MEASURING THE EFFECTIVENESS OF ALTERNATIVE METHODS OF LAB WORK ANALYSIS IN AN AP ENVIRONMENTAL SCIENCE CLASS

by

Thomas Robert Wellnitz

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 2011

STATEMENT OF PERMISSION TO USE

In presenting this professional paper in partial fulfillment of the requirements for a master's degree at Montana State University, I agree that the MSSE Program shall make it available to borrowers under rules of the program.

Thomas Robert Wellnitz

July 2011

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Lab investigations are an important part of the AP Environmental Science course. Given the 25 labs expected by the College Board and the number of students taking the class, this project sought to evaluate the increase in student understanding of environmental lab concepts from alternative debriefing methods used instead of lab reports. Four classes used different methods – use of preview materials before the lab, small group whiteboard presentations about a specific part of the lab experience, class discussion after the lab, and no treatment – to process the same lab experience rather than doing lab reports. After the treatment, students took a quiz over the lab content. Scores on the quiz were compared to determine from which treatment students better understood the associated concepts. The treatment methods rotated between classes so that each class used each of the approaches three times. Students in the class using the whiteboards had the highest class-average seven of the twelve times and were ranked second for another three trials. A possible reason for the effectiveness of the whiteboard is the small group discussion may increase the mental stimulation necessary for learning and memory.

INTRODUCTION AND BACKGROUND

Project Background

-

Teaching Experience and Classroom Environment

For the past 25 years, I instructed students in middle and high school science classes. For the most recent four years, I taught upper level courses at Northview High School, Johns Creek, Georgia. Northview High School is a Fulton County public high school, and the Fulton County School System comprises schools from the communities to the northeast and southwest of Atlanta, Georgia. Northview High School is a high achieving, suburban school. Mean SAT scores for the current graduating class are Reading -- 563, Math-- 604, and Writing -- 561; these are the third highest scores among Georgia public high schools. In the current year, 58% of seniors took at least one AP course. School enrollment for 2010-2011 is 1890. School demographics are 50% white, 40% Asian, and 10% other racial groups; 5% of the school receives free lunch.

My interest in evaluating lab activities began two years ago when I surveyed honors chemistry students about their attitudes regarding completing lab work. A large number (25%) felt that doing labs was the least effective way of learning material. Students didn't feel they learned much from completing lab reports, with 90% saying lab reports were among the least useful ways for them to learn. A summary of this survey from 2009 is included as Appendix A. Meanwhile, I was frustrated by the time needed to grade lab reports and to provide students with feedback about their work. The lab experience was not working well as a learning tool from both the student and teacher perspective.

In my capstone project, I looked at how discussion-based alternatives to lab reports affected student understanding of science and lab concepts. I studied four classes of AP Environmental Science at Northview High School. The study size was 120 students: 88% seniors and 12% juniors; 50% male and 50% female; 33% students of Asian background with the remaining 67% mostly Caucasian. Classes met five times a week, for 55 minutes. This project sought to explore both how students can increase understanding of environmental science concepts covered by laboratory investigations and experimentation, and how I, as the teacher, can spend less time evaluating their work and still provide meaningful comments for improvement. This is important both to the students and me because the College Board expects an AP Environmental Science class to conduct at least 25 lab investigations during the year.

Focus Question

My primary research question is how does the style of pre-lab preparation and post-lab analysis affect student understanding of environmental science lab concepts? Secondary questions are

1. How does individual student understanding increase when a student participates in either a small or large group discussion compared with processing the meaning of a lab experience on his or her own?

2. How well do class discussions and group poster presentations provide students with meaningful feedback on their lab experience?

3. How does the implementation of group and/or class discussion affect the time it takes me to evaluate and to grade student lab work and to manage lab equipment and experiences?

CONCEPTUAL FRAMEWORK

Before starting the action research project, I evaluated literature related to the assessments of science lab activities. Laboratory experiences form an important part of high school environmental science classes. In Georgia, identifying and investigating a problem scientifically and communicating investigation through lab reports are specified by the Georgia Performance Standards (Georgia Department of Education, 2006). The nature or importance of student laboratory experiences in high school science education has changed considerably over the past 100 years. Historically, one aspect of laboratory work was to verify scientific concepts presented in lectures. Another was learning practical skills. The science education revisions that began in the 1980s expanded the focus of student understanding from an emphasis mostly on content to one that incorporated greater development of scientific reasoning and understanding of the nature of science. In the contemporary classroom, a laboratory exercise may serve any of the above purposes. Educators, in practice, may not agree on the purpose of the lab investigation (Singer, Hilton, & Schweingruber, 2005).

Some of the change in lab emphasis originated because traditional lab work does not help students understand the principles of experimental design or data analysis (Eisenkraft, 2007). Many traditional laboratory experiences are designed so that students will work directly with material, and in the process of this hands-on experience, student misunderstandings will be replaced by explanations based more upon scientific explanation. Isolated high school science laboratory experiences do little to improve understanding of content and are no more effective than demonstrations (Singer, et al

2005; McKee, Williamson, & Ruebush, 2007). Of the students in an introductory college chemistry class, 44 % agreed that their high school lab work helped them to understand the content. Equally revealing, only 33 % agreed that the lab investigations stimulated their interest in chemistry (Mundell, 2007). A recurring problem is that while students may enjoy participating in lab investigation, long-term retention of material is low, perhaps because of a lack on mental engagement (Coan, 2005). “To many students, a ‘lab’ means manipulating equipment but not manipulating ideas” (Lunetta, 1998, as cited in Bell, 2004, p. 204).

Student laboratory experiences can be improved through conscious planning. The National Research Council (NRC) high school study offers four principles for effective lab work: labs should be “designed with clear learning outcomes in mind,” investigations “are thoughtfully sequenced into the flow of classroom science instruction,” experiments “are designed to integrate learning of science content with learning about the processes of science,” and laboratory experiences “incorporate ongoing student reflection and discussion” (Singer et al., 2005, p. 6). The NRC study examines some integrated learning experiences, the authors’ short hand for laboratory activities, that follow the above principles. Key ideas in integrated learning are that, one, lab experiences be selected for what students are expected to learn from them and, two, the labs are linked with other learning activities in the unit. Teachers should not assume students understand ideas behind a laboratory just because they have done it (Singer et al., 2005).

