TEACHING DATA LITERACY IN HIGH SCHOOL BIOLOGY

by

Patricia Lynn Jones

A professional paper submitted in
partial fulfillment of the requirements for the degree

of

Master of Science

in

Science Education

MONTANA STATE UNIVERSITY
Bozeman, Montana

July 2016

DEDICATION AND ACKNOWLEDGEMENT

This capstone project is dedicated with love and gratitude to my family. To my husband Dan, you have been a constant source of love and unwavering support throughout this process. You are my rock and the love of my life. To my children Danny, Matt, Meghan, and Jessica, you are the joys of my life. You have helped me immeasurably throughout this project. You give my life purpose, happiness and endless delight. I also want to dedicate this paper to my new loves: Josie, my beautiful daughter-in-law, my three adorable, smart and funny grandchildren Luke, Calvin and Gwyneth, my son-in-law Gus and my beautiful future granddaughter, Elizabeth. I am so lucky and grateful to have you in my life. I love you all.

I would also like to acknowledge Dr. John Graves, who guided me through this project with patience and wisdom. Your help is greatly appreciated. Thank you so much.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Time restrictions and curricular requirements in high school biology often do not allow for instruction in important data literacy skills. This capstone project was designed to develop high school biology lessons that incorporate data literacy skills into the curriculum. It was hypothesized that student engagement, confidence and skill levels would improve if authentic data was used in these lessons. Lessons were designed that incorporated each data literacy skill into a different unit of instruction. Success of the project was assessed with two skills tests, student surveys and interviews. It was shown that students' data literacy skills and confidence levels increased substantially with these lessons. In addition, student engagement and interest levels improved with the use of real data.

INTRODUCTION AND BACKGROUND

Throughout my career as a high school science teacher, I have often encountered students who find science demanding and difficult. Worse than that, many students feel that their science classes are not only boring, but irrelevant to them. There is often a lack of connection between what they are learning in class and the development of the analytical and critical thinking skills they will need for their future. One reason for this may be because many science lessons are designed so that students only learn about what others have already done. They aren't actually involved in developing new knowledge so these lessons are often uninteresting and student engagement is minimal. In addition, the required curriculum in high school biology includes so many topics that the schedule makes it problematic to include student-designed research or to incorporate instruction in data literacy skills. As a consequence, most students lack the basic analytical skills they need to interpret difficult data sets. They are often unprepared for the more demanding challenges in scientific study they will be facing in college or career.

As a new science teacher many years ago, I relied heavily on "cook-book" labs with pre-determined outcomes. The main goal for both my students and me was that they would get the right answer. As long as they followed my careful directions, they were able to learn what I thought they should learn about the topic. Every student did the same experiment, obtained similar data, and analyzed it with the same type of graph. What I failed to realize is that this is the opposite of how science actually works. Science is rarely linear in approach and often produces contradictory data, which is why it should be inherently interesting for students. Most lessons designed for high school science,

however, don't allow time for students to engage in authentic research nor do these lessons teach any data literacy skills beyond basic graphing.

Typical strategies to teach basic data literacy skills occur in a single unit of instruction, most often called the scientific method unit. Lessons in this unit are usually taught at the beginning of the school year and include acquiring data sets from textbooks or conducting labs with pre-determined results. This allows for whole-group instruction on basic graphing techniques. All students in the class have the same data and therefore, the particular skill set can be quickly incorporated into a lesson on graphing. While this method is certainly less demanding and time-consuming for both student and teacher, it stifles creativity, critical thinking and student engagement. Students may be less motivated to develop these skills because the results are meaningless to them other than for a grade. Another problem with this method is that it is frequently taught as a stand-alone unit of instruction, unrelated to any other biological concept. Students are assessed and then we move on to lessons in biology. It is as though biology and the scientific method are two separate and unrelated topics. In particular, the data literacy skills of analyzing data using basic statistics or box-and-whisker plots and arguing claims using evidence from graphs are rarely incorporated into the high school biology curriculum.

Instead of teaching these skills separately, I realized that data literacy should be incorporated throughout the year into every unit of biology instruction. To increase student engagement and focus on improving skills, instructional units can include real data either from primary sources or from student-designed experiments. While students

still need to learn the important biological concepts, instruction in data literacy can take place along with content lessons.

This project was conducted in two of my honors biology classes at Casteel High School, located in Queen Creek, Arizona. The population of Queen Creek is 29,670 and the median price for a home is $241,000. The median income in Queen Creek is $81,280. This is substantially above the median for the state of Arizona which is $47,826 (City-Data.com, 2013). Casteel High School is one of five high schools in our district and has 1,115 students (B. Naik, personal communication, December 3, 2015). Students in my honors biology classes are typically high-achieving, motivated students. The average class size is 27 students, mostly freshmen.

The search for a better method for teaching data literacy skills in high school biology led to the focus of this investigation, *How can the use of real data help students learn data literacy skills?* One sub-question of this project was, *How does incorporating data literacy into biology content lessons improve students' data literacy skills?* A second sub-question was, *How does the teaching of data literacy skills improve student engagement and confidence levels in organizing and analyzing data?*

CONCEPTUAL FRAMEWORK

Teaching science today is becoming increasingly complex as educators develop ways to help their students acquire data literacy skills. It has been decades since reading, writing, and basic math were enough to provide a living for citizens. Today's students need skills in critical thinking, creativity, teamwork, and working with data to be ready for future education and workplace demands, often referred to as 21st century skills.

Proficiency in analyzing data, in particular, is becoming more important today than ever before. In fact, the collection of data for scientific research, medicine or business is useless without the ability to analyze what the data means (Harris, 2012).

Technology has advanced so quickly that students are now able to look up facts at the click of a button. Therefore, just teaching students a body of facts is not enough. They must be able to use data to help make decisions, justify claims and solve problems. Data literacy skills include a basic knowledge of statistics and how to collect, analyze and communicate data. These are fundamental skills required in many careers in the 21st century (Harris, 2012). Yet, the need for instruction in data literacy has not been effectively implemented in today's science classrooms (Bowen & Bartley, 2014; Konold & Higgins, 2003; Vahey, Yarnall, Patton, Zalles & Swan, 2006). Many of the data literacy skills needed for success in today's careers are taught only at the collegiate level and have not been adapted to middle or high school-level students (Konold & Higgins, 2003).

Analyzing and interpreting data is one of the eight science and engineering practices described in the *A Framework for K-12 Science Education: Practices*, which outlines the *Next Generation Science Standards* (NGSS) (National Research Council, 2012). The NGSS standards state that by graduation from high school, students should be proficient in analyzing data using tools, technologies and/or models, applying concepts of statistics and probability to science and engineering problems, considering limitations of data sets when analyzing data, comparing and contrasting types of data

sets, and evaluating the impact of new data on a working explanation and /or model of a proposed process or system (NGSS Lead States, Appendix F, 2013).

The Common Core State Standards for literacy in science and technical subjects require students to effectively evaluate hypotheses, perform data analysis, and write conclusions in a science or technical text. They must also be able to verify the data when possible and corroborate or challenge conclusions with other sources of information (Common Core State Standards, 2010).

One of the most important data literacy skills is being able to read, understand and gather evidence from texts with varying levels of complexity (Common Core State Standards, 2010). A common misconception is that practicing scientists and engineers spend most of their time doing field research. Tenopir and King (2004) found that working scientists actually spend half their time reading about research that has been done about their topic or writing about results of studies they are doing. High school students must be able to read and interpret text from many sources, including primary scientific literature, which can be complex and full of difficult vocabulary. It is vital that teachers help their students gain reading skills while they are developing other data literacy skills. Teachers must be cognizant of the reading and comprehension abilities of their students. In addition, the depth of understanding high school students receive from their reading as well as how they think about a topic in biology depends on the genre of the text used in the instruction. Norris and his group found that when given an adapted primary literature article, high school students demonstrated better inquiry and critical thinking skills in evaluating the text compared to those who were given a secondary

literature version. Students who read the secondary literature version, such as what might be found in a textbook or newspaper article, however, understood the article better. Using texts with varying levels of reading complexity will help high school students develop these skills (Norris, Stelnicki & de Vries, 2012).

Students need to be able to collect their own data or use data from a secondary source and then build graphical representations that help explain the results. The NGSS states that students should be regularly engaged in investigations that produce data. Students can then participate in activities using that data to build skills in graphing, data analysis, and basic statistics. They need to know the purpose of different kinds of graphs and then make informed decisions about which is the best to way to show the results of the investigation (NGSS Lead States, Appendix F, 2013). High school students should learn to use a spreadsheet program, use basic statistical tools such as Chi-square and make graphs with error bars (Bowen & Bartley, 2014; Hougen-Eitzman, 2007; Manduka & Mogk, 2002; Webber, et. al., 2014). Hougen-Eitzman (2007) is clear about what expectations should be in a science classroom, however. The data analysis skills must always be consistent with the course content. Teachers must decide which data skills go with the particular instructional unit they are planning. Data analysis, graphing and communication of data can be included in the learning goals, but not at the expense of teaching important content. "While an introductory class presents an excellent opportunity for teaching students how to deal with data, it is important to have realistic expectations. Keep in mind this is a biology class, not a statistics class" (Hougen-Eitzman, 2007, p. 385).

Graphing skills are more than just plotting data points correctly. Webber, Nelson, Weatherbee, Zoellick, and Schauffler (2014) found that when students began to organize, graph, and interpret their data, many were unsure about what kind of graph to make. According to the study done by Webber's group, most students made bar graphs, regardless of their research question. Students also treated their work as though the graph itself was the only product they needed to produce. They didn't realize immediately that a graph's purpose is to allow them to see patterns and make arguments. Webber's group developed the Graph Choice Chart (GCC), a tool which helps students choose the appropriate kind of graph to use. The chart starts with a precisely worded research question. Making sure the research question fits the data that was collected will help students choose which type of graph best represents that data. Once they state the research question, they use the GCC to choose the type of graph they need to make. Questions on the GCC include:

- Does your question ask about the variability within a group of data points? (one group, one variable);
- Does your question compare two or more groups to decide if the groups are the same or different? (two groups, one variable);
- Does your question ask if two numeric factors are correlated? (one group, two variables); or

- Does your question ask how a total is proportioned into subgroups? (Or, what proportion a subgroup is of a total?) (one group with subgroups, one variable).

Once students have identified the type of research question, they then follow the GCC to determine the type of graph they need to make. The GCC is not an absolute set of steps. Rather, it sets up a framework for students to think about the data analysis by asking questions and applying reasoning skills to help them choose the correct type of graph to use. "Students feel empowered when they realize they have a choice about what kind of graph to use" (Webber et. al, 2014, p. 41).

Another skill necessary for data literacy is how to deal with data that does not clearly show an expected result. This type of "messy" data is often found in real investigations and students often assume their data is incorrect if it doesn't correspond to what they expected. An important issue in scientific inquiry is dealing with ambiguity in data. Students should be involved in many experiments and activities which can frequently produce inconclusive or ambiguous data. These types of results are actually common with real scientists and students need skills in learning how to interpret confusing outcomes (Bowen & Bartley, 2014; Emery, Harlow, Whitmer, & Gaines, 2014; Gould, Sunbury, & Dussault, 2014; Metz, 2005).

Because science is tentative in nature, in order to move forward, scientists must be able to argue, defend claims, and make inferences. Argumentation is an important part of the scientific process. Besides being able to analyze and interpret data, arguing from evidence is another of the eight important science and engineering practices

identified in the NGSS (NGSS Lead States, 2013). The Common Core State Standards also require students to be able to build and critique evidence-based arguments (Common Core State Standards, 2010). Supporting a claim using evidence is critical to the scientific inquiry process. Being able to critique someone else's claim using evidence is an example of the type of critical thinking skills students need to meet the challenges of the 21st century workforce (Harris, 2013; Llewellyn, 2013; McNeill & Pimentel, 2010). Students need to be able to investigate phenomena, build a body of evidence from data, graphically display it in a way that makes sense, and use that evidence to back up a claim. Alternatively, they need to be able to see when a claim cannot be backed up with evidence. "We should teach that it's normal for scientists to differ, offer alternatives, and attempt to resolve conflicts with data and additional research" (Clary & Wandersee, 2013, p. 40). The importance of the teaching of argumentation skills cannot be overstated. According to Lewellen, the words *argument* or *argumentation* appear 97 times in the 325-page *A Framework for K-12 Science Education:Practices* (Lewellen, 2013).

To build data literacy, teachers should plan skill building lessons into their current curricular units. A constant concern for teachers is making sure they are covering required content areas for district and state assessments. Data literacy skills, however, could be incorporated as part of a number of instructional units (Webber, et. al., 2014). In classroom lessons, data literacy skills can be taught by scaffolding instructional activities designed to introduce concepts, practice skills in various ways and assess student progress as each skill is mastered. In addition, there isn't a prescribed list of tasks that

must be accomplished in a particular order. There are many ways students can use to communicate information, evidence and ideas. These include making tables or diagrams, building graphs, models or interactive displays. They can also practice written and oral communication skills by participating in debates, presentations, or extensive discussions (NGSS Lead States, 2013, Appendix F).

A lesson in ecology, for example, may start with students looking at raw data from a study of a declining population of a certain species. As the lesson progresses, students read about the study from either primary scientific articles or summaries from a secondary source, and confront tables of raw data. "Because raw data as such have little meaning, a major practice of scientists is to organize and interpret data through tabulating, graphing, or statistical analysis. Such analysis can bring out the meaning of data—and their relevance—so that they may be used as evidence" (NGSS Lead States, Appendix F, 2013, p. 9). In fact, the numbers in a statistical analysis are both less interesting and less important to students without the context from which they came. Vahey, Yarnall, Patton, Zalles and Swan (2006) studied the intersection between the mathematics of statistical analysis and the context given to these numbers in social studies or science. They realized that students can only experience true data literacy if their studies are deeply anchored in the context of the problem. Frequently, context can be provided by students analyzing data from a real-world perspective, rather than using fictitious data. Students learn that data is used by scientists as well as policy makers to make decisions and therefore, it must be precise and the analysis must accurately portray what the data shows.

Instruction in data literacy skills is most useful when incorporated within existing units of instruction (Bowen & Bartley, 2014; Hougen-Eitzman, 2007). Collecting data or gathering data from reading technical texts allow students to begin an investigation. This is followed by choosing the best way to communicate that data and making decisions about which type of graph, table, model or interactive display will best show their results (Webber, et. al., 2014). They must be able to confront ambiguity or conflict within the data they compile and be able to justify claims using evidence (Emery, et. al., 2014). The need for instruction in each of these skills is critical for our students to become scientifically literate.

METHODOLOGY

The focus of this study was to determine whether incorporating data literacy instruction into an existing biology curriculum enabled students to be more engaged in their learning and improve their data literacy skills. The student population for this study consisted of 2 sections of honors biology classes with a total of 54 ninth-grade students. Data literacy instruction was embedded into the evolution and ecology curricular units. For each of the data literacy skills that were the focus of this study, unit lessons consisted of instruction in the biology content along with one of the skills. The research methodology for this project received an exemption by Montana State University's Institutional Review Board and compliance for working with human subjects was maintained (Appendix A).

The treatment period occurred over the course of two months. The four instructional units in the treatment period were each designed to teach a particular data

literacy skill set: choosing and making the right kind of graph, using and analyzing box-and-whisker plots, analyzing data using a *t*-test and understanding the purpose of a *t*-test, and arguing claims using evidence from graphs. Each skill set was taught as part of a lesson that was embedded into the biology content. This enabled students to see how important data literacy is to understanding scientific principles.

Prior to the treatment period of instruction, students were given the Data Literacy Skills Confidence Survey (Appendix B). To gauge student attitudes about data literacy, the first part of the survey asked students their opinion of the importance of data literacy, how often they used data to make decisions and their overall confidence level in their data literacy skills. The survey asked them to rate their confidence and skill levels in each of the specific data literacy skills that were the focus of this project. The numbers on the survey correlated to a scale where zero indicated no understanding of the skill and four indicated they felt that they could be a peer tutor and explain it to others. The survey was analyzed by comparing student responses before and after the instructional unit. Results of paired answers to the attitude survey questions were analyzed using a Wilcoxon Signed Rank Test to determine whether there were any significant differences between the pre and post results.

The second part of the Data Literacy Skills Confidence Survey asked students to rate their confidence level for a number of individual data literacy skills on a scale from 0-4, which corresponds to a series of statements with varying degrees of mastery. The differences in these scores were analyzed by comparing the number of students pre- and post-treatment who chose the responses, *I could perform this skill by myself* or *I could be*

a peer tutor and explain it to others. This level of response indicated that the student felt that they met expectations for independent mastery of the skill. Students indicating they did not think they had mastered the skill chose one of the remaining responses, such as *I don't understand this at all*, *I get some of it but not enough to do well on a test* or *I could perform this skill with some help*.

Prior to the start of the project, five students were randomly chosen to be interviewed regarding their attitudes and opinions about data literacy with the Pre and Post Student Interview Questionnaire (Appendix C). Students were asked what their attitudes were regarding the importance of data literacy in their lives. In addition to the general data literacy questions, they were asked about the specific skills they learned from each lesson and how the lesson could be improved. Students were asked to indicate on a scale of 1-10 whether the use of real data helped them learn the skill, how engaged they were in the lesson, and how confident they were in their ability to perform this skill. Results were analyzed for trends in the data.

All students took the Practical Skills Test before and after the instructional unit (Appendix D). Students looked at one of three data sets, made an appropriate graph to organize the data and answered questions about the results. Students also took the Data Literacy Skills Test, a multiple choice test with 12 questions (Appendix E). The multiple choice questions were designed to assess the specific data literacy skills taught in the instructional unit. There were two to three questions for each of the data literacy skill sets. The results of the pre and post scores on the Data Literacy Skills Test were analyzed by calculating the means and the normalized gain. According to Hake (1998), a

normalized gain with a value of <0.3 is considered to be *low*, values of 0.3-0.7 are medium, and values >0.7 are considered *high*.

For the skill of choosing the right kind of graph, the lesson content was incorporated around a student designed experiment in ecology and the use of the Graph Choice Chart (Appendix F). As the beginning of the ecology unit, students were given a set of ten questions that different ecologists might ask and the Graph Choice Chart (GCC) was introduced and explained (Webber et al., 2014). Students worked in groups to choose a question from the list they were most interested in, design an experiment that would answer that question and, using the GCC, chose what kind of graph they would make. Following that exercise, students were asked to define what biotic and abiotic factors would be present if they considered our school site and the surrounding land to be its own ecosystem. Students took a field trip to the desert area surrounding the high school and recorded all of the organisms they saw. After making a food web of the Casteel High School ecosystem, student groups randomly chose a type of graph. Slips of paper contained the name of one of six types of graphs described on the GCC. Whichever graph type they pulled from a basket, they had to design an experiment that would give them data that would be best shown with that particular type of graph. The hypothetical experiment had to be ecology-based and designed to incorporate one or more of the organisms they saw living in the desert surrounding the high school. To conclude this instructional unit, students took the Graph Choice Chart Quiz (Appendix G). The quiz was analyzed by looking at the average and range of scores.

A lesson on genetic variations during the evolution unit was used to teach students how to make and interpret a box-and-whisker plot. The use of box-and-whisker plots helps students understand that often the range of data collected may provide more information than just determining the mean. An important concept in the teaching of evolution is the fact that variation occurs within all species and that these variations are often heritable. It is the variation within species that allows for natural selection. In the Variation in Natural Populations Lab, students looked at one of three different examples of variation within species (Appendix H). This task was accomplished by closely observing either a sample of pinto beans, making careful measurements of size and number of spots, the number of trichomes, or hairs, on the leaves of two types of desert plants or measuring the length and number of ridges for a species of mollusk shells. For whichever of the three examples they observed, students collected data, determined the mean, median, mode, and range and made histograms for each set of data. After asking students whether this was the best way to study the variation within species, we discussed how to incorporate the full range of data into their graphs. Students were shown a box-and-whisker plot and taught how to make one. Students then made box-and-whisker plots incorporating the data they had collected on their observations. The Box-and-Whisker Plot Quiz followed this lesson (Appendix I). This quiz was analyzed by determining the mean and range of scores.

