scaffolded approach, where students are gradually exposed to more difficult learning tasks, can be an effective means of increasing students' inquiry skills (Gable, 2001). Introductory unit activities included:

- The Metric Measurement Activity
- Graphing Skills Activity
- The Lab Skills Packet
- The Scientific Method Notes.

The application pieces, where students applied the various skills in context, included:

- The Tootsie Pop Lab
- The Quicker-Picker-Upper Lab
- The Organism Observation Lab
- The Organism Investigation Lab. See Figures 3-6 for detailed handouts of these activities. The video served as a means of reviewing the scientific method. The pre- and post-tests compared students' inquiry skills and knowledge of the scientific method before and after instruction.

Data Analysis

The main data source for this study was the laboratory design question found on the inquiry unit pre- and post-test (see Figure 2, Question #13). The question included specific written components, which were analyzed separately. The target components of the question included hypothesis formation, variable control, procedure writing, data analysis description, and a labeled diagram or picture. We analyzed students' responses to each of these aspects and rated them on a four-point Likert scale. Before we analyzed and scored students' responses to the various lab components, we established inter-rater reliability between Mr. Turner and another science teacher.

Mr. Turner and a volunteer science teacher independently rated a 10% sample of the lab design questions. Both pre- and post-tests were included in this rating sample. We established inter-rater reliabilities for each aspect of the design question: scientific method knowledge (93%), hypothesis formation (79%), variable control (86%), procedure writing (70%), data collection (71%), and labeled picture (100%). For readers to gain a better understanding of this system, the Likert scale for hypothesis formation is provided below to show a representative sample of student responses for clarification. Student examples are shown in italics.

1 = Hypothesis was merely mentioned but not attempted.

Don't remember what exactly a hypothesis is.

2 = Hypothesis was attempted, but was unfocused, did not set the stage for the investigation or was otherwise not testable as stated.

I think that the female mouse will be able to run through the maze more quickly than the male mouse because she is smaller in size.

Figure 3. Tootsie Pop Experiment.

Question:

Observation: (Comment on flavor, shape, size , abnormalities.)

Hypothesis: (Place your best guess below in a complete sentence.)

Method: (Carefully define what a lick is and define your stopping point.)

Data:

Results: (use a complete sentence to report the actual number of licks it took.)

Analysis Questions: (Answer the following in complete sentences.)

1. Was your hypothesis correct or not?
2. Did you follow your method exactly?
3. Describe any problems you encountered during the activity.

Variables List:

1.
2.
3.
4.
5.
6.

Extension:

How could you answer this question in a more scientific manner? In other words, how could you achieve a reproducible answer (by eliminating variables) to the question of how many licks it takes to get to the Tootsie Roll center of a Tootsie pop? Write your answer below.

Figure 4. The Quicker-Picker-Upper Lab

This lab has been created for you to help gain proficiency with scientific methods and later become a full-fledged research scientist. Your overall goal in this lab is to determine which paper towel of the given brands is the best value. In other words, you are trying to determine which paper towel gives you the most performance for your money using absorption data (number of water drops) and price-per-sheet data. You must follow the method that you and your classmates generated in our earlier discussion. You will be assessed on the design of this lab, neatness of tables and graphs, and most of all, your conclusion. Your lab must include the following components, neatly labeled, and written in order on your own paper.

TITLE
Make your own creative title and write it in large letters in the upper margin of your paper. The title should somehow reflect the materials you are working with.

PURPOSE
A simple purpose statement a few sentences in length should be enough to describe why you are doing research on the given brands of paper towels.

OBSERVATIONS
Write a description of each paper towel, making sure to note texture, dimpling, graphics, etc.

HYPOTHESIS
Write a concise hypothesis in the form of an If ...then ... statement.

MATERIALS LIST
Your list should be complete in terms of the brand names of the towels and size of towel pieces. Include all other relevant materials necessary to accomplish this lab.

PROCEDURE
Rewrite the procedure the class has generated on your paper. It should be neatly written in a clear, step-by-step manner. Please see the example:
Step 1. Yada, Yada, Yada
Step 2. Yada, Yada, Yada
Be sure to state precisely how you will collect your numerical data.

DATA TABLE
Design your table to hold all of your data, plus that of the others. Label it carefully, as you will be using it later to make a graph and analyze the data.

Complete all steps above and get your lab checked by Mr. Turner before proceeding.