Lab work provides students with common experiences from which to build their scientific understanding. More effective labs engage students rather than have them follow directions as if from a cookbook. Properly designed laboratory investigations should have a definite purpose that is communicated clearly to students, focus on the processes of science as a way to convey content, incorporate ongoing student reflection and discussion, and enable students to develop safe and conscientious lab habits and procedures (Eisenkraft, 2007).

Research supporting the effectiveness of integrative laboratory experiences is not extensive because programs incorporating these ideas are new. Many of the studies are not fully relevant to the focus question because many of the integrated learning experiences were designed for middle school students. Fewer examples following the model exist for high school students, and the majority of these focus on physics content (Singer et al, 2005; Lynch, 2004; Bell, 2004). In addition, many teachers may not know how they can incorporate these methods (Deters, 2004). Others may see the problems but feel they don't have the time to implement new procedures because student experimentation and writing following the experiment may take longer to evaluate (Ferzli, Carter, & Wiebe, 2005; Coan, 2005). State standards that don't allow the time necessary to permit the inquiry approach may be another problem (Singer et al., 2005).

A key step in a model laboratory program involves increasing the student's mental engagement with the ideas. The more the student is thinking about the material, the more he is likely to gain. The thought process should begin as the student first starts to consider the lab. Thinking can increase even with investigations where, for safety reasons, a set procedure needs to be followed. Clough (2002) suggests adding thoughtful questions to pre-lab preparation, such as why the particular equipment is used. Many investigations can become more inquiry based by having students plan procedures, organize data, and develop their own analysis upon completion (Deters, 2004).

Individual reflection and the students' oral and written expression figure prominently in many studies of lab improvement (Singer et al., 2005; Duschl, 2003). Use of language engages more areas of the brain, making retention more likely, and as students increase use of language, many show an increased understanding of experimental methods and design. The use of prompts forces students to address questions they may not otherwise consider (Ferzli et al., 2005).

Peer-review strategies can help students become more critical of work, which may lead to desirable self-assessment (Lewis & Lewis, 2008; Bell, 2004). Student writing can be improved by using models of work completed by other students. This approach worked better when students didn't know who wrote the papers and when the papers were not from the students' class. Students were more likely to make critical, and therefore, instructive comments, when they knew they would not be criticizing another student in the room (Atkin, 2002;, Duschl, 2003). Self-assessment strategies, such as asking students to say what they thought they did or did not understand or to evaluate their own explanation of an illustration, led students to improve their performance on inquiry tasks. The wide applicability of the later may be limited because the studies involved only physics classes (Bell, 2004).

An important part of a successful lab program is the opportunity for students to discuss ideas in small groups and as a class (Lynch, 2004; Dushl, 2003). This type of a strategy may enable a teacher to evaluate student work more quickly and to provide the important needed feedback. However, finding time to allow this discussion may be difficult because as teachers point out, quality inquiry can double the time it takes to complete an investigation (Deters, 2004). Likewise, state standards tend to include more concepts than can be covered with the amount of time needed to do inquiry (Singer et al., 2005).

Teachers can get around the problems of having enough time to grade all student writings and to provide discussion time. For difficult concept labs, Coan (2005) conducts personal interviews on the day following the lab. She interviews one lab group at a time. Each student in the group needs to answer at least one question. She reports that students invested more effort and attention after the interview process began. "When students only have to write a lab report, often there is not much original thought involved, and some students participate more in the lab than others" (Coan, 2005, p. 64). Student ability to perform the chemistry calculations associated with lab concepts improved after the interview process began. Because the teacher's attention is on a small group, the remainder of the class needs an activity that engages them as they complete it autonomously. As an alternative to interviewing all lab groups, interviews can be staggered over several labs.

Many teachers favor class discussion of labs (Duschl, 2003). One high school physics teacher has students prepare white board presentations of their data. Random groups are selected to present to the class (Bleicher, 1994). Bell (2004) considers laboratory experiences that more accurately reflect the processes followed by working scientists to be beneficial in raising understanding of the processes of science. Students in a college program participate in an online discussion group for the class after an investigation (Clark & Chaudhury, 2008).

Effective laboratory investigation enhances student achievement and expands student understanding of scientific experimentation. While the theoretical foundation for effective lab implementation is in place, problems remain in making the change in classrooms. Published reports of strategies for effective lab work in the high school environmental science classroom are limited. In addition, this review finds state standards may inhibit teachers from spending the necessary time on inquiry investigations. Furthermore, the need for timely teacher feedback of student use of the science language -- both oral and written -- can become a time burden for an already busy teacher. The literature showing how teachers effectively incorporate inquiry and efficiently manage the time necessary is limited. Many studies of integrated lab experiences deal with elementary and middle school classes. Existing high school level studies emphasize physics content and experiments. While addressing gains in student understanding, these studies do not address the challenges faced by teachers who need to include numerical grades for lab work as a major portion of a semester grade.

Many of these studies emphasize the need for students to do more than write about their lab work if they are to learn more from the experience. From this research, I saw a need for students to discuss their experiences. Among the useful ideas, I especially saw how small group presentations to the class using whiteboards like posters could engage students in talking about the lab and allow for class discussion of the lab quickly following the completion of the investigation. The work with class posters described by Bleicher (1994) provided me with one of my eventual treatments. Another treatment, the use of preliminary questions to help students understand why they are following certain procedures, came from the ideas of Clough (2002).