To teach the skills of calculating and understanding the purpose of using a *t*-test, students looked at data from a study of protist species diversity in fresh water ecosystems. The Protist Diversity Lab was introduced during a lesson on populations

during the ecology unit (Appendix J). For this activity, samples of water were collected from local irrigation ditches and man-made lakes. Students chose two samples to observe for this activity. For each data point, they took a dropper of the sample of water, placed a drop on the microscope slide and counted the number of different species of organisms they saw in the water. They collected 10 data points and calculated the mean number of species for each of the samples. The following day, students learned how to calculate a biodiversity index, which they then applied to their data. Finally, the value of using a statistical tool such as a *t*-test was introduced and discussed. Students learned how to use the formula for a *t*-test to analyze the difference between the means of their two samples. The Statistical Analysis Quiz was the assessment for this unit and included questions on the value of a *t*-test and the parameters required to complete a *t*-test (Appendix K). This quiz was analyzed by determining the average and range of scores.

When students look at a graph or data set, they need to be able to make a claim about that data and justify it with evidence. This is arguably one of the most important data literacy skills. This skill was practiced using a lesson designed to study how certain environmental factors are influencing penguin populations on the Antarctic Peninsula. The Antarctic Climate Change Lesson contained five data sets describing changes in population sizes of two species of penguins in the Antarctic as well as changes in air temperature, sea ice distribution, winter snow, and krill density (Appendix L). This lesson was adapted from a similar lesson described by Constable, Sandro and Lee (2007). The lesson started with student groups of five, with each student taking on the role of a different scientist. Each student scientist was given a set of data to analyze and graph.

Students presented their findings to their group and looked at how their data set helped explain the effects of global climate change on penguin populations. Student groups were required to make a flowchart that explained their claim about what was causing the changes in populations and use evidence from their graphs to support their claims. One example of a group's flowchart showed the events leading to the population decline of Adelie penguins (Appendix M).

Data from the scores on the Practical Skills Test, the Data Literacy Skills Test, the Data Literacy Skills Confidence Survey, individual unit quizzes and student interviews were used to determine if the focus questions for this project were answered (Table 1).

Table 1
Data Triangulation Matrix

Focus Question	Data Source 1	Data Source 2	Data Source 3
Primary Question: 1. How can the use of real data help students learn data literacy skills?	Pre *and* Post Scores on Practical Skills Test and Data Literacy Skills Test	Pre *and* Post Data Literacy Skills Confidence Survey	Pre *and* Post Student Interviews on attitudes, general confidence levels and following individual unit lessons
Sub-Questions: 2. How does incorporating data literacy into biology content lessons improve students' data literacy skills?	Pre *and* Post Scores on Practical Skills Test and Data Literacy Skills Test	Comparing mean scores on Graph Choice Chart Quiz, Box- and Whisker Plot Quiz and Statistical Analysis Quiz	Pre *and* Post Data Literacy Skills Confidence Survey
3. How does the teaching of data literacy skills improve student engagement and confidence levels in organizing and analyzing data?	Pre *and* Post Data Literacy Skills Confidence Survey	Pre *and* Post Student interviews on attitudes and confidence levels	Pre *and* Post Student interviews following individual unit lessons

DATA AND ANALYSIS

The average score on the Data Literacy Skills Test before treatment was 42.02%, while the average score after treatment was 57.46% (*N*=54). Although the results show an average increase of 15.44%, the results of this test showed only a *low* normalized gain from pre-treatment to post treatment of 0.24 (*N*=54). Because there was a *low* normalized gain, there was only a small increase in the skills that were measured by this test even though there was an increase in the overall percentages. The distribution of scores shows that while the range of scores remained the same, the scores on the post-test were higher in all quartiles (Figure 1).

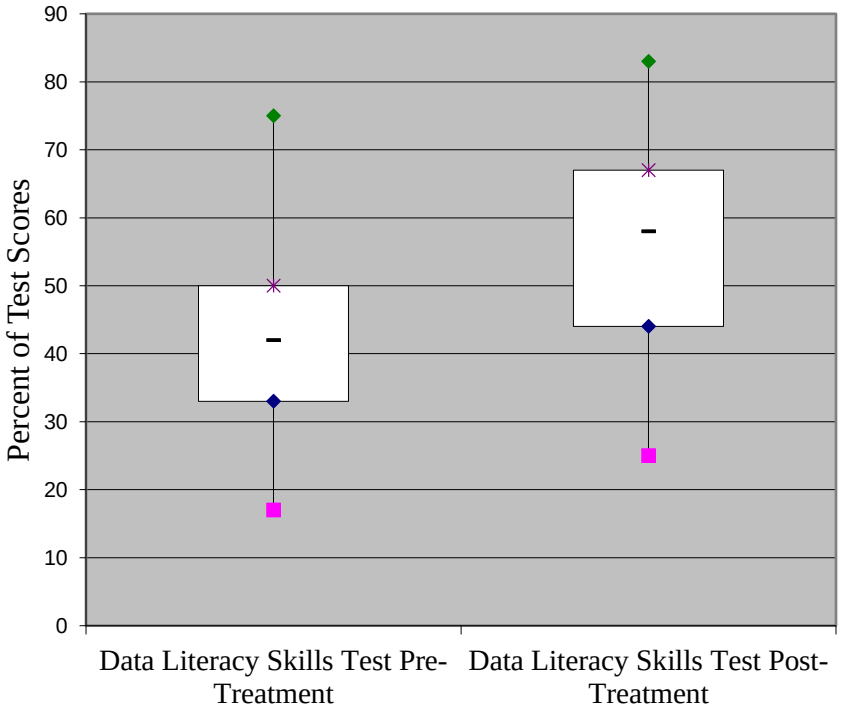


Figure 1. Box-and-whisker plot of scores from the Data Literacy Skills Test, (*N*=54).

The Practical Skills Test scores before the treatment period had a mean of 37.73%. The post test scores on the Practical Skills Test increased to a mean of 82.09%, an increase of 44.36%. In addition, the normalized gain for this test was .73 which, according to Hake (1998), is considered to be *high*. Scores on this test showed an increase in all quartiles (Figure 2).

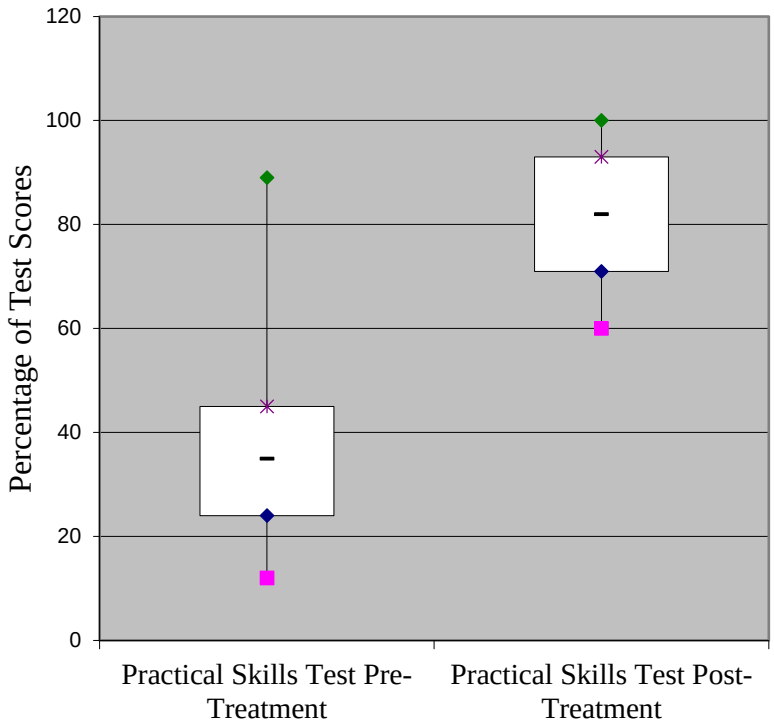


Figure 2. Box-and-whisker plot of scores from the Practical Skills Test, (*N*=54).

The Data Literacy Skills Confidence Survey asked students to give their opinions about the importance of data literacy and to rank their overall confidence level in their data literacy skills. Initial evaluation of the result of this survey showed differences in student responses pre and post treatment for all questions. However, after analyzing the results with a Wilcoxon Signed Rank Test, it was determined that these results were not statistically significant. The questions regarding student attitudes about the importance of

data literacy skills had p-values ranging from 0.59 to 1.0. The question regarding students' overall confidence levels showed a p-value of 1.0. Since all of these p-values were greater than 0.05, the null hypothesis stating that there would be no difference between pre- and post-treatment scores was accepted (Table 2).

Table 2
Wilcoxon Signed Rank Test Analysis of Student Attitudes and Overall Confidence Levels. Calculated for pre and post-treatment on the Data Literacy Skills Confidence Survey. A significance value of $\alpha=0.05$ *was used,* (N=54)

Survey Question	*N* for test	Wilcoxon Test Statistic	Critical Value (p-value)	Significant Increase?
How important is it for everyone to have data literacy skills?	4	5.0	1.00	No
How important is it for scientists to have data literacy skills?	3	0.0	1.00	No
How often do you use data to help you make decisions?	5	5.0	0.590	No
When you read an article or hear something on TV or the internet, how often do you question the statistics you have heard?	5	7.0	1.00	No
Rate your overall confidence level in your data literacy skills.	5	7.0	1.00	No

The number of students choosing the response *absolutely essential* for the question, "How important is it for everyone to have data literacy skills?" showed only a slight increase from one student pre-treatment to five students post treatment, which is consistent with the finding of no significance in the Wilcoxon Signed Rank Test (Figure 3).

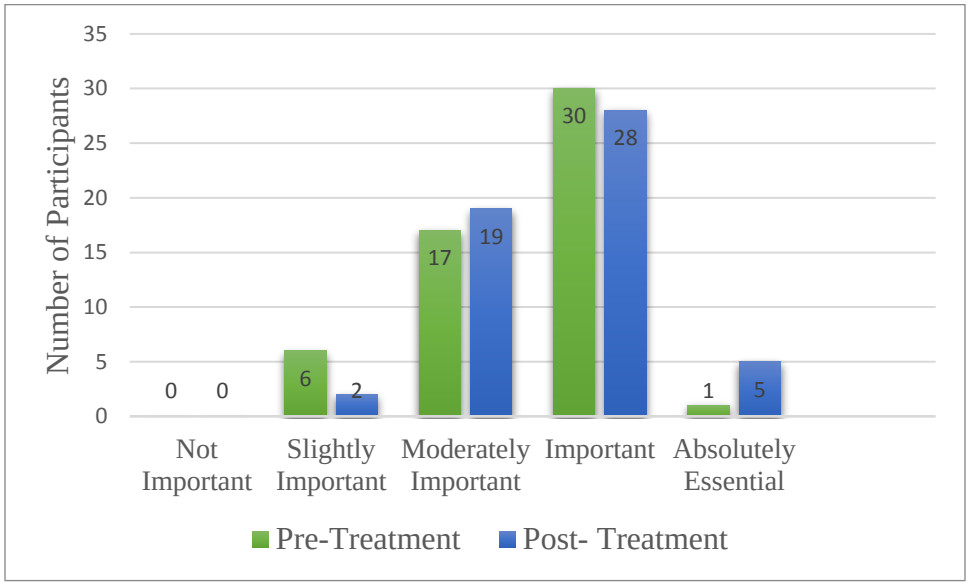


Figure 3. Student responses to survey question, *How important is it for everyone to have data literacy skills, no matter what career they work in?*, (*N*=54).

For the question regarding student attitudes about the importance of data literacy skills to scientists, results showed an increase in the number of students choosing the option *absolutely essential*. Although the Wilcoxon Signed Rank Test did not show this to be significant, the number of students choosing this option increased from 29 pre-treatment to 41 post-treatment, a 41% increase (Figure 4).

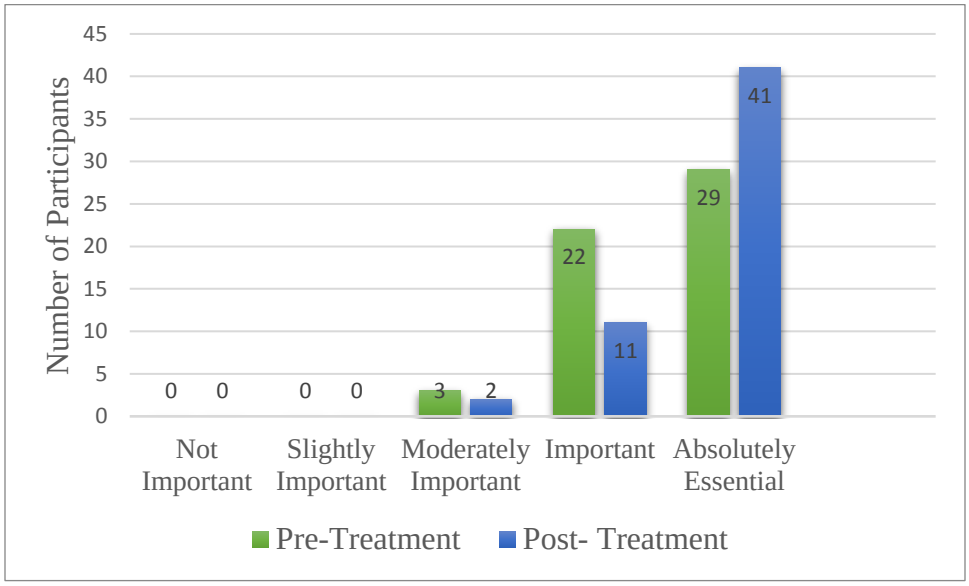


Figure 4. Student responses to survey question, *How important is it for scientists to have data literacy skills, no matter what field of science they work in?*, (*N*=54).

Only 10 students indicated they used data to help them make decisions prior to the instructional unit. These students chose the answer *very often* to the question "How often do you use data to help you make decisions?" on the survey and none chose the option *always*. After the instructional unit, 21 students chose *very often* or *always*, an increase of 110% (Figure 5).

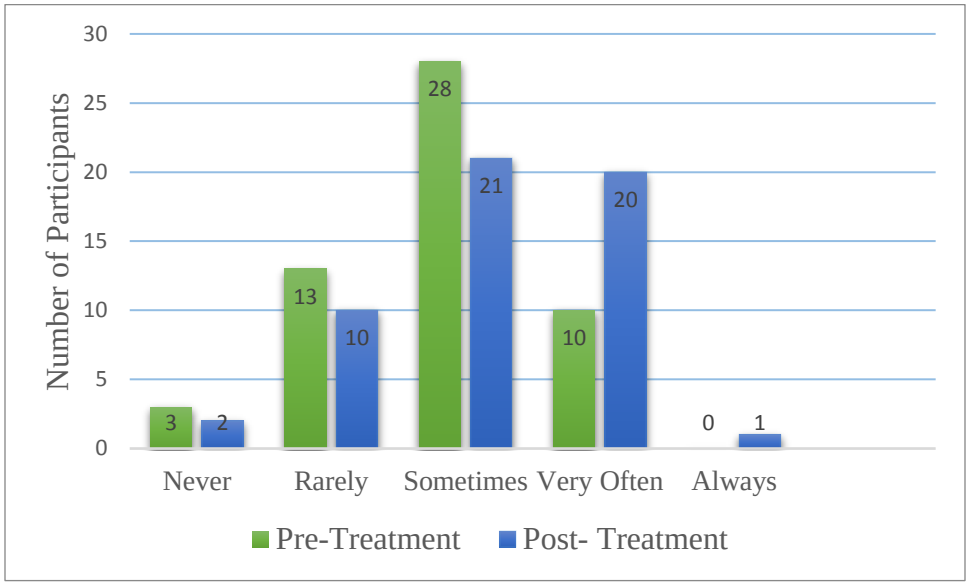


Figure 5. Student responses to survey question, *How often do you use data to help you make decisions?*, (*N*=54).

Before the instructional unit on the Data Literacy Skills Confidence Survey, 15 students rated their skill level as *above average* and none of them rated their overall skills as *excellent*. After the instructional unit, 33 students rated themselves as *above average* or *excellent*, an increase of 120% (Figure 6).

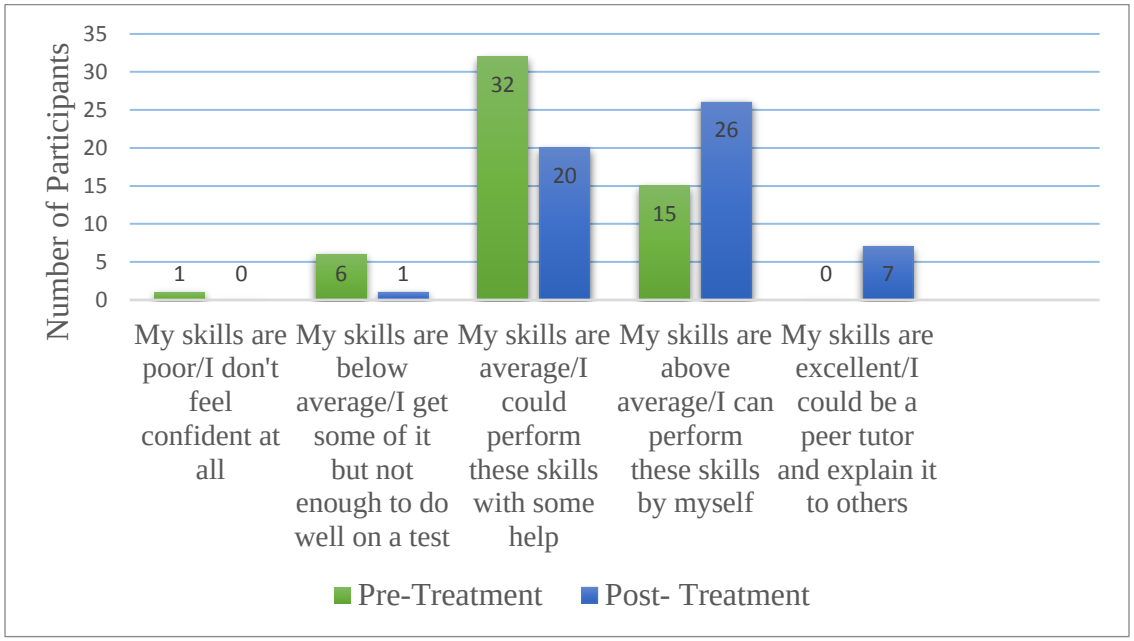


Figure 6. Student responses to survey question, *Rate your overall confidence level in your data literacy skills*, (*N*=54).

Even with the large increase in the number of students rating their skills as *excellent*, the Wilcoxon value showed that the results for this question did not show a statistical difference between pre and post treatment. The Wilcoxon value was also at odds with what students indicated when interviewed individually. Before the instructional unit began, four of the five students interviewed said they felt they were average in their data literacy skills. One student stated, “I can figure most things out if I have the instructions, but I don’t think I could do it on my own.” After the instructional unit, all of the five students said they felt either *pretty confident* or *very confident* overall in their data literacy skills.

Student responses for the second part of the Data Literacy Skills Confidence Survey showed that there was an increase in the number of students who indicated they

had met expectations for independent mastery of the skill by choosing *I could perform this skill by myself* or *I could be a peer tutor* (Table 3).