GRAPH
Using an assortment of colors and a sheet of graph paper, create a neat visual representation of your data. This must be a bar graph that is carefully labeled.

DATA ANALYSIS
Discuss the data you obtained in this lab using the data table and the graph you created. Point out the critical differences in the data while making reference (refer the reader back) to your data table or graph.

CONCLUSION
Discuss the following in paragraph form when your lab is complete. Do not add numbers or bullet points in front of each discussion point. Be clear and use proper grammar.

- Does your data support your original hypothesis? Discuss why or why not.
- What were the variables that you had to control in this lab?
- Which paper towel held the most water?
- Which paper towel held the least amount of water?
- Which paper towel is the best overall value, taking into account cost and absorbency data?
- Which paper towel would you buy now that you have this data?
- What would you do differently next time?

TOWEL DATA (Use to find overall value):

Brand	# Sheets per roll	Size of each sheet	Cost per roll	Cost per cm^3

3 = A feasible hypothesis was stated, but could have used more specific language to further focus and clarify the investigation.

Female mice with the medication of Drug X can and will run faster through a maze then a male mice.

4 = Hypothesis was focused, clearly stated, and readily testable.

If a female mouse is given Drug X then it will run through a maze more quickly than the male mice given Drug X.

Other Likert scale descriptors can be found in Figures 7-12.

Results

We analyzed pre- and post-test data and computed average raw scores for each of these targeted components for each class to determine the degree of change following the inquiry unit. Using a paired samples test, we also determined whether statistically significant differences existed between students' pre- and post-test responses. We summarized the results of each targeted component individually across the four classes studied in Tables 1 through 4.

In summary, the average scores for knowledge of scientific methods and all of the specifically targeted components either stayed the same or increased across the pre- and post- assessments. Overall, we found the largest percentage gains in Likert scale scores in the two general biology classes. However, except for the changes in overall knowledge of scientific methods, the majority of the target categories revealed little change occurred in Likert scores across the pre- and post-assessments. It is also important to note that a change in score from 2 to 3 represented an improvement from a "non-feasible" to "feasible" answer, which is significant change. The paired samples tests are perhaps more meaningful, as the calculations reveal trends in statistical significance across the pre- and post-test that were not otherwise apparent. We found more statistically significant differences in the general biology classes than in the honors biology classes.

Figure 5. Organism Investigation

This portion of the lab is to help you get to know our lab animals. Take some time to check them all out, and narrow your focus to two of your favorites. Sketch each of them and under each provide 10 detailed observations, generate five questions, and then formulate three hypotheses as If …then … statements.

Organism #1 Sketch	Organism #2 Sketch
Organism #1 Observations (10) 1. 2. 3. 4. 5. 6. 7. 8. 9. 10.	Organism #2 Observations (10) 1. 2. 3. 4. 5. 6. 7. 8. 9. 10.
Organism #1 Questions (5) 1. 2. 3. 4. 5.	Organism #2 Questions (5) 1. 2. 3. 4. 5.
Organism #1 Hypotheses (3) 1. 2. 3.	Organism #2 Hypotheses (3) 1. 2. 3.
HABITAT NOTES Organism #1	Organism #2

Discussion

Overview

In addressing the question of whether or not teacher-designed inquiry-based units are effective, we have explored the call for further research and curriculum development by classroom teachers (Keys & Bryan, 2001; Chinn & Hmelo-Silver, 2001; Huber & Moore, 2001; Neopolitan, 1999; and Bencze & Hodson, 1999). Enhancing students' ability in the area of laboratory investigation design is vitally important, as in recent years they have been asked to design several investigations within the Science WASL. Since these questions typically fall into the extended response category, they are worth a total of four points and, as such, can have a marked effect on a students' overall test score. The question used as an evaluative tool within the context of this project provides an example of what students are asked to tackle within the Science WASL. This is the last question on the pre-test, which can be viewed specifically within Figure 2. (Examples of Science WASL release items can be found at http://www.k12.wa.us/Ealrs/TopNav/WaslSupport.aspx. Just choose a grade level).

When these questions are presented on the Science WASL, students are given only a brief prompt that sets up a scientific investigation. In order to be successful, students must rely on their inquiry skills and reasoning ability to complete the investigation from the information given within the prompt. Such inquiry skill and reasoning ability may only be gained through extended laboratory experiences wherein students design and complete their own investigations (Roth & Roychoudhury, 1993).