METHODOLOGY

The research study involves four AP Environmental Science classes at Northview High School, Johns Creek, Georgia. A total of 120 students are in the classes, with class size ranging from 28 to 31. Most of the students (88%) are twelfth grade, and the remaining 12% are in eleventh grade. Most of the students have already taken biology, chemistry, and are currently taking or have already taken physics. Because it is an AP class, these students needed minimum grades of at least a “B” in other science classes before enrolling. As a result, the percentage of students who will be attending four-year colleges is probably higher than the 91% of recent Northview graduates who currently attend four-year colleges. With regards to motivation, the juniors in the class have more at stake because colleges will see their grades and scores on the AP exam during the application process. The seniors may have started the year with greater concern about their grade and possible achievement on the AP exam, but as the year progressed and more knew which college they would be attending and whether the college will accept this course for credit, overall effort from the seniors declined. This decrease in effort came from approximately 30% of the seniors.

I chose this group of four classes because it was the majority of my teaching load, and by using four classes, I could compare different treatments for the same lab. Because every lab experience is unique, the effectiveness of each treatment would be harder to analyze if the content or procedure were different. With all participants doing the same lab, any differences between classes were more likely to result from the treatment. Further, each class did each treatment three times, and this enabled me to reduce the influence of individual student differences on the results. The class sizes were large enough to get a range of responses and also to allow meaningful evaluation even if several students were absent or did not complete assessment forms. The four classes were fairly close in their grade average. Term grade averages (based upon 69 being "failing" and a maximum of 100), for the first semester, were, first period -- 78.87, second period-- 81.25, fifth period --80.44, and sixth period -- 79.51. At the end of the treatment in the second semester grades were first period --78.35, second period -- 81.94, fifth period -- 78.77, and sixth period -- 75.75.

Table 1 describes the overall research questions and data collection design.

Table 1
Research Matrix

Data question	Data source 1	Data source 2	Data source 3
1: What is the effect of pre-lab and post-lab treatment on understanding of environmental science content?	Student self-assessment survey	Lab quizzes and rankings from class comparisons	Pre- and post-lab unit assessments
2: How does post lab discussion compare with individual student understanding without discussion?	Lab quizzes and rankings from class comparisons	Pre- and post-lab unit assessments	Student self assessment surveys
3: How well do poster-sessions and group discussions provide meaningful feedback to students?	Student self assessment surveys	Lab quizzes	Pre- and post-lab unit assessments
4: How does group discussion or poster presentation affect teacher evaluation time?	Record book of time spent to evaluate	Comparison of mean student score to evaluation time	Student self assessment survey

Before beginning the treatment, all the students completed an anonymous survey (Appendix B) about their attitudes towards lab work. I designed the survey so it would allow me to see if my current students shared similar attitudes towards lab work that first inspired this project two years ago. The survey information allowed me to identify students who said they didn't find post-lab discussions to be useful. Students who responded that discussing lab work after it is over is not helpful formed the group of individual students from which I analyzed individual improvements based upon the treatment plan. Initial surveys were completed in October of 2010.

At the start of each unit, students answered a set of pre-test questions related to the content of the unit. The questions asked about material that they might have studied in an earlier class or learned through their day-to-day lives. The question content was related to both what the labs covered as well as other lessons in the unit. As a result, these surveys reinforce the main treatment evaluation but cannot by themselves answer the research questions. The treatment schedule included assessments and lab work done in November and December 2010, continuing through January and February 2011.

For the treatment plan, students in my four classes conducted the same lab activity. Each class did something different in its pre-lab preparation or post-lab analysis. One class reviewed a written background and procedure the day before the lab, and each student wrote as homework, a several sentence summary of background and procedures before the lab. All classes then performed the same lab. Afterwards, students in a second class worked in groups to discuss a specific portion of the lab, summarized this information on a large white board, and then reported their information to the entire class. The third class discussed the lab together as a class. The fourth class did nothing besides completing the lab. After the lab and respective discussion, each class completed a quiz with the same questions over the lab contents and procedures.

The following example illustrates how the treatment worked with the third lab topic -- monitoring macro-invertebrates as a measure of water quality. The second period class completed a homework assignment explaining what the presence of macro-invertebrates tells observers about water quality and how to do this type of testing. The following day, all four classes determined the water quality index of their stream; students on each team were given ten cutouts of stream organisms and then, with the use of a key, determined a numerical score representing the health of the water. On the day after the lab, the first period (no discussion treatment) and second period (preview instructions) classes took a quiz about the lab when they first came to class, then later in the day, groups in the fifth period class answered one of the following questions on their whiteboard: What are the advantages of biological monitoring? What common features are found in group 1 organisms? In group 2? In group 3? What organisms did your group find? After each group shared the information on its whiteboard with the class, the fifth period class took the same quiz as the other classes. Finally, on the same day, sixth period conducted a class discussion without the whiteboards of the same questions used by the groups in fifth period, and then the students took the same quiz as the others. Each treatment took portions of two class periods -- one class meeting for the lab and at least a portion of another day for discussion and quiz. Appendix C shows the preview material, and Appendix D displays a copy of the quiz questions.

Each class rotated through four treatment procedures with subsequent lab activities. The rotation continued until each class did each treatment method three times. Table 2 shows the treatment schedule. For example, the first period class completed the following treatment schedule: treatments 1-4 -- discussion, preview, no discussion, whiteboard session; treatments 5-8 -- whiteboard, no discussion, discussion, preview; treatments 9-12 -- discussion, whiteboard, no discussion, preview. Final analysis looked at the rankings of average class quiz scores for the twelve lab experiences.

Table 2
Treatment Schedule by Class Period

Experiment	#	Date	Preview class	Whiteboard class	Discussion class	No treatment class
Biological oxygen demand	1	11/16	5	2	1	6
Chemical water testing	2	11/30	1	6	2	5
Macroinvertebrate ID	3	12/3	2	5	6	1
Cleaning oil spill	4	12/10	6	1	5	2
Copper from malachite	5	1/6	5	1	2	6
Mining labs	6	1/17	2	6	5	1
Electrolysis	7	1/24	6	2	1	5
Corrosion labs	8	1/28	1	6	5	2
Polymer properties	9	2/2	6	2	1	5
Polymer making	10	2/4	5	1	6	2
Heat of combustion	11	2/9	2	5	6	1
Lemon battery	12	2/15	1	5	2	6

Note. "#" represents treatment number.
Dates from months 11 and 12 were in 2010. Months 1 and 2 were in 2011.
Class columns identify which class period did which treatment; e.g. period 5 had preview materials for treatment 1.