Table 3
Percentage of Students Indicating Mastery of Individual Data Literacy Skills. This was determined by the number of students choosing "I could perform this skill by myself" or "I could be a peer tutor" on the Data Literacy Skills Confidence Survey, (N=54)

Question on Data Literacy Skills Confidence Survey	Percentage of Students Confident of Mastery Pre-Treatment	Percentage of Students Confident of Mastery Post-Treatment
1. When given a set of data, I can choose what type of graph to make.	59%	72%
2. I can properly make a graph by hand.	78%	98%
3. I can make a box-and-whisker plot by hand.	24%	69%
4. I can interpret a box-and-whisker plot and understand what the data says.	26%	67%
5. When given a set of data, I could perform a *t*-test to statistically analyze and compare two means.	7%	39%
6. I understand the purpose of a *t*-test to compare two means.	7%	33%
7. When I look at a graph, I can determine the trends and make a prediction from what the data shows.	61%	94%
8. When I look at a graph, I can make a claim about what the data shows and use data as evidence for my argument.	74%	93%

Results from the lesson designed to teach students how to make and choose the right kind of graph indicated that student skill levels in this area increased slightly from pre to post treatment. Many students were confident in their ability to choose the right kind of graph before the instructional unit. However, this was before they realized how many different kinds of graphs they would be learning about. One student said, "I always

made a bar or line graph and I thought I knew which one of those to use. The Graph Choice Chart gave us many more options to think about. Now that I know about all the different types of graphs, it is a little overwhelming." Fifteen students (28%) did not feel confident in their abilities in this skill even after the instructional unit. Almost all students (98%), however, were very confident in their ability to properly make a graph following the instructional unit (Figure 7).

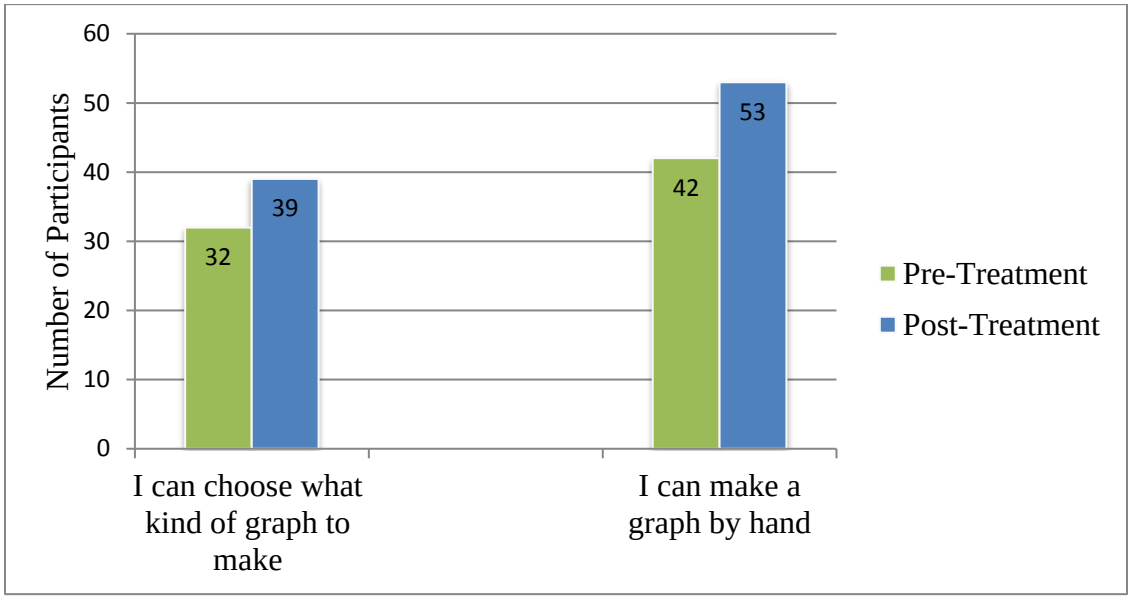


Figure 7. For the survey questions shown, number of students indicating mastery of the skill by choosing either *I can perform this skill by myself* or *I could be a peer tutor and explain it to others*, (*N*=54).

Another subunit of this project was to teach students how to make and interpret box-and-whisker plots. The results of this subunit showed a substantial increase in student confidence in this area. Student confidence levels in mastery of this skill improved from 13 students (23%) to 37 students (65%). Four of the five students interviewed after the unit indicated this was a new skill they had seen but never learned how to do. A few students indicated they would need more practice with box-and-

whisker plots to feel confident in their abilities. One student said, “I would still need to look at the directions in order to do a box-and-whisker plot, but if I had those, I would be pretty confident” (Figure 8).

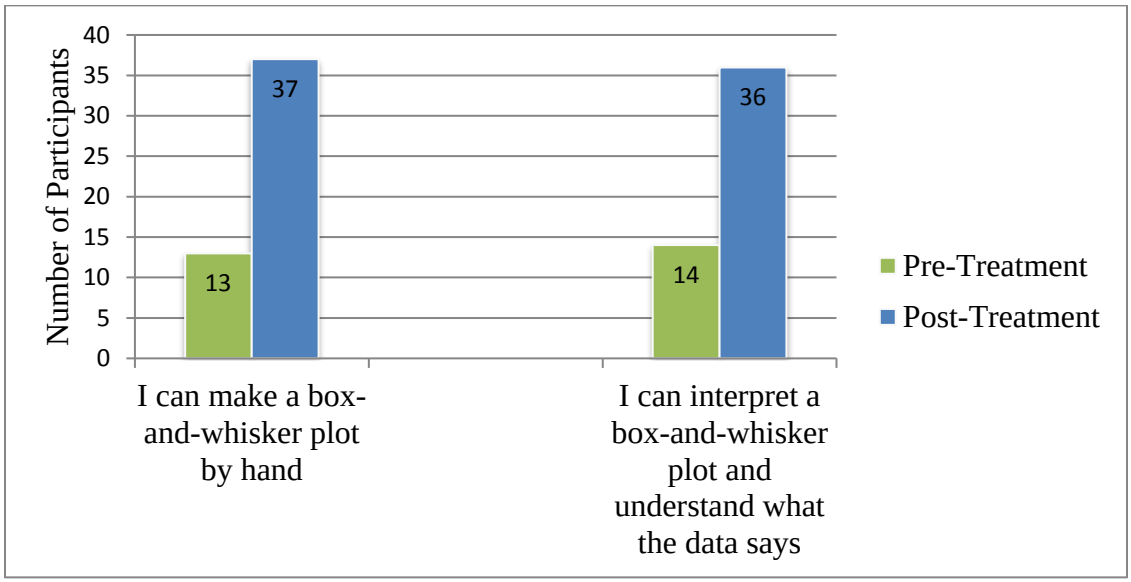


Figure 8. For the survey questions shown, number of students indicating mastery of the skill by choosing either: *I can perform this skill by myself* or *I could be a peer tutor and explain it to others*, (*N*=54).

The instructional subunit on the skill of using a *t*-test to compare two means was one of the more difficult for students. One student said, “I really need to practice this more, but now I realize that using a *t*-test to look at data makes the comparison between groups more accurate. I mean, when you just look at the data, you didn’t realize it was different.” Student confidence levels in the skills of using a *t*-test and understanding the purpose of a *t*-test rose from 4 students to 21 and 18, respectively. Although this is a substantial increase, it means that many students are still not confident in this data literacy skill. One student said, “It was confusing, but with help I could do it. I just need more practice” (Figure 9).

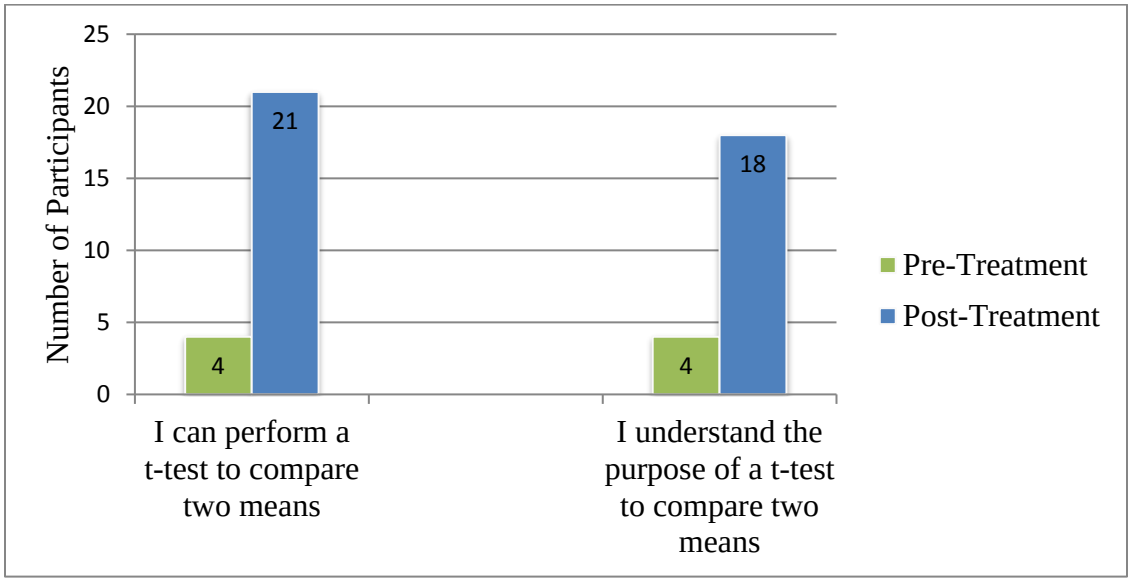


Figure 9. For the survey questions shown, number of students indicating mastery of the skill by choosing either: *I can perform this skill by myself* or *I could be a peer tutor and explain it to others*, (*N*=54).

The final instructional unit was designed to teach students the skills of analyzing data using graphs and arguing claims using evidence from graphs. Many students felt fairly confident in their abilities prior to this instructional unit, although some gains were made in confidence levels, according to the survey results. Student confidence levels in these skills were among the highest of all skills measured in the survey. All students interviewed said they felt confident in their abilities to perform this skill. A student said, "I feel pretty confident that I can make an argument from my graph. I like working in a group because we can compare data and bounce ideas off each other. It makes it easier to figure out what the data means." After the instructional unit, 94% of students indicated confidence in their mastery of the skill of analyzing data using graphs. In addition, 93% indicated mastery of arguing claims using evidence from graphs. (Figure 10).

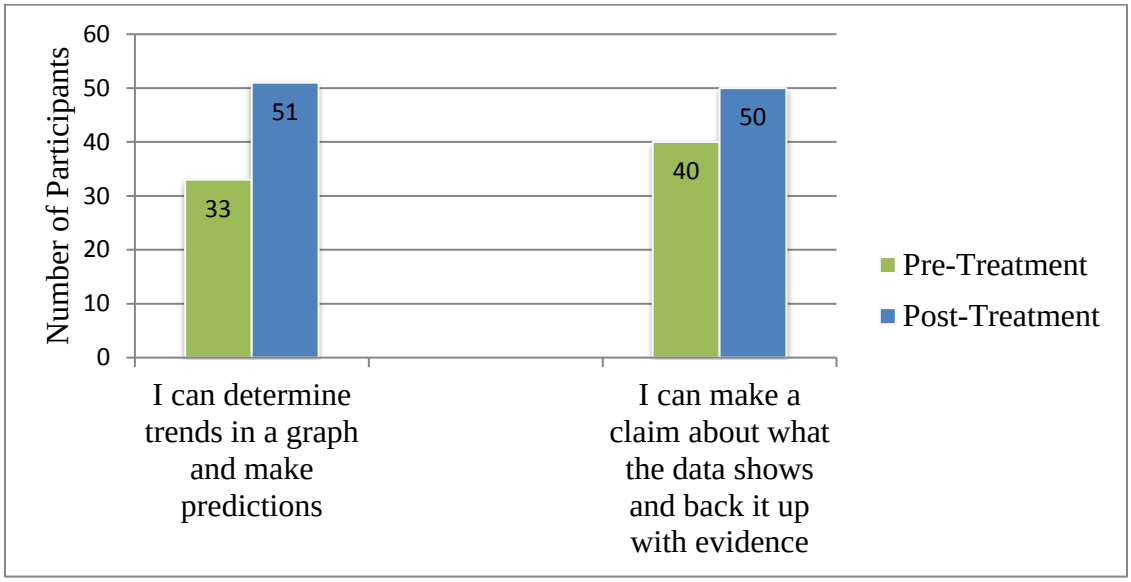


Figure 10. For the survey questions shown, number of students indicating mastery of the skill by choosing either *I can perform this skill by myself* or *I could be a peer tutor and explain it to others*, (*N*=54).

Instructional unit quizzes were given after each particular skill set was taught. For a student to be considered as having mastered the skill, they needed to achieve a score of 67% or higher on the quiz. Scores on the three instructional unit quizzes showed that at least 72% of all students scored at mastery level (Figure 11).

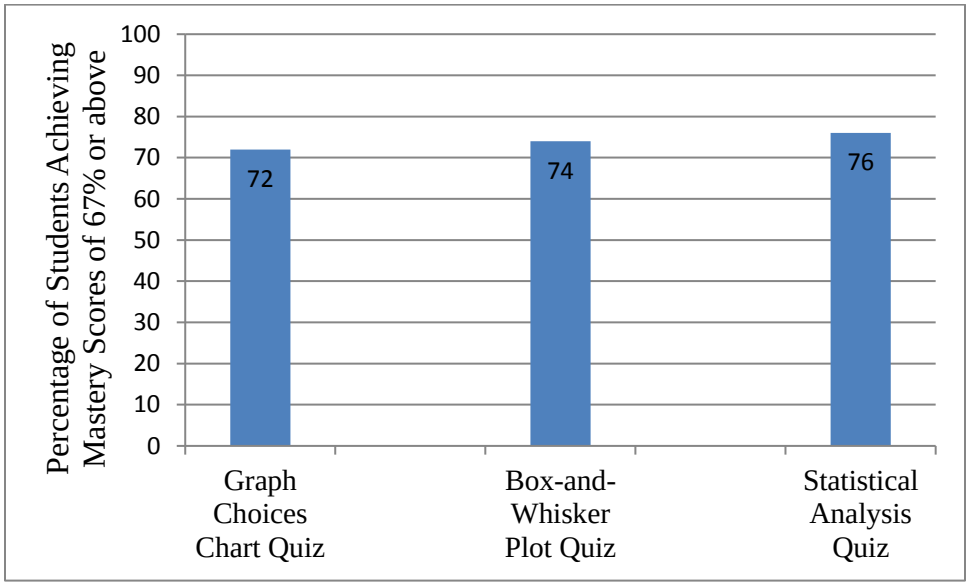


Figure 11. Percentage of students scoring mastery level of 67% or above on the instructional unit quizzes, (*N*=54).

The Graph Choice Chart Quiz had the highest number of students receiving a mastery-level score (80%), with a mean score of 72%. This is consistent with the percentage of students expressing confidence in their mastery of this skill (72%) on the Data Literacy Skills Confidence Survey. The Box-and-Whisker Plot Quiz had a mean score of 68%, with 74% of students receiving a mastery-level score. This is also consistent with the number of students expressing confidence in their mastery of this skill (69%) on the survey. The mean score on the Statistical Analysis Quiz was 69%, with 76% of students receiving a mastery-level score. This is not consistent with students' confidence levels on the survey. Only 39% of students felt confident in their mastery of this skill (Figure 12).

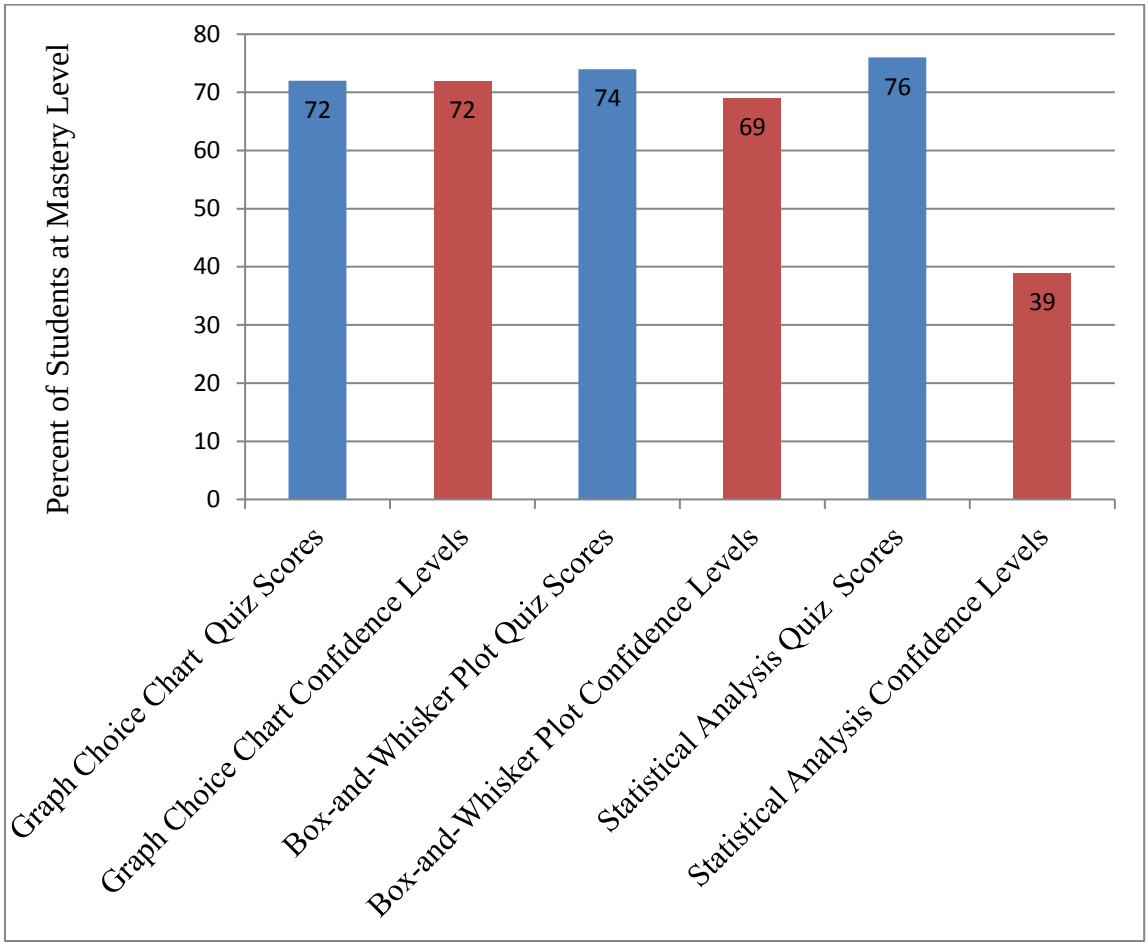


Figure 12. Percent of students scoring at mastery level on quizzes (67% or higher) and percent of students expressing confidence in their mastery of the skill. This was determined by percent of students choosing *I can perform this skill by myself* or *I could be a peer tutor and explain it to others* on the Data Literacy Skills Confidence Survey, (*N*=54).

Trends in student responses during interviews conducted after each instructional unit were analyzed for frequency. Results showed that students were engaged and interested in three of the four lessons designed to teach the data literacy skills. The lesson designed to teach students how to choose the right kind of graph was less interesting for students. One student said, “It was better than taking notes on different kinds of graphs, but it would be better if we got to do the experiment.” In all of the lessons except the graph choice lesson, students responded positively that using real data made it easier to

learn the skills. Almost all students felt at least fairly confident that they could perform the skills by themselves. Some students, however, indicated they were only confident performing the skills with help from an instruction sheet or with more practice (Figure 13).