Overall, the laboratory activities presented over the course of the unit served to increase students' understanding of science

Figure 6. Organism Investigation

This lab has been created to help you gain proficiency with scientific methods and later become a full-fledged research scientist. Your overall goal in this lab is to get to know our classroom pets a little better and then design your own experiment with one of your choice. You must carefully observe the organisms and then design your own investigation with minimal help from Mr. Turner. In other words, you are designing your very own authentic investigation with your favorite organism. You will be assessed on the design of this lab, neatness of tables and graphs and, most of all, your conclusion. Your lab must include the following components, neatly labeled, and written in order on your own paper.

TITLE
Make your own creative title and write it in large letters in the upper margin of your paper. The title should somehow reflect the organism you are working with.

PURPOSE
A simple purpose statement a few sentences in length should be enough to describe why you are doing research on this particular organism.

OBSERVATIONS
Write a description of your organism that details its appearance and observed habits.

HYPOTHESIS
Write a concise hypothesis in the form of an If ...then ... statement.

MATERIALS LIST
Your list should be complete in that it contains all of the relevant materials (complete with names and sizes) necessary to accomplish this lab.

PROCEDURE
Carefully write your own procedure in a clear, step-by-step manner. Please see the example below:
Step 1. Yada, Yada, Yada
Step 2. Yada, Yada, Yada
Be sure to state precisely how you will collect observational and numerical data.

DATA TABLE
Design your table to hold all of your numerical data. Label it carefully, as you will be using it later to make a graph, and analyze the data. Use an adjacent section for noting behavior.
Complete all steps above and get your lab checked by Mr. Turner before proceeding.

GRAPH
Using an assortment of colors and a sheet of graph paper, create a neat visual representation of your data. This may be a line or bar graph that is carefully labeled.

DATA ANALYSIS
Discuss the data you obtained in this lab using the data table and the graph you created. Point out key findings and any critical differences in the data while making reference (refer the reader back) to you data table or graph.

CONCLUSION
Discuss the following in paragraph form when your lab is complete. Do not add numbers or bullet points in front of each discussion point. Be clear and use proper grammar.

- Does your data support your original hypothesis? Discuss why or why not.
- What were the variables that you had to control in this lab?
- How did you ensure that only the independent variable was the one being tested?
- What was the most difficult part of the investigation for you?
- What would you do differently next time in designing or carrying out your lab?
- What would have made this lab easier for you?

process skills, the parts of the scientific method, and the individual written components of laboratory investigations assessed within the Science WASL. The unit materials presented were effective across all of the classes in increasing students' knowledge of the scientific method. However, the same was not true of the students' gains in experimental design capability.

Greater increases in the design component areas were evident in both general biology classes, but were not readily apparent within the honors biology classes. For example, students in the general biology classes made statistically significant gains in hypothesis formation, variable control, procedure writing, and data collection description. By contrast, Honors Biology A achieved statistically significant gains in the areas of scientific method and variable control description. The Honors Biology Group B achieved significant gains in scientific method, hypothesis formation, and procedure writing. Although statistical significance is an important focus, we cannot lose sight of important gains made by Honors Group A in hypothesis

**Figure 7. Scientific Methods Knowledge Likert Scale.
Student examples are shown in italics.**

1 = Student did not attempt question or was not able to list or describe any of the basic parts of the scientific method.
I don't remember my teacher talking about the Scientific Method

2 = Student was able to list or discuss 1-2 basic parts of the scientific method.
Steps scientists are supposed to use to solve stuff involving a hypothesis and some other stuff. I think it has like 7 or 8 steps

3 = Student was able to list or discuss 3-4 basic parts of the scientific method.
hypothesis, experiment, collect data, thesis, conclusion can't remember the other step don't remember what everything is

4 = Student was able to list or discuss 5 or more basic parts of the scientific method.
What I know about scientific methods is these following steps:
1. *You first make observations*
2. *You then state a problem or a formable question*
3. *Then form a hypothesis (testable prediction)*
4. *Test your experiment and gather data*
5. *Analyze and record the data*
6. *Then report your findings*
7. *Make a replicable or redesign your experiment*

Figure 8. Hypothesis Formation Likert scale.
Student examples are shown in italics.

1 = Hypothesis was merely mentioned but not attempted.
Don't remember what exactly a hypothesis is.