After the treatment, each student answered a quiz over the lab content, procedure, and results. Quizzes were scored, and a class average calculated. I compared the class average for each treatment and determined an overall ranking from 1-4 (highest to lowest class average) for each treatment lab. After the twelve treatment labs were completed, I compared the overall effectiveness of each treatment. If the class average for one treatment, say use of posters, consistently was the top ranked, then I could take this as evidence that this method was more effective.

Students completed the unit pretest as a post-test at the end of the unit to help provide data on their overall learning during the unit. Results from beginning to end were compared by class and by individual. After the treatment, students completed another survey about their attitudes towards labs and with more detailed questions about the treatment methods. Opinions of the treatment methods were compared to student performance on lab quizzes.

The following measures enhanced the reliability of the class analysis. First, the potential sample size for each comparison was large, with approximately 30 students doing each treatment. Second, rotating the treatment through four classes three times reduced the effect of class meeting time (e.g., beginning of day compared to end of day) and different student abilities. Third, the data for each lab were analyzed using ANOVA to determine significance of variations between the classes. Finally, student opinions of the treatment effectiveness were obtained at the conclusion of the treatment schedule.

For the students identified in the beginning survey as not finding lab discussions to be helpful, I recorded their performance on each treatment quiz. I compared their score to the average score earned by the class that had no treatment. I calculated the percentage of their score to the group average (subject score/ class average x 100%). The score allowed me to know if the individual student was above or below some reasonable value, in this case the average for that particular lab. Next, I compared the number of times each student was above average for the two discussion methods -- whiteboards and class discussions -- to the other two treatments. The goal was to have some type of objective measure as to whether the discussion was or was not helpful to these students. These data were compared to the students' opinions of the effectiveness of the treatment.

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. As a result, I needed to structure the research so that I didn't know the identities of any of the students. Students identified their work with a number of their choice and a class period identifier. An example might be "5-1718," where the "5" represented the class period and the "1718" the student's unique number. This allowed me to keep track of class data without knowing by name the student completing any quiz or survey. I could place students in a particular class, but I do not know who they were. This protected student privacy while allowing me to follow the progress of an individual student.

Originally, I planned on interviewing some students at the conclusion of the treatment as a way of obtaining more detailed response. However, after taking the IRB course, I decided I couldn't use interviews without having an informed consent form because I would know the identity of the students. Obtaining the informed consent forms at my county would require school board approval, a process that could take up to five months to receive.

Overall, the research methodology provided a large amount of data. The treatments were used twelve times, three times with each class. For each experiment except one, the sample size was greater than 75 students. The student opinion surveys highlighted a group of individuals who shared negative attitudes towards post lab discussion. The study followed these students, and the ability to compare the effectiveness of the treatment on individual students strengthens the reliability of the data analysis.

DATA AND ANALYSIS

Student Survey From Before the Treatment

Students in all four classes completed a survey (Appendix B) about their attitudes towards participating in labs and how labs can help them learn better and more effectively. This information survey was partially designed to affirm that my current students shared many of the attitudes towards labs that gave me the idea for this project two years ago. In addition, student opinion of lab discussions determined the students I would use in my study of individual performance. This section reports on the answers to the survey questions most directly related to the focus question of this capstone project.

Detailed Analysis of Specific Questions

Survey questions # 2, 3, and 4 dealt with topics related to the formation of the study topic, mainly do my current students have similar attitudes towards labs that I measured two years ago? All results represent percentages of the entire sample (N = 95) from four classes. In question 2, where students picked the phrase that best describes how they consider doing a lab as a learning tool, 15% of the respondents said labs were "unhelpful" or "very unhelpful." Question 3 asked about how helpful questions were after the lab, and 29% of the students chose "unhelpful" or "very unhelpful." Like with my 2009 study group, many students, 66% in question 4 of this preview survey, indicated that completing a lengthy, multiple-part, "traditional" lab report was unhelpful or very unhelpful in increasing learning from a lab. An ANOVA of the entire data broken into classes did not reveal significant differences between the classes. The following are the p-values: question 2, $p = 0.293$; question 3, $p = 0.471$: question 4, $p = 0.143$. Together the student answers to these questions support the need for finding alternative ways of dealing with the learning experiences associated with labs.

Question 12 asked students to pick what they considered to be the most effective way of conducting a lab. Of most relevance to the capstone project, 41% (39 out of 95) said they preferred working in a group of three to five people. Another 16% (15 out of 95) preferred “working with one other person of your choice.” In the student explanation, many stated the viewpoint that working with others was important because they often depended on the other people to explain the lab to them. The treatment data will show if group work, at least as far as students talking about the lab together, is indeed helpful as the students indicate.

Different Class Treatments

The main capstone project compared how well students that completed different lab analysis treatments did on a post-lab quiz. Table 3 summarizes the mean post-lab quiz scores by class treatments with ANOVA p-values for each experiment. The highest class-average for each experiment is indicated in **blue.** A more detailed version of the data is included in Appendix E.

Table 3

Class Period Mean, Standard Deviation, and ANOVA p-value of Post-lab Quiz

Treatment number	ANOVA p-value	Prev. mean	Prev. SD	White mean	White SD	Disc. mean	Disc. SD	NT mean	NT SD
1	0.920	3.19	1.74	**3.37**	1.47	3.17	1.55	3.08	1.35
2	0.018	3.91	1.27	**3.96**	1.08	3.85	1.54	2.81	1.17
3	0.006	3.91	1.63	4.72	1.19	**4.92**	1.35	3.03	1.07
4	0.071	5.41	1.69	**5.70**	1.69	5.52	2.00	5.29	1.83
5	0.006	**7.10**	2.37	5.00	1.81	5.96	1.93	5.65	2.04
6	0.000	4.63	2.29	**5.54**	2.57	3.85	2.58	1.80	1.21
7	0.412	4.20	2.08	4.50	2.64	4.57	3.13	**5.58**	3.17
8	0.172	**5.22**	1.54	4.55	2.26	4.00	1.83	4.29	2.10
9	0.000	7.33	1.37	**7.54**	1.71	4.75	2.31	5.39	1.58
10	0.000	**6.76**	1.48	4.95	2.19	4.22	1.63	4.71	1.59
11	0.018	3.29	2.27	**4.53**	1.91	3.70	0.95	2.38	1.45
12	0.005	4.77	1.83	**6.68**	1.70	5.55	1.97	4.54	2.34

Note. "Prev." is the treatment involving reading material in advance.
"White" is the poster discussion session.
"Disc." is the class discussion.
"NT" is no treatment.