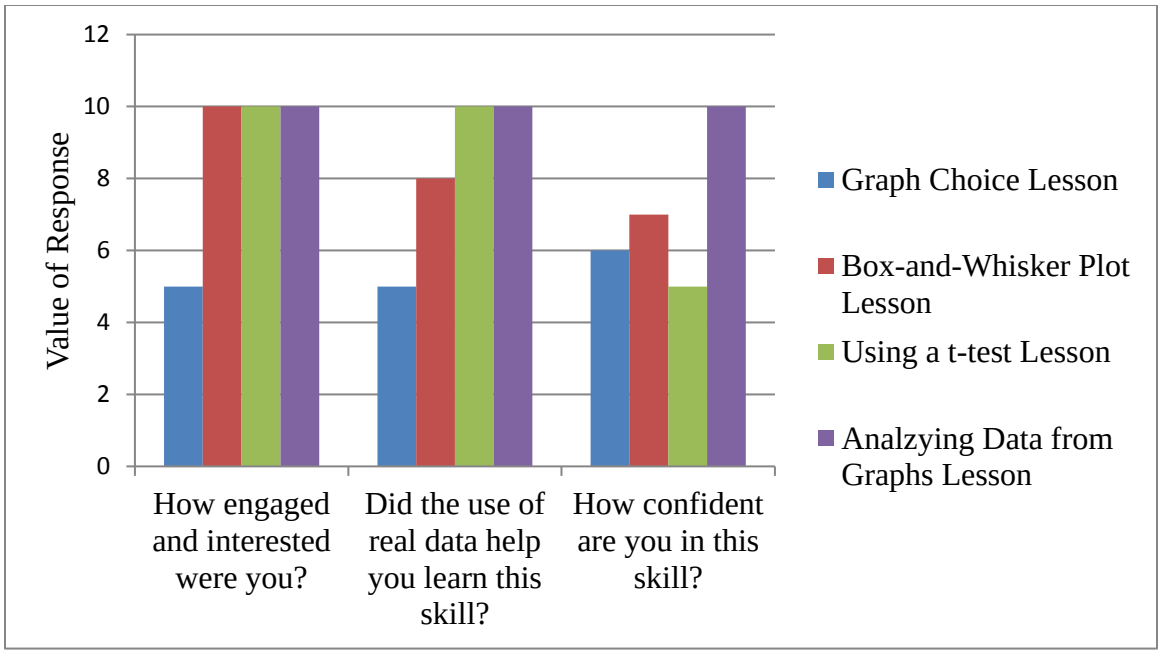


Figure 13. Student interview responses after each instructional unit. Responses were scored on a scale from 1-10, where 1 is no response and 10 is the most positive response, (*N*=54).

INTERPRETATION AND CONCLUSION

Results from the Practical Skills Test, the Data Literacy Skills Confidence Survey, and student interviews all indicated an increase in student confidence and proficiency in data literacy skills. Results from the Data Literacy Skills Test, however, did not show a significant increase. Data from individual unit quizzes have limited usefulness in evaluating the effectiveness of this project.

The Data Literacy Skills Test was a 12-question multiple-choice test which required students to interpret and analyze many different data sets. The scores from pre to post treatment showed only a modest gain of 15.44% in mean scores, with a low normalized gain of 0.24. The small increase in scores could be attributed to a couple of factors. The first thing to be considered is that students just didn't learn as much as they should have during the instructional unit of this project. However, this is not shown in any of the other assessments. The second possibility, then, is that the test does not properly measure student success in the areas that were the focus of this project.

Close analysis of questions on the Data Literacy Skills Test itself reveals some reasons why test scores did not show a substantial improvement from pre- to post-treatment. A number of different types of graphs were used in the test including scatter plots, box-and-whisker plots, double y-axis graphs, line graphs and area graphs. While the test focused on student abilities to analyze the different types of graphs, none of the other skills that were the focus of the project were included on this test. In addition, the instructional units designed to teach individual skills did not address some of the questions that were included on the test. For example, there were two questions that asked students to determine which scatter plots showed a positive relationship or a strong correlation between factors. These are important concepts students should understand when interpreting a scatter plot, but the instructional unit on graphing did not specifically teach this skill.

Other questions on the Data Literacy Skills Test required students to read some information about a case study and analyze the associated set of graphs. Students were

unfamiliar with some of the vocabulary terms used in either the readings or in the captions under the graphs. For example, one graph showed character displacement in beak size of finches before and after a drought. Students did not know what character displacement was and so the question related to that graph was confusing to them. In another example, the terms introgressive hybridization and morphological convergence were used to describe variations in patterns of songs used by male and female finches. Although these terms were defined in the reading, students may not have applied these definitions to their interpretation of what the graph showed. These are terms that had not been introduced to students prior to the test. One student told me after the post test, "I really didn't understand what the question was asking me to do." Several students asked questions about the meanings of various vocabulary terms while they were taking the post-test. Because so many of the questions on the test were confusing, it is difficult to apply the results of these test scores towards this capstone project.

In addition to the flaws in the questions mentioned above, the test did not include questions pertaining to other data literacy skills that were the focus of the instructional units for this project. There were no questions on the test requiring students to choose the right kind of graph, to show that they understand the purpose of a *t*-test, or to interpret results from a *t*-test. In summary, there were questions on this test that should not have been included and there were other questions that should have been there.

The Practical Skills Test did show substantial improvement in student skill levels. The mean scores on this test improved 44.36% with a high normalized gain of 0.73. For this test, students were given a case study with a set of data. They were required to

choose and make the right kind of graph, perform some basic statistical analysis of the data, and make a claim about what the data showed using evidence from the graph. This test required students to perform most of the basic data literacy skills that were taught in the instructional units. The only skill not assessed in this test was using and understanding the purpose of a *t*-test. Because this test was more indicative of student progress in gaining data literacy skills, results from this test are more useful. Based on these results, it shows that the instructional unit was successful in reaching the goals of this capstone project.

One of the goals of this capstone project was to see if the teaching of data literacy skills improved student confidence levels in organizing and analyzing data. The results from the Data Literacy Skills Confidence Survey from pre- to post-treatment showed only a slight increase in student attitudes regarding the importance of data literacy skills and how often they use these skills in their lives. Although these are important indicators of student growth and change in global views of the importance of data literacy, the instructional units in this capstone were not designed to affect these particular student attitudes.

Overall student confidence in their data literacy skill levels, however, did show an impressive increase. The number of students who feel they can perform these skills by themselves or that they could be a peer tutor and explain it to others improved from 28% to 61% from pre- to post-treatment. The instructional unit was effective at improving students' overall confidence levels.

The Data Literacy Skills Confidence Survey also asked students about their confidence regarding individual data literacy skills. It was satisfying to see that there was an increase in the number of students who feel they have mastered the skill in every category, although the results varied from slight to large improvements. There were only small improvements in the skills of choosing and making the right kind of graph and using evidence from a graph to support a claim. This is primarily because many students already felt confident in these areas before the instructional unit.

The most substantial improvement in confidence levels was seen in students' assessments of their ability to make and interpret a box-and-whisker plot, perform a *t*-test, and understand the purpose of a *t*-test. It was also gratifying to see that almost all students felt confident in their abilities to properly make a graph (98%), determine trends and make predictions from graphs (94%), and to make a claim about what the data shows and use data as evidence for their argument (93%). The instructional unit was effective at improving student confidence levels in individual data literacy skills.

Individual unit quizzes were given after each of the lessons that taught a particular data literacy skill. These included the Graph Choice Chart Quiz, the Box-and-Whisker Plot Quiz, and the Statistical Analysis Quiz. There were no pre-tests given for these individual unit quizzes, so there is no way to measure differences from pre to post treatment. Therefore, considering the scores on these individual quizzes towards gaining an understanding of the effectiveness of the instructional unit is questionable. If it is considered that a 67% score on the quiz shows mastery of the skill, then approximately 25% of students did not show mastery of the skills measured in these quizzes. In reality,

an argument could be made that a 67% score on a quiz or test does not actually show mastery. Assigning a particular percentage as a passing score to measure mastery is really an arbitrary measurement of student success. Analyzing results from the individual unit quizzes has a limited function in understanding successful improvement toward the capstone project's goals.

A more important measure of student success came from analyzing trends in the interviews conducted with students after each instructional unit. The statements made by students after each unit were important indicators of the success of each lesson and gave insights into why some lessons were more effective than others.

The Graph Choice Chart lesson, for example, was not as interesting or engaging for students as it could have been. A couple of students thought that this lesson was difficult and somewhat tedious. They also said, however, that this was because the concept itself was uninteresting to them, not necessarily that the lesson designed to teach it was boring. One important insight gained was that it would have been more interesting if the lesson involved real experimentation. Instead, for this lesson, students were required to design an experiment and chose what type of graph they would use, but they didn't actually perform the experiment. A suggestion for improving this lesson was to allow students to design and perform an actual experiment related to the environment around the school. One student said, for example, "Any real data is better. It makes it more interesting."

The effectiveness of using real data to help students learn the various data literacy skills was a common theme in all of the student interviews. Students really appreciated

the chance to apply a new skill to data they had collected themselves. This was especially true for the protist diversity lab where students calculated the difference between two means using a *t*-test. “This was a very cool lab and using the *t*-test to analyze the data was interesting. Once you collected your data, you wanted to find out how it ended up and whether your results were significant or not.” Another student agreed that using real data was inherently more interesting than learning the skill from a given set of data. “I know I could have learned how to do a *t*-test or make a box plot without collecting the data myself, but it was more interesting. I also really understand now why you would need to do a *t*-test or make a box plot to look at your data.” Student interviews revealed that the use of real data did help students learn data literacy skills. Not only did the use of real data increase their skill levels and confidence, it also helped them be more engaged and interested in the lesson.

Overall, the results of the assessments for this capstone project indicate that incorporating data literacy lessons into a biology curriculum improves student confidence levels and helps students gain proficiency in data literacy skills. By incorporating the use of real data into lessons on biology concepts that they were already learning, students were more engaged and interested in their learning.

VALUE

The most significant thing I learned from this capstone project was the importance of using real data in my biology lessons. Every lesson was enhanced by incorporating either the collection of data or the use of real data from actual research. Students were excited and engaged with each lesson. Because they were actively involved, they began

to see that the entire process of collecting, organizing and analyzing the data is important. Each step was an important part of helping them understand the message the data is telling them. By incorporating data literacy skills along with biology concepts, instruction in these skills flowed naturally and didn't seem to be separate lessons. The process was authentic and real to my students. One student told me, "Learning this way was fun and easy. It wasn't too hard, so I didn't get frustrated and give up." Another told me that these lessons were "more personal because I got the data myself. I feel like it is important to me now in a way that it wouldn't be if I didn't do it myself."

The process was not without some problems along the way. I realized that the graph choice lesson was not nearly as effective as I had hoped. Because students only designed possible experiments they could do but didn't actually perform the experiments, they didn't realize the true value of using different kinds of graphs to display their data. Students were not as engaged or interested in finding the solutions to the problem of figuring out the best type of graph to use because they didn't have to. The scenario they proposed was just hypothetical. It wasn't a real experiment and therefore, it didn't mean anything to them. They weren't invested in the process other than to get a grade or do well on the quiz. Besides, they thought they already knew what type of graph they should use. They really thought that a line graph or a bar graph would be enough for basically any type of experiment. Past experience and previous science classes had taught them that line graphs and bar graphs would be all they would ever need to learn about. One student told me, "Really, no kid wants to learn about graphs. I just always used a line or bar graph as a default. I never really thought about it." It appeared that students didn't

really want to learn the value of using different kinds of graphs or what parameters would lead them to choose a box-and-whisker plot or a scatter plot instead of their go-to choice of a bar graph or line graph.

Many lessons I have used in the past started with hypothetical data or a data set with predetermined outcomes. This type of lesson is so common in biology classrooms and I should have known from the start that using hypothetical data would be less effective with my students. When I do this lesson again, I will have my students actually design and conduct their own experiments first. My students told me that they thought the Graph Choice Chart was a great tool. I will continue to use it because it will help my students learn that there are great alternatives to display data from different kinds of experiments. However, it became clear to me that there is no substitute for using real, authentic data to help students understand the importance of learning about the different kinds of graphs.

The lessons designed to incorporate data literacy did take a little more time to complete than a simple lesson on biodiversity or genetic variation. Time is often a concern when a lesson is more open-ended or inquiry-based. Every minute of class time with students is important and effort must be made to have the labs set up in advance and outline clear expectations in advance for what students will need to do. Although class time could be shortened greatly by giving students sample data, as I discovered with the graph choice unit of instruction, this is boring and irrelevant to students. Since class time is such an issue, there are ways of shortening the instructional unit. For one thing, I would introduce the Graph Choice Chart at the beginning of the year. Then, as the year

progressed, and we did different experiments for different units, I would have them learn a new kind of graph. By the end of the year, hopefully, they will have used the chart so often that it will become second nature to them. Students could receive a quick introduction to graphing using Excel and do their graphs as homework, which would use less class time than graphing by hand.

Although I truly believe data literacy skills are important, I cannot justify teaching them at the expense of teaching important biological concepts. Therefore, they absolutely must be taught only in conjunction with the existing biology curriculum. My students are still required to do well on standardized state science exams and I would be doing them a disservice to ignore these important topics. This capstone project has allowed me to develop some excellent lessons that teach biological concepts along with data literacy skills. In the future, I think that many more of my biology lessons could be rewritten to incorporate additional practice. I realize that my primary goal is to help my students learn biology concepts, but my hope is that I can also help them become scientifically literate citizens ready for their future.

Since data literacy skills are never taught in a stand-alone class, it is important that teachers from middle school to high school incorporate these skills into their own curriculum at every level. At the high school level, students in earth science, chemistry, physics and Advanced Placement courses would all benefit from specific data literacy skill instruction. If lessons such as these that I developed for my biology class were readily available, teachers could more easily integrate these into their own repertoire of lessons.

Prior to this capstone project, my curriculum did not include basic statistical analysis tools or more advanced graphing skills. As a result of this project, my own data literacy skills have improved and I feel more confident in my ability to teach these skills. I learned the value of incorporating the acquisition and use of authentic data into my lessons as much as possible. I realize that it is critical to take the time to allow these types of labs and experiments if I wish to fully engage my students in their learning. This capstone project really taught me that it is the process of science that is most important. For my students to become scientifically literate citizens, they need the opportunity to learn and practice the skills that real scientists use.

REFERENCES CITED

Atkinson, A., Siegel, V., Pakhomov, E., & Rother, P. (2004). Long-term decline in krill stock and increase in salps within the Southern Ocean. *Nature*, *432*, 100-103.

Bowen, M. & Bartley, A. (2014). *The Basics of Data Literacy: Helping Your Students (and You!) Make Better Sense of Data.* Arlington, VA. NSTA Press.

British Antarctic Survey. (2004, November 18). Antarctic Meteorology Online from the British Antarctic Survey. Retrieved from https://legacy.bas.ac.uk/met/metlog/

Cary Institute for Ecosystem Studies. (2015). Changing Hudson Project | Cary Institute of Ecosystem Studies. Retrieved from http://www.caryinstitute.org/educators/teaching-materials/changing-hudson-project

Clary, R. & Wandersee, J. (2013). Arguing history. *The Science Teacher, 80*(5), 39-43.

Common Core State Standards (2010). Reading standards for literacy in science and technical subjects 6-12. *Common Core State Standards Initiative.* Retrieved from http://www.corestandards.org/wp-content/uploads/ELA_Standards. pdf

Constible, J., Sandro, L., & Lee, R.E. Jr. (2007). A Cooperative Classroom Investigation of Climate Change. *The Science Teacher, 74(6),* 37-43.

Emery, K., Harlow, D., Whitmore, A., & Gaines, S. (2014). Confronting ambiguity in science. *The Science Teacher, 82(2),* 36-42.

Gould, R., Sunbury, S., & Dussault, M. (2014). In praise of messy data. *The Science Teacher, 81(8),* 31-36.

Grant, P., & Grant, R. (2008). Fission and fusion in Darwin's finches. *Philos Trans R Soc Lond B Biol Soc, 363*(1505), 2821-2829. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18508750

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *Am. J. Phys*, *66*(1), 64. doi:10.1119/1.18809

Harris, J. (2012, September). Data is useless without the skills to analyze it. *Harvard Business Review,* Retrieved from https://hbr.org/2012/09/data-is-useless-without-the-skills.

Hougen-Eitzman, D. (2007). Teaching data analysis and presentation (as well as some statistics) in an introductory biology class. In O'Donnell, M. A. (Ed.), *Tested*

studies for Laboratory Teaching, Volume 28, Proceedings of the 28th Workshop/Conference of the Association for Biology Laboratory Education (ABLE), (385-388). Retrieved from www.ableweb.com

Intergovernmental Panel on Climate Change (IPCC). (2007). *The physical science basis: Summary for policy makers.* Working Group I Report. Retrieved from www.ipcc.ch

Konold, C. & Higgins, T. (2003). In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to Principles and Standards for School Mathematics,* Reston, VA: National Council of Teachers of Mathematics, pp. 193-215. Retrieved from https://www.srri.umass.edu/sites/srri/files/Konold_Higgins2003.pdf

Llewellyn, D. (2013). Making and defending scientific arguments: Strategies to prepare your students for the new wave of curricula reform. *The Science Teacher, 80(5),* 34-38.

Manduka, C. A., & Mogk, D. W. (2002, April). Using data in undergraduate science classrooms: final report on an interdisciplinary workshop held at Carleton College. Northfield, MN. Science Education and Resource Center. Retrieved from http://serc.carleton.edu/usingdata

McNeill, K T., & Pimentel, D. S. (2010). Scientific discourse in three urban classrooms: The role of the teacher in engaging high school students in argumentation. *Science Education, 94(2),* 203-229. Retrieved from DOI 10.1002/sce.20364

Metz, S. (2005). Embracing the messiness of data. *The Science Teacher, 73(7),* 8.

National Research Council. (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas.* Committee on a Conceptual Framework for New K-12 Science Education Standards. Board on Science Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

NGSS Lead States. (2013). *Next General Science Standards: For States, By States.* Washington, DC: The National Academies Press.

Norris, S., Stelnicki, N., & de Vries, G. (2012). Teaching mathematical biology in high school using adapted primary literature. *Research in Science Education, 42(4),* 633-649.

Palmer LTER Data Archive. (n.d.) Retrieved December 6, 2015, from http://pal.lternet.edu/data

Smith, R., Fraser, R. & Stammerjohn, S. (2003). Climate variability and ecological response of the marine ecosystem in the Western Antarctic Peninsula (WAP) region. In *Climate variability and ecosystem response at Long-Term Ecological Research (LTER) sites,* eds. Greenland, D., Goodin, D., & Smith, R.

Tenopir, C., & King, D., (2004). *Communication patterns of engineers.* Hoboken, NJ: Wiley.

Webber, H., Nelson, S., Weatherbee, R., Zoellick, B.,and Schauffler, M., (2014), The Graph Choice Chart: A tool to help students turn data into evidence, *The Science Teacher, 18(8),* 37-43.

APPENDICES

APPENDIX A

IRB EXEMPTION APPROVAL FORM

INSTITUTIONAL REVIEW BOARD

For the Protection of Human Subjects
FWA 00000165

MONTANA STATE UNIVERSITY

960 Technology Blvd. Room 127
c/o Immunology & Infectious Diseases
Montana State University
Bozeman, MT 59718
Telephone: 406-994-6783
FAX: 406-994-4303
E-mail: cherylj@montana.edu

Chair: Mark Quinn
406-994-5721
mquinn@montana.edu
Administrator:
Cheryl Johnson
406-994-6783
cherylj@montana.edu

MEMORANDUM

TO: Patricia Jones and John Graves

FROM: Mark Quinn, Chair Mark Quinn CJ

DATE: December 10, 2015

RE: *"Teaching Data Literacy in High School Biology" [PJ121015-EX]*

The above research, described in your submission of December 10, 2015, is exempt from the requirement of review by the Institutional Review Board in accordance with the Code of Federal regulations, Part 46, section 101. The specific paragraph which applies to your research is:

_X_ (b) (1) Research conducted in established or commonly accepted educational settings, involving normal educational practices such as (i) research on regular and special education instructional strategies, or (ii) research on the effectiveness of or the comparison among instructional techniques, curricula, or classroom management methods.

_X_ (b) (2) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, unless: (i) information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and (ii) any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability, or be damaging to the subjects' financial standing, employability, or reputation.

____ (b) (3) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior that is not exempt under paragraph (b)(2) of this section, if: (i) the human subjects are elected or appointed public officials or candidates for public office; or (ii) federal statute(s) without exception that the confidentiality of the personally identifiable information will be maintained throughout the research and thereafter.

____ (b) (4) Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available, or if the information is recorded by the investigator in such a manner that the subjects cannot be identified, directly or through identifiers linked to the subjects.

____ (b) (5) Research and demonstration projects, which are conducted by or subject to the approval of department or agency heads, and which are designed to study, evaluate, or otherwise examine: (i) public benefit or service programs; (ii) procedures for obtaining benefits or services under those programs; (iii) possible changes in or alternatives to those programs or procedures; or (iv) possible changes in methods or levels of payment for benefits or services under those programs.