2 = Hypothesis was attempted, but was unfocused, did not set the stage for the investigation, or was otherwise not testable as stated.
I think that the female mouse will be able to run through the maze more quickly than the male mouse because she is smaller in size.

3 = A feasible hypothesis was stated, but could have used more specific language to further focus and clarify the investigation.
Female mice with the medication of Drug X can and will run faster through a maze then a male mice.

4 = Hypothesis was focused, clearly stated, and readily testable.
If a female mice is given Drug X then it will run through a maze more quickly than the male mice given Drug X.

Figure 9. Controlling for Variables Likert scale.
Student examples are shown in italics.

1 = Did not list or discuss any relevant variables, controlled or uncontrolled.
Control for@ least two variables: I am not sure what this means

2 = Listed or discussed at least one possible variable, but did not clearly state why it needed to be controlled.
To control at least two variables the scientist needs to give some mice drug x and some mice nothing so he can tell the time difference between male and female mice

3 = Listed or discussed at least one relevant variable and clearly stated why it needed to be controlled.
Control Variables: You need four exactly-the-same mazes, two for the female mouse and two for the male mouse because first you have to have exactly-the-same mazes so they are running the same course. Then two for each because if you want to test more than once you would either have to clean and disinfect the mazes or have more than one for each.

4 = Listed or discussed at least two relevant variables and clearly stated that they needed to be controlled.
Variables, and controlling them: Use identical mazes, or one maze but after each use you'd need to clean it out, this creating a constant environment for each trial
After giving drug x to the mice, before running the maze they should be given 20 minutes before running the maze in another constant, control environment such as a glass case where temperature can be controlled. The mice should be given the same quarters to allow the drug to set in because different environments may effect the effects of the drug

formation and Honors Group B in variable control. While not statistically significant, their answers improved from "non-feasible" to "feasible." In other words, they now had a better grasp of important skills and concepts as evidenced by their scores. The reasons behind the honors groups' apparent lag in performance as compared to the general classes' deserve further discussion.

Figure 10. Procedure Writing Likert scale.
Student examples are shown in italics.

1 = Did not attempt to write a procedure, or simply gave non-specific advice as to how it should be done without describing the specifics of a possible lab set-up.
The doctor should tell how hes going to test his hypothesis while taking data

2 = Procedure is incomplete, misleading, lacks necessary detail or, would otherwise not be readily followed by another scientist.
build maze
give mouse A chem. X
have mouse A run maze
collect data
give mouse B chem. X
have mouse run maze
collect data

3 = Procedure is mostly clear, complete, and could be followed, but lacks minor details that would help clarify the investigation.
Buy 2 female mice and 2male mice. Feed all four mice Drug X. Take 1 female mice& 1 male and time them going through the maze. Then use the other male& female and do the same thing. Compare the results of the 2 sets of mice

4 = Procedure is very clear, complete, and ample detail is provided such that the directions could readily be followed by another scientist.
Procedure: both a male and female mouse was given drug x. The male mice was then put into a maze where it will be timed of how long it will take to complete the maze. Then the female mice was put into the same maze and timed how long it will take to complete the maze. Each trial was recorded. The procedure was then repeated two more trials where both mice times were added the averaged to see whether the male mice was quicker than the female mice.

Differences in Performance

Prior Experience

Presumably, the students in the honors classes may have used reasoning skills better than their counterparts in the general classes. Since reasoning ability was not measured within the scope of this study, it is difficult to explain why this result occurred. Perhaps previous experiences with activities that build reasoning skills allowed many of the honors students to perform well on the pre-test. The majority of students from the two honors classes posted higher scores on the pre-test than did most of the general biology students. If individuals within the honors group also did well on the post-test, it is difficult to discern to what extent their ability level increased due to exposure to the unit materials. Many students within the honors classes would likely have scored well on the experimental design questions within the Science WASL without ever having been exposed to the unit materials. Given the greater differences in attendance and performance from pre- to post-test for the general biology students, however, the same statement might not hold true.

Figure 11. Data Collection Likert scale. Student examples are shown in italics.

1 = Student did not attempt this portion of the procedure or never specifically described how data was to be collected.
How data will be collected – he Could Collect DATA With A Data Table

2 = Student states that data will be collected within the procedure, but does not specifically describe how it is collected or what data is to be written down.
run each mouse through the maze at least 3 times, record your findings. A week later run them through the maze again 3 times, record your findings.