The ANOVA p-values show that there were significant variations ($p < 0.05$) between the quiz results for eight of the twelve experiments. This suggests that the different treatment methods had some effect on the students' performance. In addition, for five of the treatments (#6, 9, 10, 11, and 12), the median quiz score from the highest ranking class – each doing some type of treatment – was more than one standard deviation higher than the mean score from the "no treatment" class. Of the treatments, the class doing the whiteboard poster discussion earned the highest quiz average seven of the twelve times. This compares with three highest scores for the preview presentation and one highest score for each of the discussion and no treatments.

For each lab, the treatment used by the students in the period was ranked by the class average on the assessment quiz, with the highest average score receiving a "1" and the lowest receiving a "4." Figure 1 shows the overall rankings of the four treatments over the twelve lab activities.

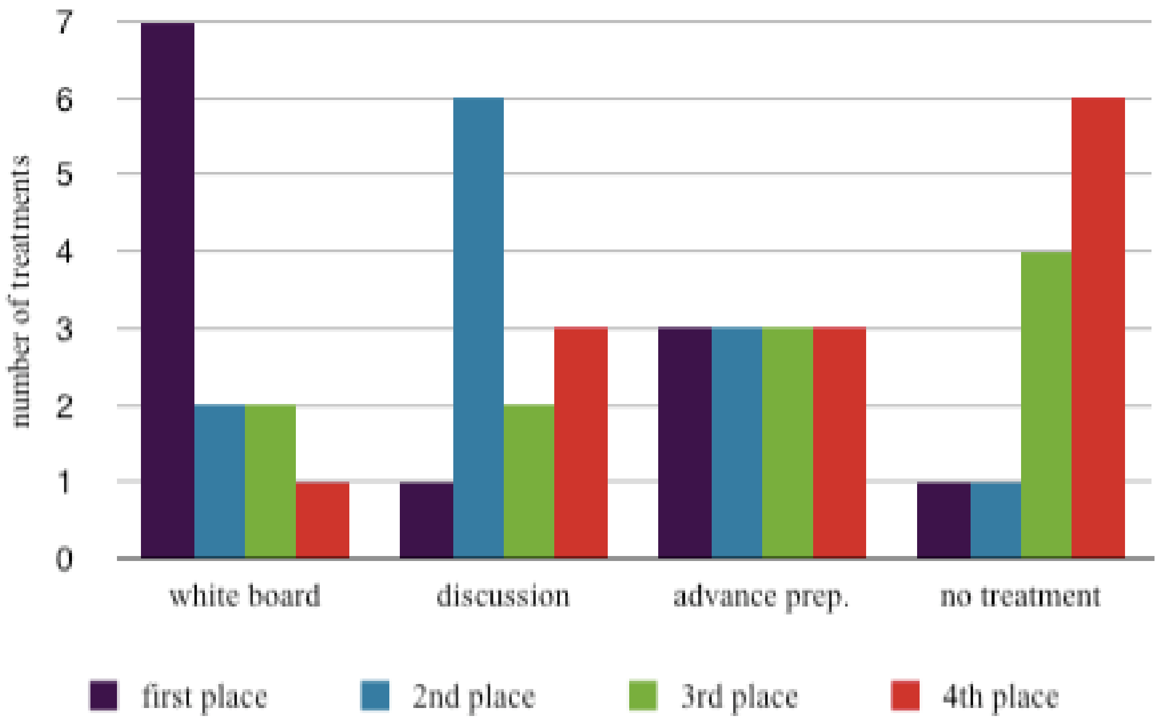


Figure 1. Rankings of Quiz Scores by Treatment Type. *N*= 12.

Classes conducting the whiteboard discussion treatment method earned the highest quiz scores on seven of the twelve labs -- 58% of the treatments. Whiteboards finished second in another 17% of the labs. For 75% of the treatments, students in classes where the whiteboards were used earned one of the highest two class quiz score averages. Another way of processing the strength of the treatment is to assign 4 points for a first place finish, 3 for a second, 2 for a third, and 1 for fourth. Using this method over the twelve labs, the whiteboards end up with a weighted average of 3.2, the advance preparation -- 2.6, the group discussion 2.4, and no discussion -- 1.8. The ANOVA for these rankings shows significant difference between groups ($p = 0.021$). This approach shows the strong performance of the classes using the whiteboard in discussions.

Overall, discussions with either the white boards or as a class resulted in higher student achievement on post-lab quizzes than when no discussion was held. The discussion format probably clarifies the important information for students, and the white board may increase student memory through both visual and kinesthetic stimuli.

While the data as reported in Figure 1 support the use of white boards as an effective post-lab discussion, a more detailed discussion of the results is necessary to eliminate results being affected by the four different sample groups, such as student ability or time of class meeting. Each class alternated its treatment method with every lab. Appendix F contains a more detailed breakdown of the results by class. What is apparent is that the whiteboard method worked well in all the classes. In each class, at least one of its top scores came when the students presented their group results with the whiteboards, and at least two of the whiteboard presentations per class period coincided with a first or second ranking. The data also show that treatment was more important than the class, as the distribution of mean quiz scores among classes is fairly even. This can be seen quantitatively by calculating the weighted average for each placement. As when done with the treatment above, a first place score earns 4 points and a fourth place earns 1 point. When this is done to the classes, first period averages 2.4 (out of possible 4), second period averages 2.7, fifth period averages 2.6, and sixth period averages 2.3. The ANOVA of these means gives $p = 0.891$, an indication that the rankings are not significantly different.