_____ (b) (6) Taste and food quality evaluation and consumer acceptance studies, (i) if wholesome foods without additives are consumed, or (ii) if a food is consumed that contains a food ingredient at or below the level and for a use found to be safe, or agricultural chemical or environmental contaminant at or below the level found to be safe, by the FDA, or approved by the EPA, or the Food Safety and Inspection Service of the USDA.

Although review by the Institutional Review Board is not required for the above research, the Committee will be glad to review it. If you wish a review and committee approval, please submit 3 copies of the usual application form and it will be processed by expedited review.

APPENDIX B

DATA LITERACY SKILLS CONFIDENCE SURVEY

Please answer all questions as honestly as possible. This survey will not be graded. Your participation in this survey is entirely voluntary and no penalty will be given if you do not participate. *NOTE: For the purposes of this survey, data literacy skills include: being able to create the appropriate type of graph, being able to make and interpret data from box-and-whisker plots, analyze data using statistics and being able to argue a claim based on evidence from data. Please circle your response.*

1. How important is it for everyone to have data literacy skills, no matter what career they work in?
    0 = Not important at all
    1 = Slightly important
    2 = Moderately important
    3 = Important
    4 = Absolutely essential
2. How important is it for scientists to have data literacy skills, no matter what field of science they work in?
    0 = Not important at all
    1 = Slightly important
    2 = Moderately important
    3 = Important
    4 = Absolutely essential
3. How often do you use data to help you make decisions? (For example, how often do you look at data when you are deciding what to purchase?)
    0 = Never
    1 = Rarely
    2 = Sometimes
    3 = Very Often
    4 = Always
4. When you read an article or hear something on TV or the internet, how often do you question the statistics you have heard? (For example, if you hear that a certain product produces amazing results, do you hear that and question whether the data supports their claim?)
    0 = Never
    1 = Rarely
    2 = Sometimes
    3 = Very Often
    4 = Always
5. Rate your overall confidence level in your data literacy skills. (For example, how comfortable and confident do you feel about your own data literacy skills?)
0—My skills in this area are poor/I don’t feel confident at all.
1—My skill level is below average/ I get some of it but not enough to do well on a test.
2—My skill level is average/ I could perform these skills with some help.
3— My skill level is above average. I can perform these skills by myself
4—My skill level is excellent. I could be a peer tutor and explain it to others.

PART 2: *For this section of the survey, you will be rating your own data literacy skills. Please continue to be honest with your responses. This survey will not be graded.*
For the following questions, please rate your level of confidence with the following skills using this scale:

0—I don't understand this at all/I've never heard of this.
1—I get some of it but not enough to do well on a test.
2—I could perform this skill with some help from my team.
3—I can perform this skill by myself/ I am ready for the test
4—I could be a peer tutor and explain it to others.

1. When given a set of data, I can choose what type of graph to make.
   0 1 2 3 4
2. I can properly make a graph by hand.
   0 1 2 3 4
3. I can make a box-and-whisker plot by hand.
   0 1 2 3 4
4. I can interpret a box-and-whisker plot and understand what the data says.
   0 1 2 3 4
5. When given a set of data, I could perform a t-test to statistically analyze and compare two means.
   0 1 2 3 4
6. I understand the purpose of a *t*-test to compare two means.
   0 1 2 3 4
7. When I look at a graph, I can determine the trends and make predictions about what the data shows.
   0 1 2 3 4
8. When I look at a graph, I can make a claim about what the data shows and use data as evidence for my argument.
   0 1 2 3 4

APPENDIX C

PRE AND POST STUDENT INTERVIEW QUESTIONS

Please note: Your participation in this interview is entirely voluntary and not required. No penalty will be given if you do not participate or if you do not want to answer one or more of the questions.

1. How confident do you feel in your ability to communicate and analyze data? In other words, how confident do you think you are at making graphs, analyzing data, using basic statistics, and making claims from evidence in the data?

2. Do you feel it is important that you have data literacy skills? Why or why not?

3. Where would you use data literacy skills in your life? Can you give me an example?

ADDITIONAL INTERVIEW QUESTIONS FOLOWING EACH LESSON:

1. What did you learn from today's activity?

2. How did the activity improve your skill in _____(the skill we worked on)?

3. How could this activity be improved to help you learn this skill better?

4. How engaged and interested were you in the lesson today? Rate your level of engagement in the lesson from 1-10.

5. How would you compare how interesting and engaging this lesson was compared to other lessons from this class?

6. Did the use of real data help you learn this skill easier? If so, in what ways did the use of real data help you? Rate from 1-10 how the use of real data helped you learn this skill.

7. Why do you think the skill you learned today is important in science?

8. How confident do you feel in your ability to perform this skill? Rate from

1-10 how confident you are in your ability to perform this skill.

9. Is there anything else you can tell me about your participation in the activity today?

APPENDIX D

PRACTICAL SKILLS TEST

DIRECTIONS: CHOOSE ONE OF THE FOLLOWING CASE STUDIES:

1. A researcher for the Arizona Department of Agriculture was measuring the effectiveness of applying natural fertilizer such as composted manure to tomato crops. His experiment involved planting several rows of tomatoes: 1/3 of which received no fertilizer, 1/3 received a commercial fertilizer called TopFlight, and 1/3 received composted manure. All of the tomatoes were the same age and variety. All were planted in the same plot of soil and received the same amount of water and sunlight. At the end of the growing period, he harvested tomatoes from 8 representative plants of each type and recorded their yield in pounds of tomatoes. Here is the result of this study:

No Fertilizer, lbs.	TopFlight fertilizer, lbs	Composted Manure, lbs.
10	12	14
9	11	8
11	9	8
7	13	9
9	10	10
6	11	8
10	12	7
5	9	11

1. What is the mean number of pounds for each treatment?

No fertilizer_____TopFlight fertilizer_____Composted Manure_____

2. What is the range for each treatment?

No fertilizer_____TopFlight fertilizer_____Composted Manure_____

3. Determine the median for each treatment:

No fertilizer_____TopFight fertilizer_____Composted Manure_____

4. What is the best type of graph to show this data? Choose from: scatter plot, line graph, multiple line graph, bar graph, stacked bar chart, box-and whisker plot

5. Graph the data.

7. Which is the best method to grow tomatoes?

_____

8. Justify your claim with evidence from the graph.

_____

_____

_____

2. A student in Honors Biology wanted to determine if there was a relationship between the arm and leg lengths of students in their class. Here are the results of this investigation:

Student	Arm Length (cm)	Leg Length (cm)
1	74	81
2	81	89
3	64	70
4	68	76
5	76	85
6	79	86
7	64	70
8	83	91
9	75	84
10	88	95
11	61	66
12	84	93
13	71	78
14	80	86
15	76	85

1. What is the mean for arm length? _____for leg length? _____

2. What is the range for arm length? _____for leg length? _____

3. What is the median arm length?_____median leg length?_____

4. What is the best kind of graph to show this data? Choose from: scatter plot, line graph, multiple line graph, bar graph, stacked bar chart, box-and whisker plot

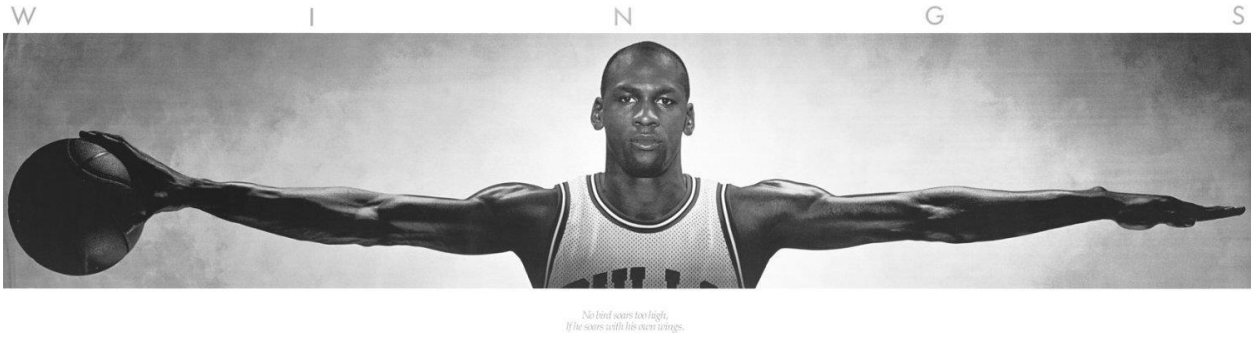
5. Graph the data.

6. What is the relationship between arm length and leg length?

_____

_____

_____



3. Josh has been training his dogs to compete in a local dog show. He is trying to determine which dog would be the best to compete in the obstacle course. Dogs run over, under and through various obstacles and the fastest dog to complete the course correctly wins. Here are the results of the trials he ran with his four dogs. Each dog ran a short obstacle course five times (with rest times in between). Josh timed each dog in seconds.

Trial	Rusty	Lizzie	Toby	Chance
1	14 s	17 s	8 s	10 s
2	13 s	10 s	9 s	8 s
3	13 s	16 s	6 s	9 s
4	15 s	8 s	8 s	11 s
5	17 s	9 s	7 s	9 s

1. What is the mean time for each dog?

Rusty_____Lizzie_____Toby_____Chance

2. What is the range in time for each dog?

Rusty_____Lizzie_____Toby_____Chance

3. What is the median time value for each dog?

Rusty_____Lizzie_____Toby_____Chance

4. What is the best kind of graph for this data? Choose from: scatter plot, line graph, multiple line graph, bar graph, stacked bar chart, box-and whisker plot

5. Graph the data

6. Which dog should Josh take to the obstacle course race? _____

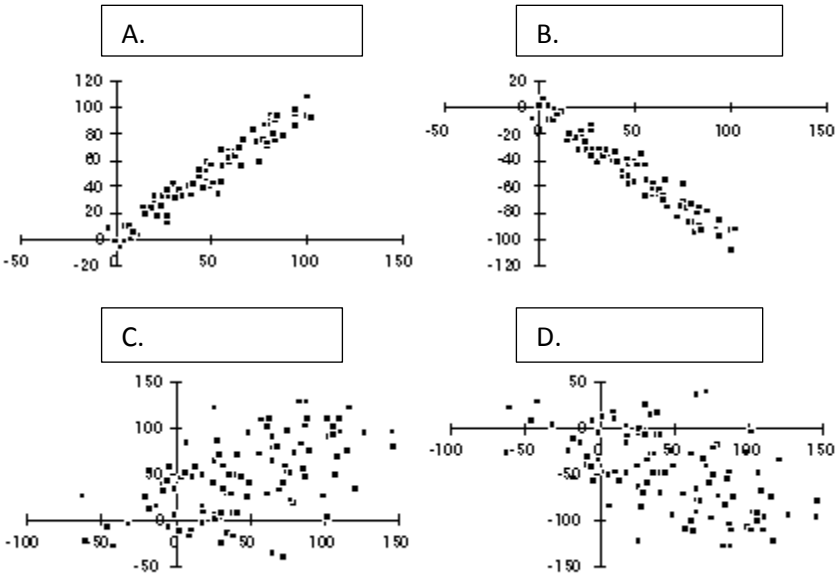
7. Justify your claim with evidence from the graph._____
_____
_____

APPENDIX E

DATA LITERACY SKILLS TEST

Multiple Choice

For Questions 1 and 2, observe the following graphs:



_____1. Which of the graphs shown above indicate a positive relationship between variables?

A. A only

B. B only

C. C only

D. A and C

E. B and D

_____2. Which of the graphs shown above show a strong correlation between variables?

A. A only

B. B only

C. A and B

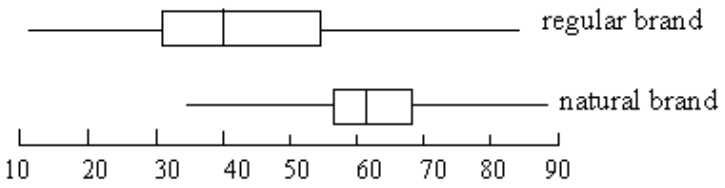
D. A and C

E. B and D

For Questions 3 and 4, study the following graphs:

For the following graphs, a grocery store manager asked 200 of his customers to rate their favorite peanut butter on product quality. The scale measures customers responses as to which peanut butter brand they would buy most often.

Comparing quality ratings between brands of peanut butter



_____3. What is the lower quartile of quality ratings in the natural brand of peanut butter?

A. 58-62

B. 34-58

C. 34-62

D. 34-68

_____4. Why is there a wider box for the regular brand than the natural brand?

A. There were more people taking the survey.

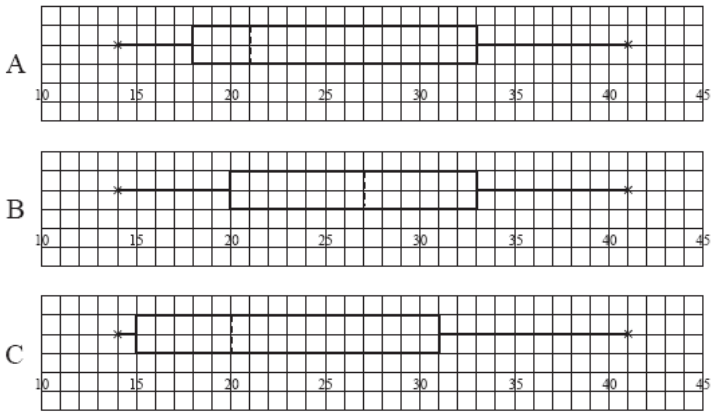
B. There were fewer people choosing the regular brand.

C. There was a wider range of responses from the survey.

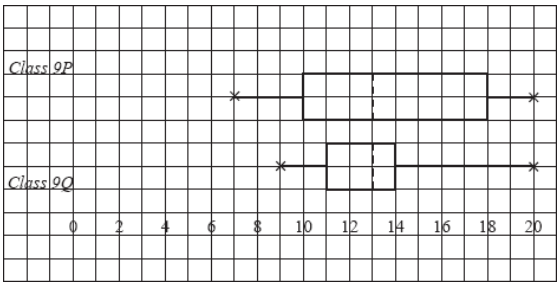
D. More people preferred the regular brand.

_____5. Choose one of the box and whisker plots for the following sample:

17 22 18 33 14 36 39 41 25 31 18 19 16 21 21



For Question 6, read the following information: A math test is given to two 9th grade classes. Here is a graph showing the results of the test.



_____6. What can you tell about the results of this test when you compare the two classes?

A. Class 9Q did better overall on this test because the scores of all students were more closely the same.

B. Class 9P did better overall on this test because the range of scores was the greatest.

C. Both classes had the same median score, so there is no significant difference between the two classes.

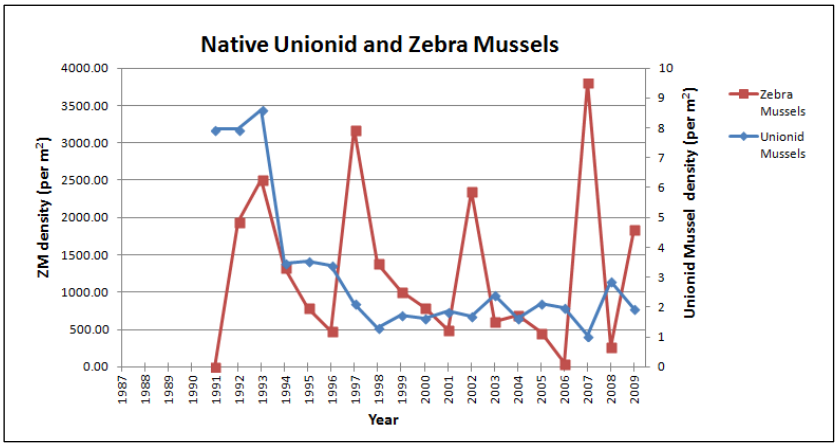
D. Class 9P did better overall because more of the students scored in the upper quartile.

For Questions 7-9, read the following passage.

Scientists have been studying the effects of the invasive Zebra Mussel (*Dreissena polymorpha*) on native populations of fish, invertebrates, plankton and other mussels. This small clam-like mollusk was introduced accidentally into the Great Lakes in 1988. Since then, their population numbers have increased exponentially and they have spread to waterways all over the eastern United States and as far south as Arkansas. Female Zebra Mussels can lay up to 500,000 eggs in one year and they have few natural predators or diseases. These fingernail-sized animals are filter feeders, filtering out food particles from the water. They compete directly with many species of native animals. The Cary Institute of Ecosystem Studies has been gathering data on the Hudson River in New York since 1991. All graphs are from: http://www.caryinstitute.org/educators/teaching-materials/changing-hudson-project/ecosystems-action-population-community-dynami-3

Figure 1: Changes in population of native mussels and Zebra mussels from 1991-2009. →

NOTE: Zebra Mussel (ZM) density is measured at the peak density for that time period.



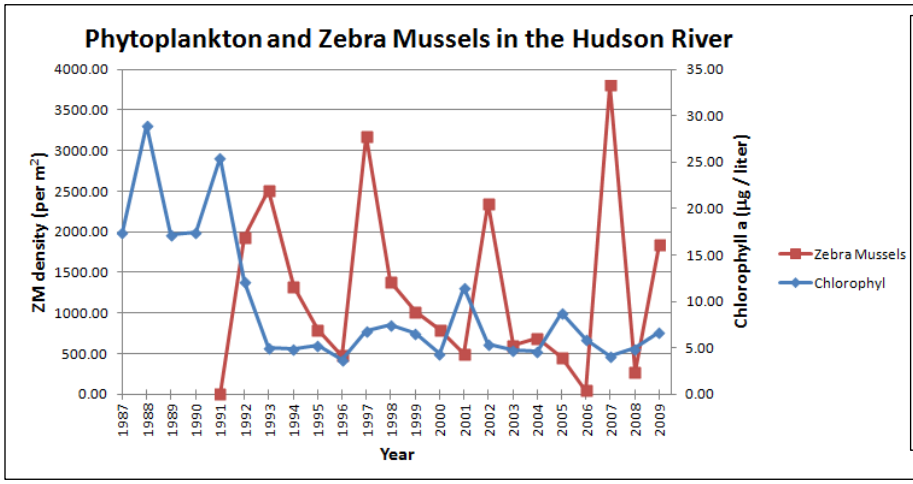


Figure 2: Chlorophyll levels measure the number of phytoplankton, photosynthetic organisms that are the base of the aquatic food chain.

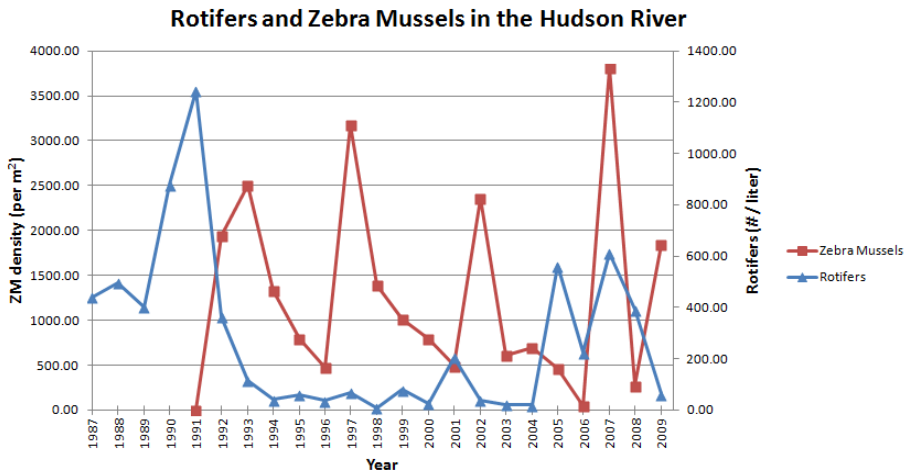


Figure 3: Rotifers are microscopic zooplankton that feed on phytoplankton. In turn, rotifers become food for invertebrates and fish.