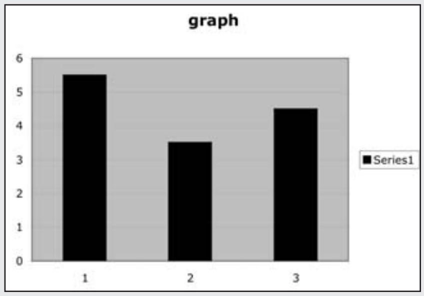
3 = Student states how data are collected and what data are to be written down within the procedure, but does not mention what measurement tools are utilized.
Put mouse in the maze and measure how much time it take to complete it
Record timings

4 = Student specifically states how data are collected and what data are to be written down while including the measurement tools utilized.
Put the male mouse through the maze and time him using a stopwatch. Record time. Repeat steps 1-4 for each female mouse.

Figure 12. Labeled Picture or Diagram Likert scale. Student examples are shown in italics.

1 = Did not attempt to draw a diagram or picture.
(Nothing can be shown for example here)

2 = Drew a picture or diagram, but was not considered helpful in giving a scientist a clearer picture of how data was to be organized or what was to happen in the experiment.
(The following is a re-creation of a student's work using Excel)

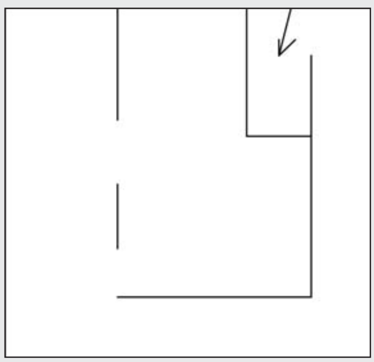


3 = Drew a data table which clearly depicted how data was to be organized for the experiment. (The following is a re-creation of a student's work using drawing tools.)

Data Collection:

	time		*time*
male mouse #1		*female mouse #1*	
male mouse #2		*f. mouse #2*	
male mouse #3		*f. mouse #3*	
total speed		*total speed*	
average speed		*average speed*	

4 = Drew a picture or diagram of the experimental setup and labeled it in such a way that a scientist would have a clear picture of how the experiment could be accomplished. (The following is a re-creation of a student's work using simple drawing tools.)



Attendance Issues

Generally speaking, attendance was better for the honors classes than for the general classes. A review of attendance records and completed inquiry assignments revealed that students who missed critical lab design and discussion days typically fared worse on the post-test as compared to others. This premise held especially true if the absent students failed to make up lesson or inquiry lab work before the post-test was taken. Perhaps adding a reward system for good attendance may improve student performance, especially on critical days.

Incomplete Lab Reports

Another issue relating to the lab work in this unit was the fact that many students, across the all of the participating classes, did not complete their assigned lab reports. Many had carried out interesting investigations, but failed to take sufficient data, graph the data, or complete the analysis questions. This was especially troubling, as we believe that it was difficult for students to grasp the full outcome of a scientific investigation if they never analyzed the data gathered. Without the analysis, students could not draw valid conclusions about their work. In essence, they never answered the question they strived so hard to formulate in the beginning.

We felt that students were given ample time to complete all parts of the labs in class and then were able to finalize the labs as homework. However, many students simply chose not to complete the graph and written lab portions as homework. Their choice not to do so may have adversely affected their performance on the investigation design question within the post-test. A follow-up review of the two major inquiry labs completed within the unit, the Quicker-Picker-Upper Lab and the Organism Lab, revealed an interesting statistic. Taking all classes into account, 22% of the two labs turned in did not contain written conclusions. Not having completed the full data analysis may have caused them unnecessary difficulty on the lab design question wherein a clear description of data collection and organization were required. We plan to devote more in-class time to completing these sections, as well as using peer review as an accountability strategy.

Strengths & Limitations of This Study

As with all classroom-based studies, the extent to which we can generalize our results is limited to these classes. Although

Table 1. Summary of Honors Biology A Data.

	Pre-test Average Score (out of 4)	Post-test Average Score (out of 4)	Average Change in Score (out of 4)	% Change in Score	Statistically Significant?
Scientific Methods	2.1	3.6	+1.5	+38%	Yes
Hypothesis	2.8	3.0	+.2	+5%	No
Variable Control	2.1	3.0	+.9	+23%	Yes
Procedure	2.5	2.6	+.1	+2%	No
Data Collection	3	3	0	0	No
Picture	1.6	1.6	0	0	No

Table 2. Summary of Honors Biology B Data.