To evaluate further the overall effectiveness of each treatment, the class quiz score mean was compared to the mean score for the entire sample (Appendix G). The class using the whiteboard discussion had a mean higher than the sample mean eight times (experiments 1, 2, 3, 6, 7, 9, 11, and 12). The preview class treatment exceeded the sample mean seven times (experiments 1, 2, 5, 6, 8, 9, and 10); the discussion group exceeded the mean five times (experiments 1, 3, 5, 7, 12), and the no treatment three times (experiments 1, 2, 7). The number of times that the whiteboard exceeded the mean sample score reinforces the sense that using the whiteboards aided student understanding.

Students completed a second survey (Appendix H) at the end of the treatment that asked them to evaluate the four different treatments. Table 4 shows the numerical results for each of the methods and which one the students preferred.

Table 4
Student Post-Treatment Opinions Regarding Helpfulness of Various Discussion Methods

Category Responses are % where N = 80	Very helpful	Helpful	Unhelpful	Very unhelpful
Preparing and watching white-boards discussion	6%	58%	29%	8%
Discussing the lab as a class after the experiment	40%	46%	6%	3%
Doing advanced reading before the lab explaining the main idea	14%	41%	35%	9%
Doing the lab but not discussing it afterwards	1%	16%	28%	49%

The students rated the class discussion more helpful than using the whiteboards, which in turn were more helpful than the preview materials. Consistent with the class quiz results, the “no discussion” treatment was overwhelmingly the least helpful. While not completely in agreement with the already discussed quiz results, the student opinions trend in the same direction of the quiz scores, that is, discussion (whether whiteboard or class) is better than no after lab discussion (the preview group and the no discussion treatment). No attempt was made to analyze these opinions by class because numbers of respondents were similar. For example, in question 1, the breakdown by class of helpful responses was 10, 11, 11, and 14.

The second category of questions from the final survey asked students to compare the effectiveness of class discussions to the white board presentations. Approximately 50% (39 out of 80) preferred the class discussions to the whiteboard presentations. One quarter (21 out of 80) preferred the whiteboards to the discussions, and the remaining 25% (20 out of 80) said class discussions and whiteboards were equally helpful. The students recognized that discussion helped them do better on the quizzes even if their choice of discussion style did not match their performance. This observation points to a perception difference in the white board experience that will be discussed further in the following interpretation section.

Within the second category of final questions, ranking the preview treatment compared to the discussion methods was left out because three of the spring semester labs had no preview materials. School closure due to an ice storm prevented distribution of materials for one lab, and for two others, adequate prepared materials could not be found in time to make them available before the labs. In addition, my teaching journal included several mentions of students leaving the printed preview materials in the room and recorded five to seven students per experiment who didn’t do the preview homework. As a result, much of the treatment schedule for the last six labs was a comparison between whiteboard, class discussion, and two classes of no other treatment. While the deviance from the research plan made it harder to evaluate adequately the preview plan as a learning method, it doesn’t weaken the case for the whiteboards. The three labs for which the preview materials were not available were numbers 5, 8, and 10. In each case, the class that should have received the preview materials had the highest quiz average. Had the preview materials been available and been effective, this could have increased the class quiz average for the preview period, but it would not change the 1-4 ranking of the whiteboard and discussion nor affect the data as show in figure 1.

Individual Student Data Analysis

On the initial survey, students were asked to rate how helpful were teacher and class discussion of the lab concepts after doing the lab. Students who checked off "unhelpful" or "very unhelpful" became the pool of individuals to follow after the treatment was completed. For these students, three types of data were analyzed. First, their quiz scores for each treatment were compared to the average score for the "no discussion" class. Notice was made when the students did better after a whiteboard or class discussion compared with no discussion. Second, their scores on the pre-unit and post-unit concept test were compared to see if they showed improvement as a result of the unit. As already mentioned, the pre- and post- concept test can support that learning took place, but it can't say the student learned from the lab work. Finally, students' attitudes toward the helpfulness of discussion and whiteboard as recorded on the final survey were compared to their initial statements. Appendix I lists the data for individual students.