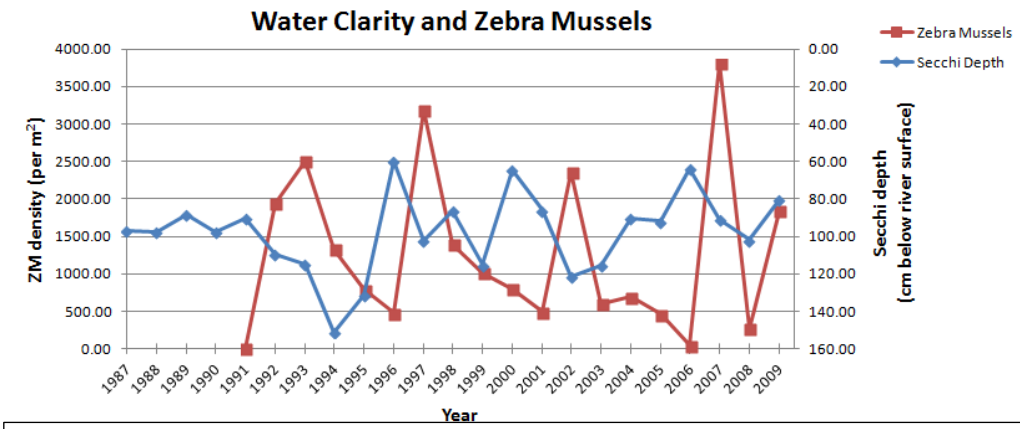


Figure 4: In the graph above, 'Secchi depth' is a measure of how deep one can see an object in the water. Since the depth was measured as cm down from the river surface, the values are plotted from the top of the graph. This means that higher values are at the bottom of the graph (i.e. you can see the object when it's farther down).

A. Water clarity decreased as Zebra mussel populations increase.

B. As Zebra mussels increased, phytoplankton decrease.

C. Both native and Zebra mussels have highly variable population trends.

D. There is a positive correlation between Zebra mussels and rotifers.

_____8. Which year had the clearest water in the Hudson River?

A. 1996

B. 1994

C. 1991 and 2006

D. 2007

_____9. A water sample taken in 2009 would have which of the following:

A. 650 Rotifers/liter and 100 Zebra mussels/m2

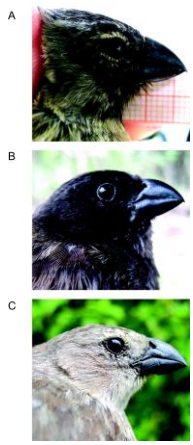
B. 1800/m^2 Zebra mussels and 7 µg/liter of chlorophyll

C. 4.5/m^2 of Zebra mussels and 2/m^2 Unionid native mussels

D. 1800/m^2 Zebra mussels and 4.5 m^2 of Unionid native mussels

For Questions 10-12, read the following passage and study the graphs.

Darwin's ground finches are small birds that live on the Galapagos Islands, 600 miles off the coast of Ecuador. They have been frequent study organisms for evolutionary biologists because they live in an isolated environment and cannot breed with mainland species. The following graphs are from a study done by biologists studying the hybridization of several species of finches on the Galapagos. The following is an excerpt from the abstract of this study:



This study addresses the causes and evolutionary consequences of introgressive hybridization in the sympatric species of Darwin's ground finches (*Geospiza*) on the small island of Daphne Major in the Galápagos archipelago. Introgressive hydridization is the change in the gene pool when hybrid organisms mate frequently with members of one of the parent species. Hybridization occurs rarely (less than 2% of breeding pairs) but persistently across years, usually as a result of imprinting on the song of another species. Hybrids survive well under some ecological conditions, but not others. Hybrids mate according to song type. The resulting introgression increases phenotypic and genetic variation in the backcrossed populations. By Peter and Rosemary Grant, published 2008. All graphs from: http://rstb.royalsocietypublishing.org/content/363/1505/2821

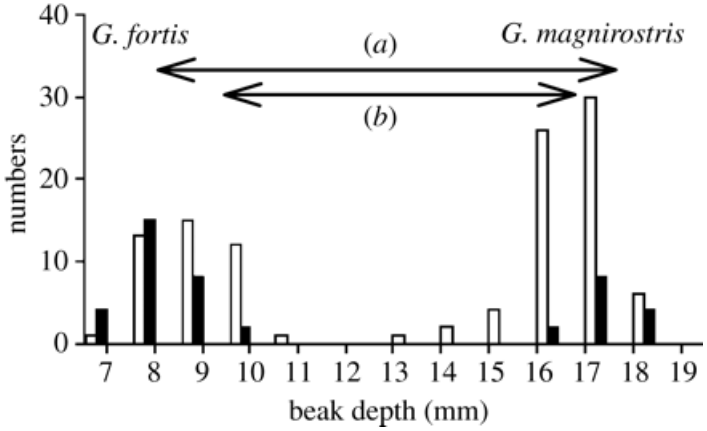


Figure 1. Character displacement (changes in genetic traits) on Daphne Major Island as a result of differential survival during the drought of 2003–2004. White bars are non-survivors and black bars are survivors. (a) = after the drought; (b) = before the drought.

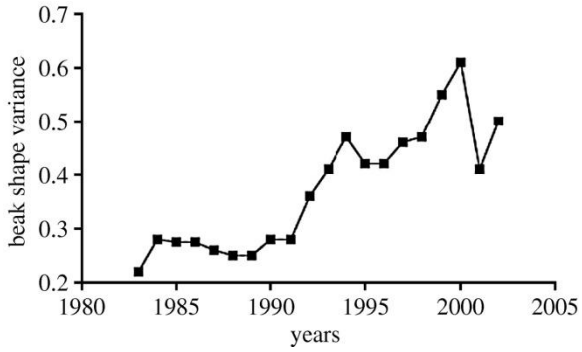


Figure 2. Increase in beak shape variance in *G. scandens* on Daphne Major Island as a result of introgressive hybridization with *G. fortis*. Adapted from Grant *et al*. (2004).

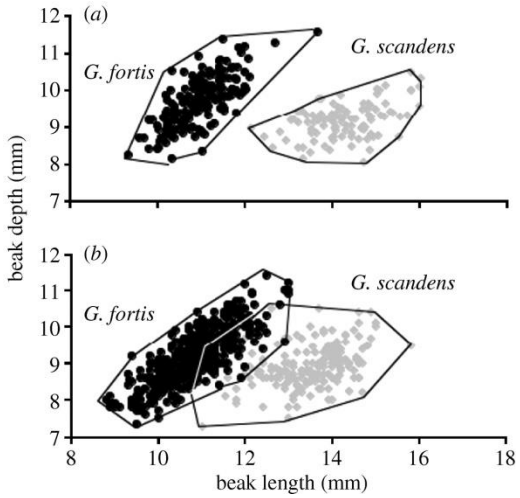


Figure 3. Morphological convergence (when physical traits become more alike) on Daphne Major Island as a result of introgressive hybridization. Polygons enclose members of each population identified by song (males) or the song of mates (females). (*a*) 1978–1982 and (*b*) 1990–2003.

_____10. According to the graphs above, which species exhibited changes in beak depth following the drought of 2003-2004?

A. *G. magnirostrus*

B. *G. fortis*

C. *G. scandens*

D. Both *G. magnirostrus* and *G. fortis*

_____11. *G. scandens* shows changes in beak shape primarily due to what factor?

A. Interbreeding and hybridization with *G. magnirostrus*

B. Changes in song patterns of males and females

C. Introgressive hybridization with *G. fortis*

D. Selective pressure from lack of food due to drought conditions

_____12. Which statement is true about the finches on Daphne Island?

A. *G. fortis* and *G. splendens* showed introgressive hybridization by the year 1982.

B. Between 1995 and 2003, *G. scandens* showed a decline in beak shape variance.

C. There were more *G. magnirostrus* survivors after the drought than *G. fortis* survivors.

D. *G. magnirostrus* and *G. fortis* differed more in mean beak depth after the drought than before the drought.

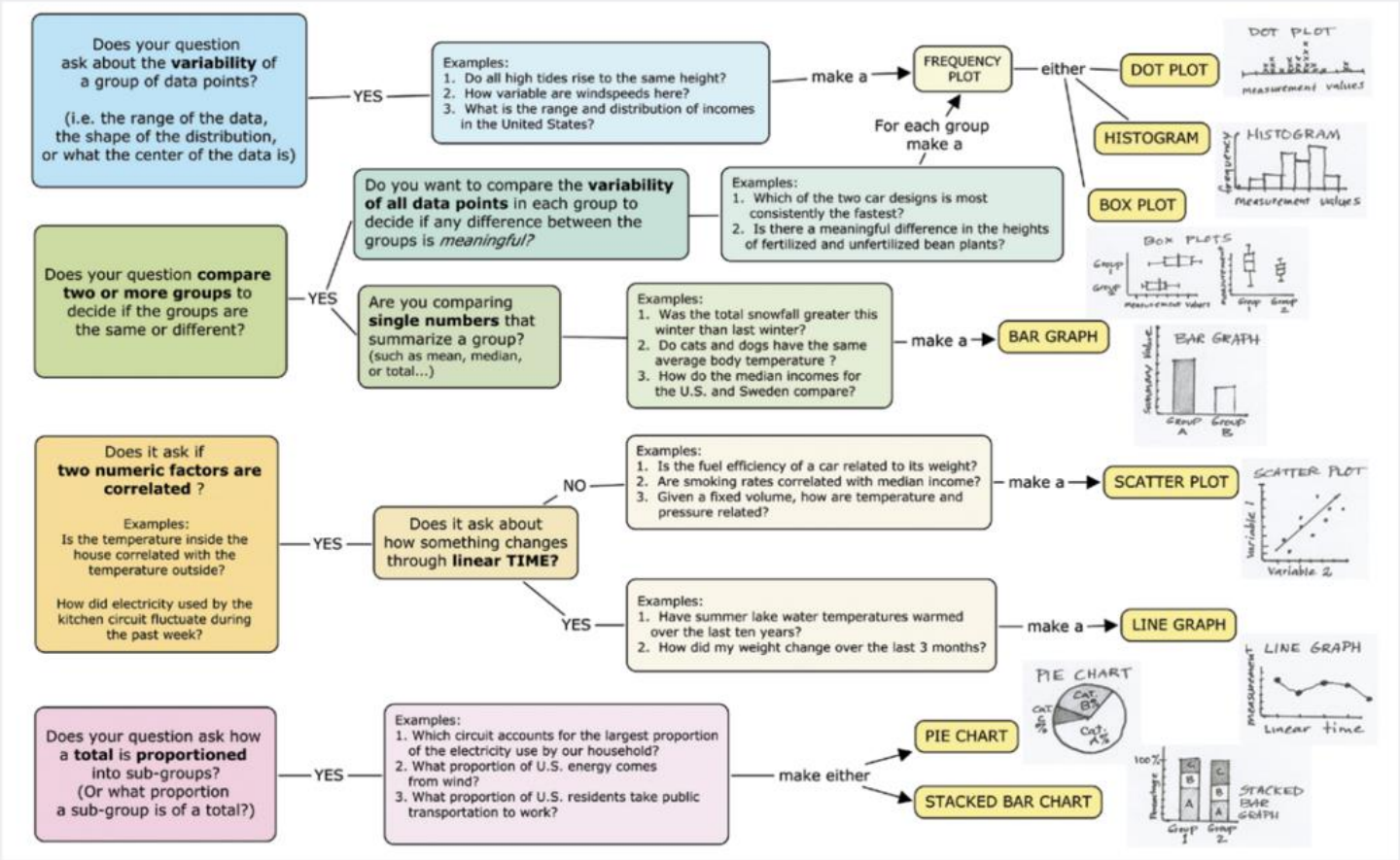
APPENDIX F

GRAPH CHOICE CHART

FIGURE 2

The Graph Choice Chart.

What question would you like to explore? Write your question as a complete sentence.



APPENDIX G

GRAPH CHOICE QUIZ

Using your Graph Choice Chart, choose which kind of graph you would use to graph the following types of data:

1. How have ocean temperatures changed in the last 40 years? _____

2. What is the most popular music genre among high school students? _____

3. How does the amount of fertilizer affect how tall tomato plants grow? _____

4. What percentages of people develop diabetes at different ages? _____

5. Look at the following table: What kind of graph should you make? _____

Summer Camp Attendance Last Week				
Ages	0–5	6–11	12–17	18–23
Frequency	15	25	17	19

6. Look at the following table: What kind of graph should you make?_____

Proportion of each grade attending the Valentine's Day Dance	
Freshmen	110
Sophomores	85
Juniors	70
Seniors	65

APPENDIX H

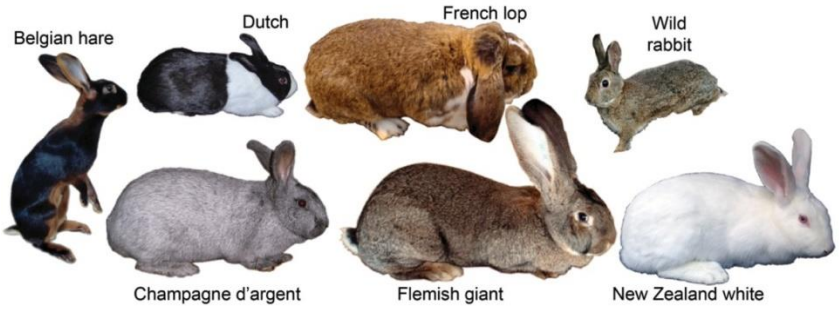
VARIATION IN NATURAL POPULATIONS LAB

VARIATIONS IN NATURAL POPULATIONS LAB

BACKGROUND INFORMATION:

Every living thing—plant or animal, microbe or human being—has a set of characteristics inherited from its parent or parents. Since the beginning of recorded history, people have wanted to understand how that inheritance is passed from generation to generation. Within the last 75 years, we have been unlocking the "secret" to inheritance by learning more about DNA and genes. Scientists have also begun to appreciate that heredity holds the key to understanding what makes each species unique.

A species is a group of organisms that can interbreed and produce viable, fertile offspring. In other words, organisms that can produce living offspring that are healthy and can live long enough to produce their own living, healthy offspring belong to the same species. While this seems like an easy definition, in the real world, it is sometimes difficult to tell whether organisms that are observed belong to the same species or to different, but closely related species. Why is this so? It is because in the real world, organisms within populations often have variations in their traits.



Are all of these rabbits the same species?

If you found them in the wild, would you know?

A population of organisms almost never consists of individuals that are all exactly alike. Even though you can see that all of the organisms shown above are rabbits, what traits would you use to describe a rabbit to someone who had never seen one? While it seems like this would be a simple thing to do, coming up with an "average" rabbit might prove to be difficult to define.

Despite our tendency to describe organisms by choosing the average traits, for some traits there remains considerable variation among individuals in a population and this variation is biologically very important. For example, the coats of some tigers are more striped than others and some leopards have more spots.

Charles Darwin was the first to understand the importance of such variation in organisms. He was the first to recognize that the *variation* in traits within a population of organisms is one of the most important things that biologists should be observing. Prior to Charles Darwin's work, biologists spent their time trying to catalogue, describe and name all of the different species in the world. For decades, naturalists (what biologists used to be called) would go out in the wild, observing, drawing and collecting examples of different species. They spent most of their time trying to locate the "best" example of the species to bring back for their collections. They would disregard those individuals they thought weren't perfect representations of the species. Darwin, however, observed those variations in traits and realized that these differences are the beginnings for variations in species.

For this lab, you will be observing variations in traits of different species. These include pinto beans, isopods and trichomes in desert plants. In addition to recognizing the extent of variations, you will be determining the standard deviation for those traits within the population and graphing your data using a box-and-whisker plot.

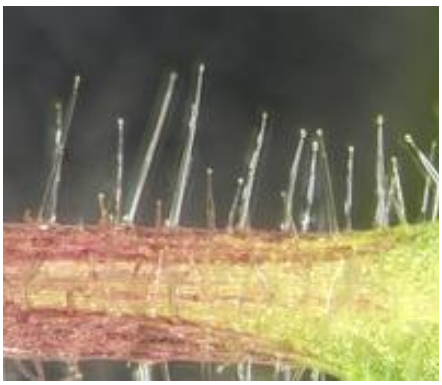
PROCEDURE:

STEP 1: CHOOSE ONE OF THE FOLLOWING FOR YOUR EXPERIMENT

A. VARIATIONS IN LENGTH AND SPOTTING PATTERNS IN PINTO BEANS. Count out 20 pinto bean seeds. With a ruler, measure the length of each bean in millimeters. Count the number of dark spots on each bean. Record data on DATA TABLE 1.

B. VARIATIONS IN MUSSELS SHELL SIZE. Collect 20 mussel shells from stock collection. With a ruler, measure the length of each shell in millimeters. Count the number of ridges on each mussel shell. Record data on DATA TABLE 2.

C. VARIATIONS IN TRICHOME DENSITY. NOTE: Do not touch the plants with your hands during this experiment. Carefully pick up 5 plants of Species A using tweezers. Take the plants back to your lab desk on paper towels. Carefully cut off 4 leaves from the bottom of each of the stems. Observe each leaf under your dissecting microscope. Count the number of trichomes on the petiole of each leaf. (The petiole is the "stem" of the leaf). Repeat this procedure with Species B. Record your data on DATA TABLE 3.



Trichomes are hair-like structures found on the leaves and stems of many plants. Trichomes exist in a wide variety of forms. (Cotton fibers are actually trichomes from the ovary of a cotton plant).

Trichomes have a wide variety of functions, depending on the plant species. Trichomes may be used to protect the plant from herbivory, to reduce water loss by transpiration, increase survival from freezing temperatures, or to protect the plant from UV light.

PART 2: BASIC STATISTICS

1. Determine the mean, median, mode and range for each of your data sets.

2. For each set of data you have recorded, you will be preparing a histogram. For each data set, look at the numbers you have recorded in your data table and decide how to group the data into at least 5 "sets" or groups of numbers. You will then record on DATA TABLE 4 the "sets" of numbers and the number of individuals that fit into each of these "sets" of numbers. Using these data, you will prepare a histogram for each of your data sets

3. Prepare box and whisker plots to show the variation in the data for each of your data sets.

DATA TABLE 1: PINTO BEANS—For each bean, record length in millimeters and number of spots.