	Pre-test Average Score (out of 4)	Post-test Average Score (out of 4)	Average Change in Score (out of 4)	% Change in Score	Statistically Significant?
Scientific Methods	2.5	3.9	+1.4	+35%	Yes
Hypothesis	2.1	2.8	+.7	+18%	Yes
Variable Control	2.5	3.0	+.5	+13%	No
Procedure	2.1	3.4	+.1.3	+33%	Yes
Data Collection	3	3.4	+.4	+10%	No
Picture	1.7	1.5	-.2	-5%	No

Table 3. Summary of General Biology C Data.

	Pre-test Average Score (out of 4)	Post-test Average Score (out of 4)	Average Change in Score (out of 4)	% Change in Score	Statistically Significant?
Scientific Methods	1.5	3.7	+1.2	+30%	Yes
Hypothesis	2	3	+1.0	+25%	Yes
Variable Control	2.1	2.9	+.80	+20%	Yes
Procedure	1.7	2.7	+.1.0	+25%	Yes
Data Collection	2.1	2.9	+.80	+20%	Yes
Picture	1.4	1.4	0	0%	No

Table 4. Summary of General Biology D Data.

	Pre-test Average Score (out of 4)	Post-test Average Score (out of 4)	Average Change in Score (out of 4)	% Change in Score	Statistically Significant?
Scientific Methods	1.4	3.7	+2.3	58%	Yes
Hypothesis	1.9	3.1	+1.8	45%	Yes
Variable Control	1.8	3.1	+1.3	33%	Yes
Procedure	1.7	2.9	+1.2	30%	Yes
Data Collection	2	3.2	+1.2	30%	Yes
Picture	1.3	1.9	+.6	15%	No

other schools and classrooms differ in many ways, we expect that science teachers will relate to some of the issues examined in this study and find the discussions useful. We, on the other hand, have discovered much about inquiry instruction and using pre- and post-assessments.

With respect to inquiry-based instruction, we concur with the notions that they are time-intensive, less predictable in nature, and require more teacher facilitation than traditional laboratory exercises. Despite Mr. Turner's best efforts to prepare for multiple scenarios, the number of students who did not complete assignments surprised us. Since students' in-class participation and enthusiasm seemed positive, finding partially-completed assignments was disappointing to us.

Pre- and post-test design is always fraught with complications. Do you use the same questions on the pre- and post-test? Do we use different questions but employ parallel design? How much time should lapse between each assessment? In our case, we used the same questions on the pre- and post-tests and administered them six weeks apart. Based on our examinations of answers and errors, we are confident that enough time had passed. However, we cannot rule out the possibility that some improvement may be due to decreased sensitivity to the test. In other words, students became better test-takers after six weeks of instruction.

Overall, Mr. Turner and I were impressed by students' improvement in the five-targeted areas. We plan to continue exploring how inquiry-based activities affect student performance on investigational problems. In addition to refining the materials in this study, we hope to implement other activities across the academic year. As we build more cases across the biology curriculum, we hope to discover more themes related to the relationship between inquiry-based activities and student performance.

Implications for Future Research

Clearly more research is necessary to better understand the methods that best prepare science students for statewide science assessments. Barring any change from the current state mandates, the level of performance on such assessments will affect administrators,' teachers,' and students' lives directly. More teachers should pick up the action research torch and share their findings such that all can benefit. Teachers should be encouraged by administrators to take an in-depth look at the science lessons they present each day to see what improvements might be made.

Furthermore, larger and more diverse groups of high school students should be studied to determine how effective inquiry-based units are in preparing them for standardized tests. More research is needed to determine to what extent gender and ethnicity affect student learning through inquiry. Case studies of successful teachers who implement inquiry are needed so that others can get a better feel for the exciting atmosphere created by kids engaging in inquiry science. We advocate that such teachers be videotaped throughout the course of their units so that all science educators, novice to experienced, can see how inquiry can be successfully implemented. Perhaps seeing the engagement of students in inquiry activities might persuade traditionally-minded science teachers to attempt inquiry methods in their classrooms.

Conclusion

Standardized tests, such as the Science WASL, continue to be one of the driving forces behind curriculum change in Washington State. Student achievement on this particular measure has never been more important. Ultimately, the reputation of the school and the jobs of administrators and teachers may hang in the balance of the degree of student success on this state mandated standardized test. Given the high-stakes nature of such tests, and the research-based effectiveness of inquiry, it seems imperative that educators begin and continue to implement more inquiry activities within their classrooms. In so doing, students stand to gain a better grasp of the nature of science while attaining the valuable content knowledge and reasoning skills necessary for success on standardized tests.