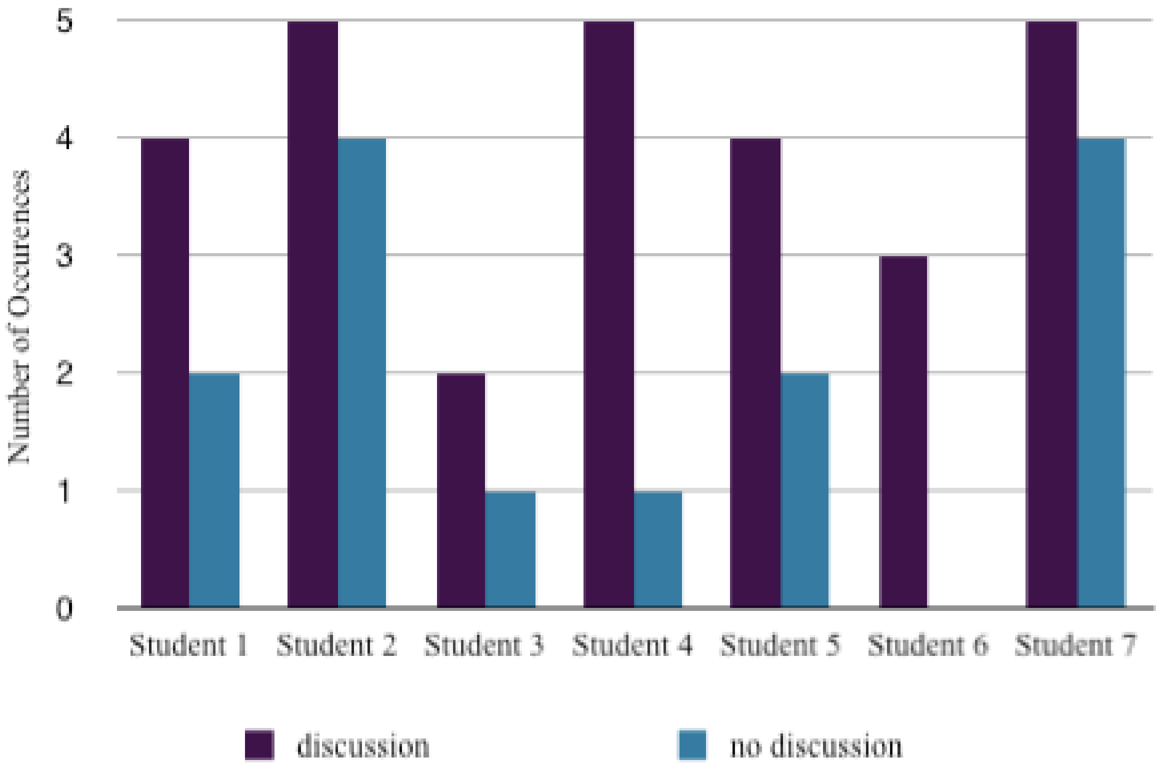


Figure 2. Individual Comparisons of Number of Quiz Scores Above Mean Score Of No Treatment Class.

As the data in Figure 2 above show, the whiteboard and class discussions made an impact on a few of the students who initially did not feel that post lab discussions were helpful. Seven students out of 18 scored better than the mean quiz score when involved with either discussion method than they did with no discussion. One of these (student 2) earned two quiz scores at least twice as high as the average during two whiteboard sessions. Two others student (students 5 and 7) earned one score double the mean while using a whiteboard. While student 2 usually did better than average, the only times the scores were double the average were with the white boards. Student 6 only earned above average scores when using a whiteboard, earning three above average scores during the three whiteboard treatments. One of these students commented on the final survey that, "The whiteboards helped because I could see the answers in my head when I took the quiz." Another said, "I paid more attention to the other groups because I wanted to tell if they did a good job." A third said, "Discussions helped because they covered all the answers on the quiz." It seems that for some students hearing and/or seeing answers helps them to remember better.

Most of the others in this study of individuals did equally well with or without discussion. Four students out of 18 were the exceptions with either more above average scores with "no discussion" treatments or more below average scores occurring with "discussion" treatments. Eleven (*N*=18) of the sample earned above average quiz scores at least 50% of the time; seven of these, with no discussion, earned at least four scores (out of possible 6) over the average. This suggests that a number of the students in this sample pool of those not finding discussion helpful are there because they understand the concepts well. Some of the scores above, especially the doubling of average scores, suggests that even the brighter students gain some benefit from the post-lab white board discussion.

The results from the pre- and post- content test show that nearly all the students in this sample improved in their understanding throughout the unit. All the students (14 out of 18) improved their score on the test by at least four points from the beginning to end. Two exceptions to this were the students who had the two highest scores on the pretest, and two other students who remained at the same level of comprehension. Interestingly, the later two students were two who showed measurable gains from the lab discussions. Perhaps for these last two, the discussion helped them have a short-term memory of the ideas, but that this didn't help them remember it long term.

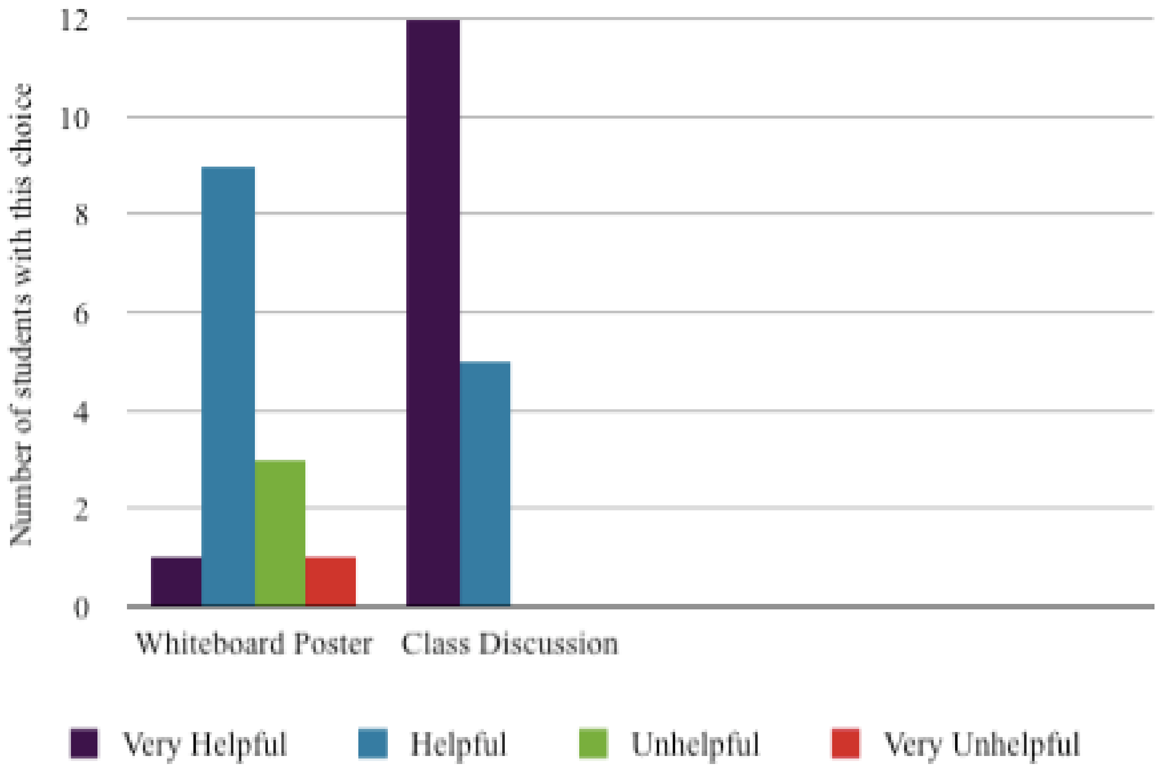


Figure 3. Student Opinion From Final Survey About Helpfulness of Post Lab Discussions. *N*= 17.