Bean:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Length, mm																				
Number of spots																				

DATA TABLE 2: MUSSELS—For each mussel shell, record length in millimeters and number of ridges

Shell:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Length, mm																				
Number of ridges																				

DATA TABLE 3: TRICHOMES—For each species, record number of trichomes on the petiole of each leaf

sample:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Number of trichomes, Species A																				
Number of trichomes, Species B																				

DATA TABLES FOR HISTOGRAMS—DESIGN YOUR OWN DATA TABLES HERE THAT WILL SHOW YOUR "SETS" OF NUMBERS AND HOW MANY INDIVIDUALS WILL FIT INTO EACH "SET" OF NUMBERS):

Data Set 1:

_____Mean_____Median_____Mode_____Range_____

"Sets" of Numbers:	Number of Individuals in that "set" of numbers:

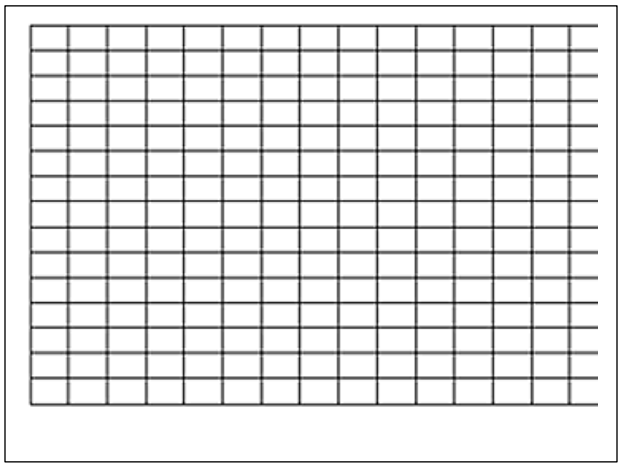
Data Set 2:

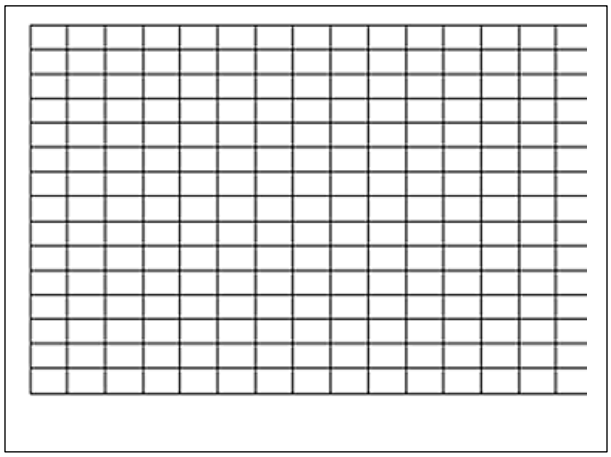
_____Mean_____Median_____Mode_____Range_____

"Sets" of Numbers:	Number of Individuals in that "set" of numbers

HISTOGRAMS:

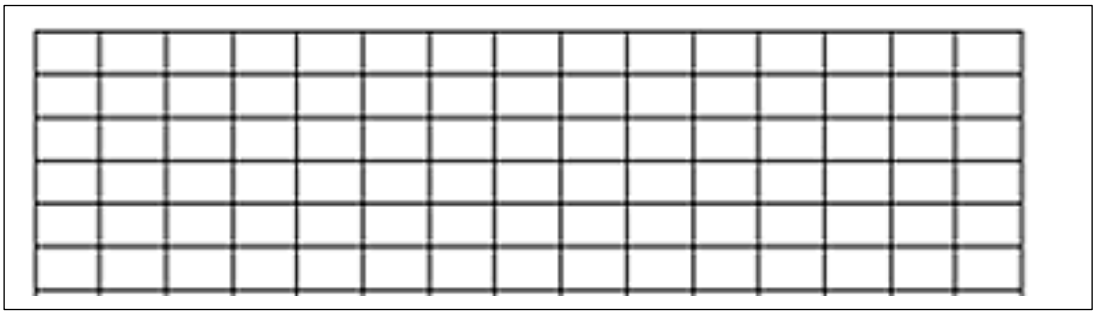
TITLE:_____ TITLE:_____



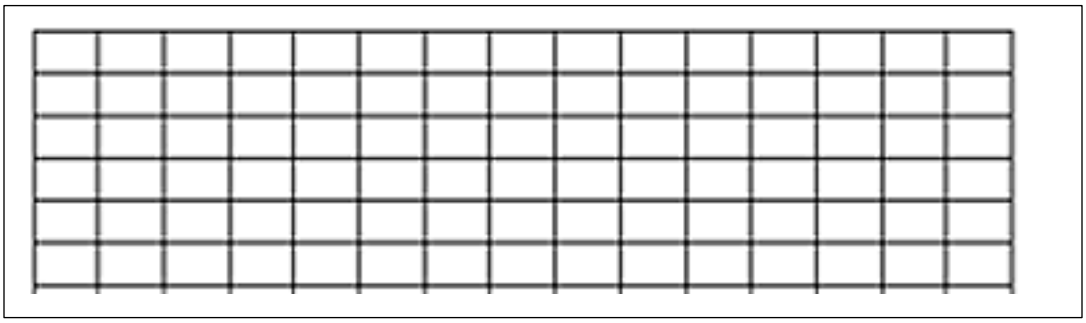


BOX-AND-WHISKER PLOTS:

TITLE:_____



TITLE: _____



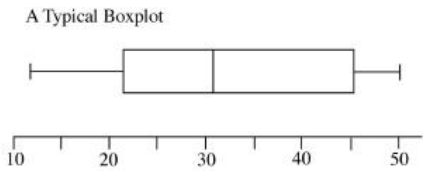
ANALYSIS:

1. How did the traits you measured vary between individuals?

2. How could this variation in traits allow for differential survival rates in these individuals? In other words, how could the differences in traits allow some individuals to have a better chance of surviving in their natural world?

BOX-AND-WHISKER PLOTS

Scientists often need a clear way to show how the variations in their data set are distributed. They may choose to show this with a box-and-whisker plot. Often simply called a Box Plot, this type of graph is used to show the distribution of data by dividing the data into four groups with the same number of data points in each group. The box contains the middle 50% of the data points and each of the two whiskers contains 25% of the data points.



Each one of the vertical lines represents an important number related to the data set:

A. The first and last lines (leftmost and rightmost) are drawn at the lowest and highest data values.
B. The three lines that form the box are drawn 25%, 50%, and 75% of the way through the data.
C. These five numbers, the least, the 25%, the 50%, the 75%, and the greatest, together are called a *5 number summary*.
HOW TO MAKE A BOX-AND-WHISKER PLOT:

1. First, put all of your data points on a line in order from lowest to highest values.

2. Then, cut the list into 4 equal parts. The numbers that are at these locations are called Q1 (Quartile 1), Q2, Q3. NOTE: If the quartile number is in between 2 numbers, take the average of these 2 numbers.

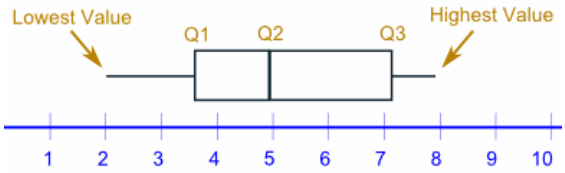
Example: Data points are: 5, 8, 4, 4, 6, 3, 8

Put in order from lowest to highest: 3, 4, 4, 5, 6, 8, 8

Cut the list into quarters: 3, 4, 4, 5, 6, 8, 8
Q1 Q2 Q3

Results: Q1 = 4; Q2= 5 (also the median); Q3 = 8 . The Interquartile Range (IQR) is between Q1 and Q3. The **IQR** is **4** for this set of data.

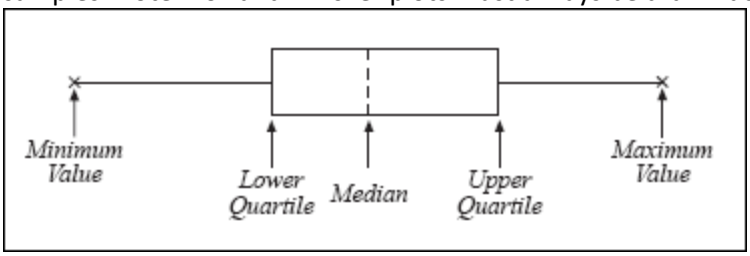
3. Put all of your values into a box and whisker plot as shown below:


Lowest Value
Q1
Q2
Q3
Highest Value
1
2
3
4
5
6
7
8
9
10

APPENDIX I

BOX-AND-WHISKER PLOT QUIZ

A box and whisker plot is based on the *minimum and maximum values*, the *upper and lower quartiles* and the *median*. This type of plot provides a good way to compare two or more samples. Note: Box and whisker plots must always be drawn accurately to scale

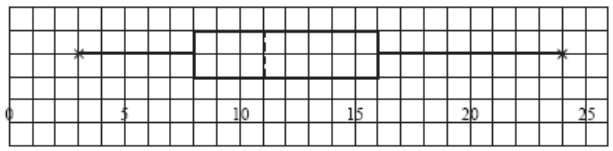


1. A sample has the following data:

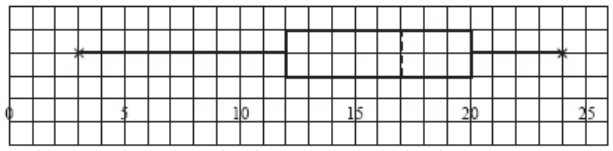
Minimum	3
Range	21
Median	17
Upper quartile	20

Which is the appropriate box-and-whisker plot?

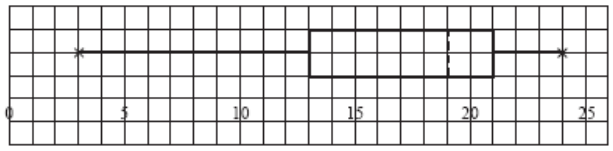
A



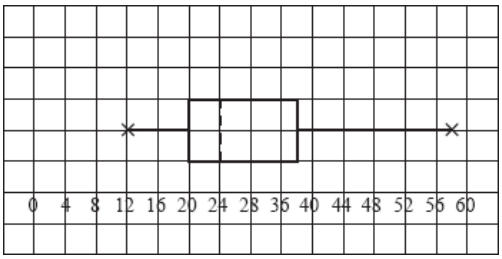
B



C



2. For the following box-and-whisker plot, determine the following data:



2a. Range = _____

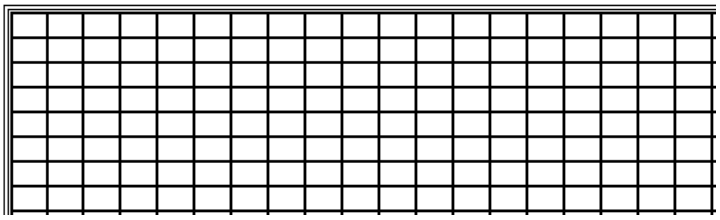
2b. Median = _____

2c. Lower Quartile=_____

2d.Interquartile range_____

3. Draw a box-and-whisker plot of the following data:

5 7 1 9 11 22 15



APPENDIX J

PROTIST DIVERSITY LAB

DIVERSITY OF PROTISTS IN FRESHWATER PONDS

BACKGROUND INFORMATION:

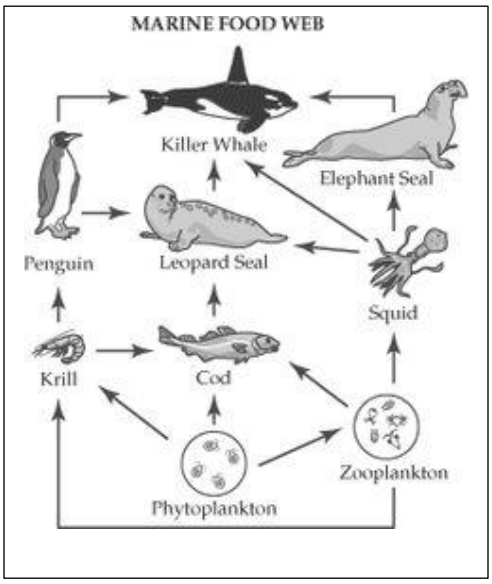
Protists are organisms that belong in the Kingdom *Protista*. This is the most diverse kingdom of life. These organisms are all eukaryotic and almost all are single-celled. They cannot be classified as plants, animals, or fungi, so they are all "lumped" together in one kingdom: *Protista*. Their classification is more based on what they ARE NOT, rather than descriptions of what they ARE.

Because most protists are single-celled, almost all live in aquatic habitats. Species include *Amoeba, Euglena, Paramecium*, algae, slime molds, and thousands more. A few species, such as *Plasmodium* (which causes Malaria) are parasites and can cause diseases in other species. However, most protists are essential to ecosystems. In fact, protists form most of the base of the food chain in all aquatic habitats because many species of photosynthetic protists are the producers in these ecosystems.

Protists occupy different niches in aquatic ecosystems. Some species, such as algae are photosynthetic and are the producers of aquatic ecosystems.

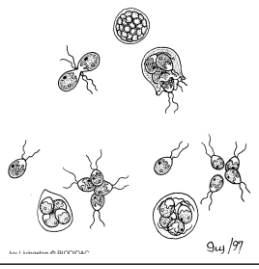
Other species of protists are heterotrophs and feed on bacteria or algae. These, in turn, become food for other animals in the ecosystem.

Together, these organisms may be called phytoplankton or zooplankton. Study the diagram at right to see examples of both

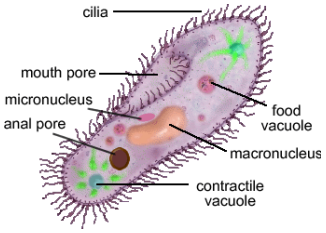


Protists can be divided into three main categories: Plant-like (mostly algae species); Animal-like; and Slime Molds. Within each category, there are thousands of different species. The three general categories are shown by these representative organisms:

PLANT-LIKE PROTISTS



ANIMAL-LIKE PROTISTS



SLIME MOLDS

DETERMINING THE BIOLOGICAL DIVERSITY OF PROTISTS IN TWO POND ECOSYSTEMS

INTRODUCTION

Today you will have the opportunity to observe cells which are actually individual organisms-- unicellular or colonized Eukaryotes. These organisms, which are members of the kingdom *Protista*, come in many shapes and sizes and exhibit diverse behaviors. As a student scientist, you will practice your observational and microscope skills as you investigate the activities of these creatures. The 17th-century observation skill that Anton van Leeuwenhoek is most known for was his identification of these tiny creatures. Leeuwenhoek called these organisms "animalcules" by observing them using a simple compound light microscope that he built himself. This lab should give you a sense of what he witnessed in his studies.

Biodiversity of species is an important concept in the study of ecosystems. The more diverse the organisms are in an ecosystem, the more stable it is. If something in the environment changes, an ecosystem with more diversity is more likely to recover quickly.

For today's lab, you will be determining the biodiversity index of protists in two different ponds. A pond can be considered its own ecosystem in that there are a number of different species living and interacting in the pond with the water, oxygen, sunlight and nutrients available. A biodiversity index is determined by the following formula:

$$\frac{\text{Number of different species}}{\text{Total number of organisms}} = \text{BIODIVERSITY INDEX}$$

In addition to determining the biodiversity index of these two ecosystems, you will learn how to use a *t*-test to determine if the results you get from your data are statistically valid. A *t*-test is a statistical test used to compare two means (averages of two sets of data) to see if they are statistically different from each other.

STUDENT OBJECTIVES

1. Prepare and observe wet-mount slides of mixed protists.
2. Calculate the biodiversity index of protists in two different pond ecosystems.
3. Use a *t*-test to determine if results from data are statistically valid.

PROCEDURE

1. Choose which two pond water samples you will be observing. Write a hypothesis here for which pond you think the Biodiversity Index will be higher and why:

   _____
   _____
   _____
   _____

2. Collect a sample of water from the bottom of the container of your first pond sample (Pond 1). Put 1 drop onto a microscope slide and place a cover slip on top. (CAUTION! COVER SLIPS ARE VERY SHARP—HANDLE WITH CARE!).

3. Observe the sample under the microscope at low power. Look for organisms that are moving, then focus on that area. Move the objective to medium power.
4. When you see organisms at medium power, count the TOTAL NUMBER OF ORGANISMS you see and record on your data sheet.
5. Then, decide how many DIFFERENT SPECIES of protists there are in this sample and record on your data sheet.
6. Wash off the slide and collect another sample.
7. Repeat Steps 2-6 until you have observed 5 samples from Pond 1.
8. Repeat Steps 2-6 until you have observed 5 samples from Pond 2.

CALCULATIONS:

9. Calculate the MEAN for number of total organisms in each pond. Calculate the MEAN for total species in each pond.
10. Calculate the Biodiversity Index for each pond. Use the formula from the Background Information section. Use the mean numbers you calculated in step 9.
11. Follow the instructions to perform the *t*-test to determine if your samples are statistically different from each other.
12. Answer the Analysis Questions with complete sentences.

ANALYSIS:

1. **In which location did the pond organisms have the highest Biodiversity Index? NOTE: If there was no statistical difference between the two locations, indicate that here as well:**_____

2. **Why do you think that location had the most diversity of organisms? If there was no statistical difference between the two locations, write an explanation of why you think that was the case**_____

3. **What adaptations did the organisms you observed today have that allowed them to be perfectly suited for this environment? Describe at least 3 adaptations:** _____

DRAW 3 OF THE ORGANISMS YOU OBSERVED HERE:

DATA TABLE

POND 1 LOCATION: _____

SAMPLE NUMBER	TOTAL NUMBER OF ORGANISMS SEEN IN SAMPLE:	TOTAL NUMBER OF SPECIES SEEN IN SAMPLE:
1		
2		
3		
4		
5		
MEAN (AVERAGE NUMBER):		

POND 2 LOCATION: _____

SAMPLE NUMBER	TOTAL NUMBER OF ORGANISMS SEEN IN SAMPLE:	TOTAL NUMBER OF SPECIES SEEN IN SAMPLE:
1		
2		
3		
4		
5		
MEAN (AVERAGE NUMBER):		

BIODIVERSITY INDEX FOR EACH POND: MEAN number of species/ MEAN total number oorganisms

POND 1: LOCATION_____ _	POND 2: LOCATION_____ _
BIODIVERSITY INDEX CALCULATION:	

***t*-TABLE CALCULATION RESULTS:**

SIGNIFICANT _____NOT SIGNIFICANT_____

CALCULATING A *t*-TEST TO ANALYZE DATA SETS

BACKGROUND INFORMATION: From: The Basics of Data Literacy by M. Bowen and A. Bartley

A *t*-test is a statistical test used to compare two means to see if they are statistically different from each other. What the *t*-test is doing is determining what the likelihood is that the difference between the two means happens because of chance or because of the variable you tested.

In simple terms, the *t*-test compares how much data "scatter" there is for each variable and then compares how different the means are in relation to that "scatter". The *t*-test can determine how likely it is that differences between the two means is due to random chance.

Here are the conditions that must be met for the results of the *t*-test to be valid:

1. **The data scatter is reasonably the same for the two categories (the variation between the data is close to the same).**
2. **More of the data points are in the "middle" of the data set than at the far ends.**
3. **The data points are randomly chosen (you didn't choose which data points to include in your calculations).**
4. **The replicates in the two treatments need to be independent of each other. (In other words, you can't use the *t*-test to measure before and after treatments of the same individuals).**

HOW TO USE THE *t*-TEST: Follow the steps outlined on the next page. You will be determining a *t*-statistics number. Then look at the Critical Values chart shown below. Make sure you look at the correct degrees of freedom, which you will have calculated on the worksheet.

1. **If the *t*-statistic you calculated is *less than* the critical value in the table below, (for the correct degrees of freedom, which you calculate on the worksheet), then the difference between the two means is NOT statistically significant. This means that the two means varied only by random chance.**
2. **If the *t*-statistic you calculated is *greater than* the critical value in the table below (for the correct degrees of freedom), then the difference between the two means IS statistically significant at 5%. This means that we are 95% confident that the difference between the means is a real one and not due to random chance.**

CRITICAL VALUES FOR *t*-TEST:

Degrees of Freedom	a = 0.05
1	12.71
2	4.30
3	3.18
4	2.78
5	2.57
6	2.45
7	2.36
8	2.31
9	2.26
10	2.23

***t*-TEST WORKSHEET**

	GROUP 1: (a) data _____		GROUP 2: (b) data _____	
Data points:	**(a)**	**(a) X (a)**	**(b)**	**(b) X (b)**
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
Sum	**(c)**	**(d)**	**(e)**	**(f)**
Count	**(g)**		**(h)**	

Step 1: Enter data in table in rows (a) and (b).

Step 2: Square (a) and put in column (a) X (a); square (b) and put in column (b) X (b).

Step 3: Sum columns (a), (a) X (a), (b), and (b) X (b) and put results on Sum row.

Step 4: Count measures in column (a) and (b) and enter them on the Count row. (how many data points you have).

Step 5: Calculate: $\frac{(c) \times (c)}{(g)}$ = _____(i) and $\frac{(e) \times (e)}{(h)}$ = _____(j)

Step 6: Calculate: (d) - (i) = _____(k) and (f) – (j) = _____(l)

Step 7: Calculate: (k) + (l) = _____(m)

Step 8: Calculate: $\frac{(m)}{(g) + (h) - 2}$ = _____(n)

Step 9: Calculate: (n) X sum of: $\frac{(1)}{(g)} + \frac{(1)}{(h)}$ = (n) X (_____+ _____) = _____(o)

Step 10: Calculate: √ (square root) of (o) = _____(p)

Step 11: Calculate: $\frac{(c)}{(g)}$ = _____(q) and $\frac{(e)}{(h)}$ = _____(r)

Step 12: Calculate: (q) – (r) = |_____| (s) Note: Absolute value of (s)

Step 13: Calculate: *t*- statistic = $\frac{(s)}{(p)}$ = _____

Step 14: Calculate Degrees of Freedom = (g) + (h) – 2 = _____

Check the appropriate box: Significant _____ Not Significant _____

APPENDIX K

STATISTICAL ANALYSIS QUIZ

<u>Using a *t*-test to analyze data.</u>

A *t*-test is a statistical tool that allows you to determine if the differences you see between two sets of numbers are significant or not. A statistical test allows the scientist to state with more certainty what differences and patterns they have found in their data. A *t*-test is used when you have data with only two test variables you are comparing.