Although inquiry science has been around for decades, a pressing question remains. How can teachers effectively foster inquiry skills in all students such that they gain the proficiency necessary to tackle standardized tests? Though not specific to the Science WASL, much of the available literature on this topic illustrates that teaching through inquiry can be a sound way to increase attitude and achievement. The data gathered within this study were concurrent with those found in the research literature. Clearly, more research is necessary to investigate the effectiveness of the inquiry materials successful science teachers have most often utilized. It is of utmost importance to determine which of those materials can get students to a high degree of proficiency within a limited amount of time. Action research projects conducted by educators themselves are greatly needed if we are to give students a positive educational experience and, at the same time, ensure their success on the standardized science tests.

References

Ad-Marbach, G. & Sokolove, P.G. (2000). Can undergraduate biology students learn to ask higher level questions? *Journal of Research in Science Teaching, 37*(8), 854-870.

Bencze, L. & Hodson, D. (1999). Changing practice by changing practice: Toward more authentic science and science curriculum development. *Journal of Research in Science Teaching, 36*(5), 521-539.

Chinn, C.A. & Hmelo-Silver, C.E. (2001). Authentic inquiry: Introduction to the special section. *Science Education, 86*, 171-174.

Chinn, C.A. & Malhotra, B.A. (2001). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education, 86*, 175-218.

Freedman, M. (1996). Relationship among laboratory instruction, attitude toward science, and achievement in science knowledge. *Journal of Research in Science Teaching, 34*(4), 343-357.

Gable, C. (2001). Effectiveness of a scaffolded approach for teaching students to design scientific inquiries. Doctoral dissertation, Colorado University, *Dissertation Abstracts International, 63*(3-A), 966.

Gibson, H.L. & Chase, C. (2002). Longitudinal impact of an inquiry-based science program on middle school students' attitudes toward science. *Science Education, 86*, 693-705.

Huber, R.A. & Moore, C.J. (2001). A model for extending hands-on science to be inquiry-based. *School Science and Mathematics, 101*(1), 32-43.

Keys, C.W. & Bryan, L.A. (2001). Co-constructing inquiry based science with teachers: Essential research for lasting reform. *Journal of Research in Science Teaching, 38*(6), 631-645.

National Research Council. (1996). *The National Science Education Standards*. Washington, DC: National Academy Press.

Neopolitan, J. (1999). Teachers' beliefs about redesigning instruction to meet new standards through action research. Paper presented at the Annual Meeting of the Association of Teacher Educators, Chicago, Illinois.

No Child Left Behind. (2003). Overview Section. Retrieved April 13, 2007 from http://www.nclb.gov/next/overview/index.html.

Olsen, L. (2001). Holding schools accountable for equity. *Leadership, 30*(4), 28-31.

Office of Superintendent of Public Instruction. (2005). Washington high school graduation requirements. [Online]. Retrieved April 13, 2007 from http://www. k12.wa.us/Communication/NewsInformation/.

Office of Superintendent of Public Instruction. (2003). Grade 10 sample mini science WASL. [Online]. Retrieved April 13, 2007 from http://www.k12.wa.us/assessment/WASL/ScienceAssessment.aspx.

Partnership for Learning. (2003). Standards question and answer. [Online]. Retrieved April 13, 2007 from http://www.partnership4learning.org/edreform_faqs.htm.

Roth, W.M. & Roychoudhury, A. (1993). The development of science process skills in authentic contexts. *Journal of Research in Science Teaching, 30*(2), 127-152.

Schneider, R.M., Krajcik, J., Marx, R.W. & Soloway, E. (2002). Performance of students in project-based science classrooms on a national measure of science achievement. *Journal of Research in Science Teaching, 39*(5), 410-422.

Udovic, D., Morris, D., Dickman, A., Postlethwait, J. & Wetherwax, P. (2002). Workshop biology: Demonstrating the effectiveness of active learning in an introductory biology course. *Bioscience, 52*, 272-281.

Yerrick, R., Parke, H. & Nugent, J. (1997). Struggling to promote deeply rooted change: The filtering effect of teachers' beliefs on understanding transformational views of teaching science. *Science Education, 81*,137-159.