To make sure you can use a *t*-test with your data, it must meet the following conditions:

1. You will be testing two means to see if they are statistically different from each other. So, you need two sets of data points that you have determined the mean for each.

2. The data scatter is reasonably the same for the two categories (the variation from the mean is about the same).

3. There is more data toward the middle (median) than there is at the farthest points away from the middle. (There is a reasonably normal distribution).

4. The data are randomly chosen (this means you didn't choose the data you included).

5. The replicates in the two treatments need to be independent of each other (the data cannot be before and after on the same individuals).

Note: After performing a *t*-test on your data, determine the p-value. The standard to measure it by is 0.05. If the p-value is less than 0.05, then there is no statistical difference between your two samples.

Quiz Questions:

1. A loving dog owner brings their Labrador to the veterinarian's office. When he lumbers onto the scale, Fido weighs 88 pounds. The doctor tells them they must put Fido on a diet. When the owner asks why, the veterinarian explains his dog is very overweight. The veterinarian tells Fido's owner: "Here are the weights for the healthy Labradors I have seen in my office:"

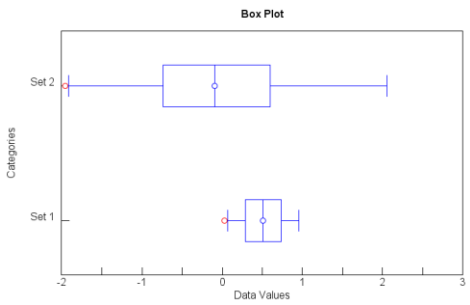
64 79 73 69 71 68 65 76

What is the mean for these numbers? _____(average weight for a Labrador)

How far away from the mean is Fido? _____

Should the loving dog owner put Fido on a diet? _____Why or why not?

2. Will the data shown in the box-and-whisker plot below meet the criteria for a *t*-test? Why or why not?



3. Angela performs an experiment where she tests the growth of plants with two different kinds of light—a grow light and sunlight. She grows her plants for the same amount of time and realizes that both sets of plants have average growth rates that are very similar. To make sure she didn't calculate incorrectly, she runs a *t*-test on her data. She discovers that the p-value is .15. Angela decides that there is no significant difference between the two types of light. Is Angela correct? _____Why or why not? _____

APPENDIX L

ANTARCTIC CLIMATE CHANGE LESSON

ANTARCTIC CLIMATE CHANGE LESSON

Adapted from:

Constible, J., L. Sandro, R.E. Lee, Jr. 2007. A Cooperative Classroom Investigation of Climate Change. *The Science Teacher.*74:6.

Background: Antarctic Climate Change

At the global level, strong evidence suggests that observed changes in Earth's climate are largely due to human activities. At the regional level (such as in Antarctica), the evidence for human-dominated change is somewhat less clear. Scientists have a particularly difficult time explaining warming trends in Antarctica—a region with a relatively short history of scientific observation and a highly variable climate. Regardless of the mechanism, however, climate change is having a dramatic impact on Antarctic ecosystems.

Air temperature data indicate that the western Antarctic Peninsula has warmed by about $3^{o}C$ in the last century. Although this relatively short-term record is only from a few research stations, other indirect lines of evidence confirm the trend. The most striking of these is a shift in penguin communities. Adelie Penguins, which are dependent on sea ice for their survival, are rapidly declining on the Antarctic Peninsula despite a 600-year colonization history. In contrast, Chinstrap Penguines, which prefer open water are increasing dramatically. These shifts appear to be due to changes in the amount, timing, and duration of sea ice.

Why is sea ice so important to Adelie Penguins?

1) Sea ice is a feeding platform for Adelies. Krill, the primary prey of Adelies, feed on microorganisms growing on the underside of the ice. For Adelie penguins, which are relatively slow swimmers, it is easier to find food under the ice than in large stretches of open water.
2) Sea ice helps control the local climate. Ice keeps the Peninsula cool by reflecting solar radiation back to space.
3) Sea ice acts as a giant cap on the ocean, limiting evaporation. As sea ice declines, cloud condensation nuclei and moisture are released into the atmosphere, leading to more snow. This extra snow often does not melt until Adelies have already started nesting; the resulting melt water can kill their eggs.

ANTARCTIC CLIMATE CHANGE LESSON

Day 1: Library Research and Specialist Group Work

Students will be assigned to a Home Group of 5 students. Within their home group, each person will choose from 1 of 5 specialists' roles: Ornithologist, Oceanographer, Meteorologist, Marine ecologist, or Fisheries biologist. The class will then go to the library.

Each person will do the following research in the library:

1. Go to www.units.muohio.edu/cryolab/education/AntarcticLinks.htm
2. Click on the General Information, Antarctica: Wikepedia article.
3. Scroll down the Wikipedia article to the Flora and Fauna section. Answer the following questions on a separate piece of paper:
    a. List 7 invertebrate animals that live in Antarctica
    b. What kinds of penguins live in Antarctica? Which is the most southern species of penguin?
    c. What kinds of seals live in Antarctica? Which one was hunted extensively?
    d. What is the keystone species for Antarctica? What animals feed on it?
    e. Why was it necessary to pass legislation to regulate over-fishing in the Antarctic?

4. Click on this website: http://lrs.ed.uiuc.edu/students/downey/project/penguins.html
5. On the same paper you completed with number 3 above, answer the following questions:
    a. Name 3 adaptations penguins have to protect them from the cold.
    b. What is the typical food of penguins?
    c. What are the primary predators of penguins in Antarctica (both land and sea predators)?
    d. Draw a table as follows and fill in the information on the two species:

Species	Average size	What habitat in Antarctica do they prefer?	Describe their mating/family life
Adelie Penguins			
Chinstrap Penguins			

Scientists' Roles:

Each Home Group contains 5 different specialists:

Ornithologist: A scientist who studies birds. Uses visual surveys(from ship or land), diet analysis, and satellite tracking to collect data on penguins.

Oceanographer: A scientist who studies the ocean. Uses satellite imagery, underwater sensors, and manual measurements of sea ice thickness to collect data on sea ice conditions and ocean temperature.

Meteorologist: A scientist who studies the weather. Uses automatic weather stations and visual observations of the skies to collect data on precipitation, temperature, and cloud cover.

Marine ecologist: A scientist who studies the relationship between organisms and their ocean environment. Uses visual surveys, diet analysis, and satellite tracking to collect data on a variety of organisms, including penguins.

Fisheries biologist: A scientist who studies fish and their prey. Collects data on krill during research vessel cruises.

After doing the above individual research in the library, students will then work with the other students in the class who share the same specialty as they do. (Ex: all Ornithologists will work together, all Oceanographers, etc.) All students will meet with their own Specialist Group and do the following:

Together with your Specialist Group, read the information on your dataset. Each member of the Specialist Group will graph the data so they can take the information back to their Home Group. After everyone has graphed the data, discuss the trends in the data and interpret the graph together.

Day 2: Building a Timeline

1. The following day, in class, students will work with their Home Group. Each member of the Home Group will make a brief presentation to the group.

2. The group's job will be to take the information each person brings to the discussion and prepare a <u>flowchart</u> of events. The groups will be given a set of cards with which to prepare the flowchart. Throughout this process, you must use the weight of evidence to construct the flowcharts. In other words, each idea should be accepted or rejected based on the amount of support it has.

3. Students will turn in their graphs and the completed flowchart made by the home group.

Day 3: Class Discussion (or Homework):

1. Discuss the following questions as a class: How has the Antarctic Peninsula ecosystem changed in the last 50 years? What are the most likely explanations for these changes? Is there sufficient evidence to support these explanations? Why or why not?

2. What information from your Specialist dataset did you find most surprising? What information was the most important in building your flowchart? What information was lacking—what other questions would you like to find out about?

3. Why did we need to consult scientists from many different fields of expertise to come to a conclusion? Which dataset was most important in making your decision about the flowchart?

Extension Activity:

For homework, look up the references provided in the bibliography for your dataset. Find data for the years 2004 to the present. Include it in your graph. How has the data changed in the last 6 years?

Ornithologists: Adélie penguin dataset

Data Source: Smith, Fraser, and Stammerjohn, 2003.

• Adélie penguins spend their summers on land, where they breed. They spend winters on the outer extent of the sea ice surrounding Antarctica, where they molt their feathers and fatten up.

• Adélies are visual predators, meaning they need enough light to see their prey. Near the outer part of the pack ice, there are only a few hours of daylight in the middle of the winter. There is less sunlight as one moves further south (closer to land).

•On the western Antarctic Peninsula, Adélie penguins mostly eat krill, a shrimplike crustacean.Several countries have been harvesting krill since the mid 1960s.

• Adélie penguins need dry, snow-free places to lay their eggs. They use the same nest sites each year and at about the same time every year. Heavy snowfalls during the nesting season can bury adult Adélies and kill their eggs.

•Female Adélies lay two eggs, but usually only one of those eggs result in a fledged chick (fledged chicks have a good chance of maturing into adults). The two most common causes of death of eggs and chick are abandonment by the parents (if they cannot find enough food) and predation by skuas (hawklike birds).

•In the water, Adélies are eaten mostly by leopard seals and killer whales.

•Adélies can look for food under sea ice because they can hold their breath for a long time. They are not as good at foraging in the open ocean, because they cannot swim very fast.

•Adélie penguins have lived in the western Antarctic Peninsula for at least 644 years.

Year	#Breeding pairs of Adélie penguins
1975	15,202
1979	13,788
1983	13,515
1986	13,180
1987	10,150
1989	12,983
1990	11,554
1991	12,359
1992	12,055
1993	11,964
1994	11,052
1995	11,052
1996	9,228
1997	8,817
1998	8,315
1999	7,707
2000	7,160
2001	6,887
2002	4,059

Oceanographers: Sea ice dataset

Data Source: Palmer LTER Data Archive (http://pal.lternet.edu/data/dataset_catalog.php)

•In August (mid winter), sea ice covers over 18 X 106 km^2, or 40% of the Southern Ocean (an area larger than Europe). In February (mid summer), only 3 X 106 km^2 (about 7%) of the ocean is covered by sea ice.

•Sea ice keeps the air of the Antarctic region cool by reflecting most of the solar radiation back into space.

•Open water absorbs solar radiation instead of reflecting it and converts it to heat. This heat warms up the atmosphere.

•Sea ice reduces evaporation of the ocean, thus reducing the amount of moisture that is released to the atmosphere.

•As sea ice melts, bacteria and other particles are released into the atmosphere. These particles form condensation or freezing nuclei, which grow into rain or snow.

•Rain helps to stabilize the sea ice by freezing on the surface.

•Sea ice can be broken up by strong winds that last a week or more.

•An icebreaker is a ship used to break up ice and keep channels open for navigation. Icebreakers were first used in the Antarctic in 1947.

Year	Area of sea ice extending from the Antarctic Peninsula (km^2)
1980	140,298
1981	136,511
1982	118,676
1983	88,229
1984	85,686
1985	78,792
1986	118,333
1987	142,480
1988	90,310
1989	44,082
1990	79,391
1991	111,959
1992	110,471
1993	94,374
1994	103,485
1995	95,544
1996	86,398
1997	100,784
1998	73,598
1999	79,223
2000	79,200
2001	69,914

Meteorologists: Winter snow dataset

Data source: Antarctic Meteorology Online, British Antarctic Survey (www.antarctica.ac.uk/met/metlog/).

•In the winter, most of the precipitation in the western Antarctic Peninsula occurs as snow. There is an even mix of snow and rain the rest of the year.

•It is difficult to accurately measure the amount of snowfall in the Antarctic because strong winds blow the snow around.

•The Antarctic Peninsula has a relatively warm maritime climate, so gets more rain and snow than the rest of the Antarctic continent.

•Most of the rain and snow on the Peninsula is generated by cyclones from outside the Southern Ocean. Cyclones are areas of low atmospheric pressure and rotating winds.

•When there is less sea ice covering the ocean, there is more evaporation of the ocean and therefore more moisture in the atmosphere.

•As sea ice melts, bacteria and other particles are released into the atmosphere. These particles form condensation or freezing nuclei, which grow into rain or snow.

Year	% of precipitation events that are snow
1982	49
1983	67
1984	72
1985	67
1986	81
1987	80
1988	69
1989	69
1990	68
1991	72
1992	70
1993	70
1994	83
1995	77
1996	74
1997	81
1998	81
2000	71
2001	90
2002	82
2003	76

Marine ecologists: Chinstrap penguin dataset
Data source: Smith, Fraser, and Stammerjohn 2003.

•Chinstrap penguins breed on land in the spring and summer and spend the rest of the year in open water north of the sea ice. The number of chinstraps that successfully breed is much lower in years when the sea ice does not melt until late spring.

•Chinstraps mostly eat krill, a shrimplike crustacean.

•Whalers and sealers overhunted seals and whales, which also eat krill, until the late 1960s.

•Chinstraps primarily hunt in open water, because they cannot hold their breath for very long.

•The main predators of chinstraps are skuas (hawklike birds), leopard seals, and killer whales.

•Chinstraps will aggressively displace Adélie penguins from nest sites in order to start their own nests and may compete with Adélies for feeding areas.

•Although chinstrap penguins have occupied the western Antarctic Peninsula for over 600 years, they have become more numerous near Palmer Station only in the last 35 years.

Year	# of breeding pairs of Chinstrap penguins
1976	10
1977	42
1983	100
1984	109
1985	150
1989	205
1990	223
1991	164
1992	180
1993	216
1994	205
1995	255
1996	234
1997	250
1998	186
1999	220
2000	325
2001	325
2002	250

Fisheries biologists: Krill dataset

Data source: Atkinson et al. 2004

•Krill are keystone species, meaning they are one of the most important links in the Antarctic food web. All the vertebrate animals in the Antarctic either eat krill or another animal that eats krill.

•Krill eat mostly algae. In the winter, the only place algae can grow is on the underside of sea ice.

•Several countries have been harvesting krill since the mid 1960s.

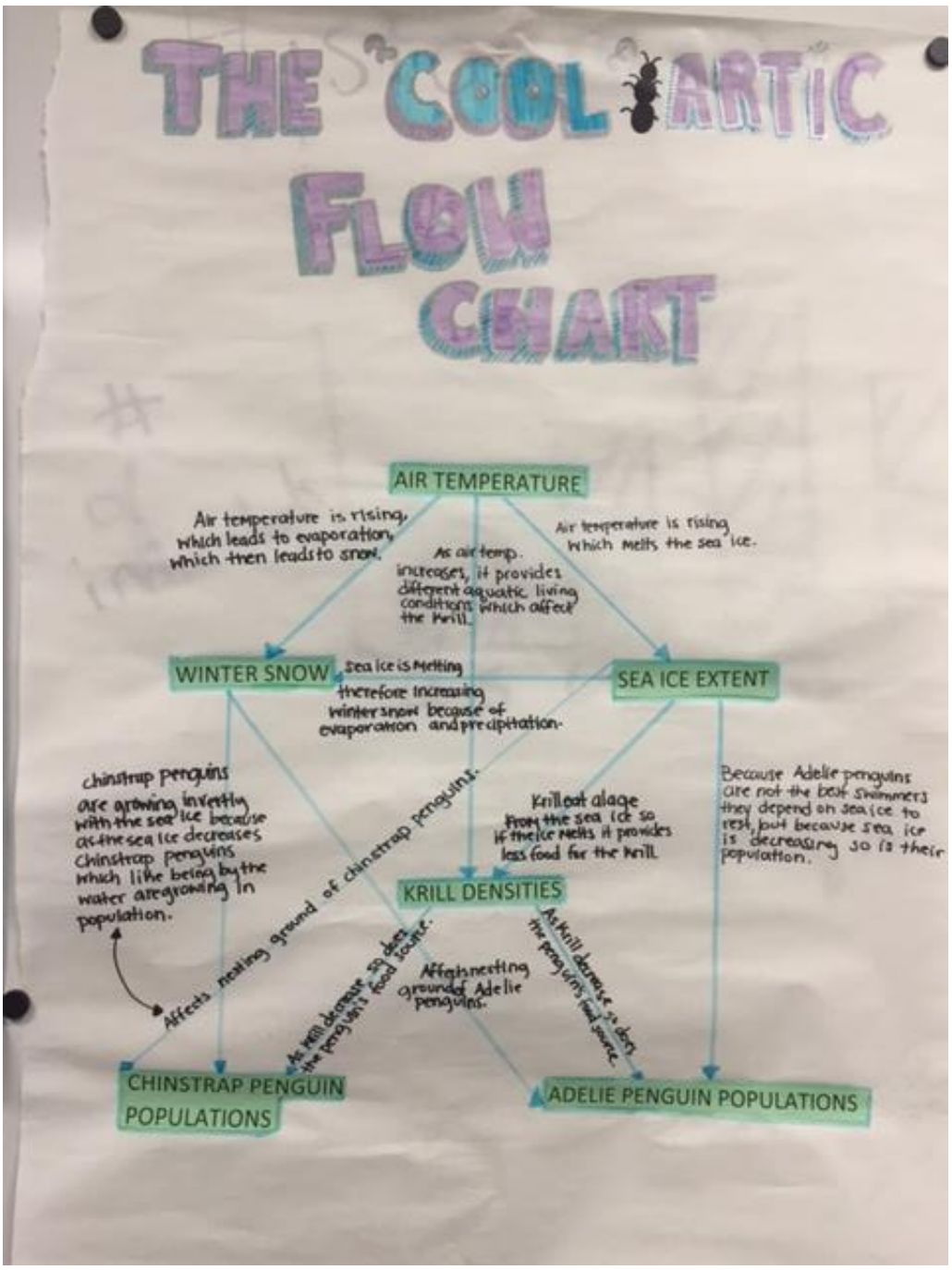
•Ultraviolet radiation is harmful to krill, and can even kill them. Worldwide, ozone depletion is highest over Antarctica.

•Salps, which are small, marine animals that look like blobs of jelly, may compete with krill for food resources. As the salt content of the ocean decreases, salp populations increase and krill populations decrease.

Year	Density of krill in the Southern Ocean ($no./m^2$)
1982	91
1984	50
1985	41
1987	36
1988	57
1989	15
1990	8
1992	7
1993	22
1994	6
1995	9
1996	31
1997	53
1998	46
1999	4
2000	8
2001	81
2002	8
2003	3

APPENDIX M

STUDENT SAMPLE OF TIMELINES


THE COOL ARTIC
FLOW CHART
AIR TEMPERATURE
Air temperature is rising, which leads to evaporation, which then leads to snow.
As air temp. increases, it provides different aquatic living conditions which affect the krill.
Air temperature is rising which melts the sea ice.
WINTER SNOW
Sea ice is melting therefore increasing winter snow because of evaporation and precipitation.
SEA ICE EXTENT
chinstrap penguins are growing in rettln with the sea ice because as the sea ice decreases chinstrap penguins which like being by the water are growing in population.
Krill eat algae from the sea ice so if the ice melts it provides less food for the krill.
Because Adelie penguins are not the best swimmers they depend on sea ice to rest, but because sea ice is decreasing so is their population.
KRILL DENSITIES
Affects nesting ground of chinstrap penguins.
As krill decrease so does the penguin's food source.
Affects nesting ground of Adelie penguins.
As krill decrease so does the penguin's food source.
CHINSTRAP PENGUIN POPULATIONS
ADELIE PENGUIN POPULATIONS