Perhaps the strongest indication that the discussions were helpful came from the final survey results: Figure 3 above. Ten out of 14 (71%) of the individuals in this sample, selected because they indicated initially not finding discussion helpful, said the poster session was helpful or very helpful. Furthermore, the entire sample pool (*N*=17) said the class discussions were helpful or very helpful with 71% choosing very helpful. Interestingly, three of the respondents whose quiz scores showed the most improvement from whiteboards evaluated the whiteboard sessions as “unhelpful.” These observations, along with other inconsistencies between students’ ranking, highlight the unreliability of student opinions, reinforcing the need for educational data collection to include more than one measurement tool, especially if one tool is student opinion.

Degree of Feedback For Students

The third research question asked how well poster and class discussions provided meaningful feedback to students. This question became harder to triangulate as the treatment took place because there was no easy way to return student work without compromising the needed anonymity. Work presented on whiteboards couldn't count as a grade, and what students presented often was of lower quality than I would see on graded assignments. The student self-assessment data are reported above in Table 4. As stated earlier, over 50% of the end of treatment respondents, found either the poster discussion or class discussion to be helpful. Some of the students' comments on the final evaluation are worth reviewing. With regards to whiteboards one said, "I couldn't zone out like I sometimes do in lectures. I actually had to contribute and write things down." The whiteboards were helpful because "hearing the information presented again and hearing it explained in different ways." Another stated about the helpfulness of whiteboards, "We got to focus in depth on one aspect of the lab and our peers taught us other aspects, so it saved everyone time and trouble, but we still learned."

Student final comments about helpfulness of class discussions were generally less elaborate than the ones about the whiteboards. Many said something similar to this one; class discussions were good because "the teacher is the one guiding the course." Another common viewpoint was "the discussion told me all I needed to know for the quiz." One liked the discussions because "we didn't have to do work." This last comment agrees with my teaching log. On at least three occasions, I wrote that few students were involved in the discussion, with most of the input coming from fewer than five students.

Usually, only one or two people would ask questions. My sense of the class discussion frequently felt like the students were waiting for me to tell them the information. This matches with student data from the preliminary survey in which students favored lab formats where they were given the answers either by the teacher or other students in their group. I suspect this research question would yield different results depending on whether the student or the teacher was defining “meaningful feedback.” Many of the students care more about their grade and less about improving the quality of their work than I, as a teacher, might.

Group Discussion Effects on Teacher Evaluation Time

The final research question asked how group discussions affected teacher evaluation time. This question, too, changed in the course of the planning and treatment process. The original intent had students evaluating their lab reports based upon feedback from the discussion, and then I would use the student’s grade as a preliminary grade. As the treatment evolved to protect student identity, lab reports were dropped from the plan. There were, however, meaningful data that I can use from this study to address the question of how can this treatment affect the time of grading lab reports. Triangulation will be hard because most of the valuable data are anecdotal pieces recorded in my logbook. Probably the most important idea to come from the treatment is the overall value shown by whiteboard and class discussions as a way of improving student understanding of lab ideas compared with doing no group work. Seeing how student-understanding increases with a discussion can reduce future lab evaluation work as some labs could be discussed using whiteboards instead of with a lab report. I am interested in seeing how the quality of whiteboards might change if students knew they would be receiving a grade for doing this.

Without any doubt, the discussion and post lab quiz processed the experience for the students much faster than I previously dealt with lab reports. A discussion and quiz were completed in no more than two days, while grading lab reports could take as along as a week. As recorded in my logbook, the white board discussion took usually two or three times the time necessary for a group discussion. Preparation of white boards could run close to 20 minutes; then to have each group present took as much or more time than the group discussion in other classes. Finally, the requirements for student confidentiality opened up a way to save teacher time. Students scored many of the post-lab quizzes. During their once a semester “lab assistant” time, they were given a set of papers from another class period. Because only the originating student knew what identity number was chosen, the grader didn’t know the student by name. I can see a way of using student assistants to grade work in the future. Students could grade work from another class, and I could use the grades if I knew the student selected identification number.

The data represent different treatments conducted with over 100 students in four classes. This allowed comparison of the treatment method between classes. In addition, the effect of the treatment on individual students was followed. Student surveys allowed objective treatments to be compared to subjective student opinion. Finally, the treatment allowed some comparison of the time necessary to implement and the degree to which it reduced teacher preparation time.

INTERPRETATION AND CONCLUSION

Conclusions

This research project began with a goal of improving student comprehension of concepts covered by lab work. The main treatment involved four AP Environmental Science classes, each conducting labs in four different styles. One class received preliminary background before the lab day. A second class did the lab, and then in small groups, students discussed a portion of the lab and presented their findings to the class using large whiteboards. After the lab, a third class reviewed the main ideas in a teacher-led class discussion. The fourth class did the lab with no other treatment. All the students conducted the same lab but differed in the how the lab was discussed when finished. At the end of the treatment, all the students took a quiz covering the main ideas in the lab.

The primary research question asked whether student understanding of environmental science concepts improved with pre- or post-lab treatments. The data show that post-lab discussions improve student understanding (Figure 1). Moreover, the class where students presented findings with whiteboards consistently did better than other treatment plans (seven out of twelve highest class quiz average). Of the other five experiments, the whiteboard class finished second three times, and when it finished third, the results were not significantly different. The reliability of this is shown by all classes earning top quiz scores (1-2 per class) when using whiteboards while the overall average ranking for all twelve treatments was statistically similar (mean score 2.3 – 2.7, $p = 0.891$). Data from individual students ($N = 18$) show the helpfulness of the whiteboard discussion in terms of increased quiz scores. Three had quiz scores more than double the overall mean when participating in the whiteboard discussion; another student only exceeded the mean score on the three uses of the whiteboard.

The second research question asked whether students gained more from discussion after the lab as compared with no discussion. The data from the treatment plan support this, as the two discussion styles produced higher quiz scores than no discussion after the lab (Whiteboard mean 3.2, discussion 2.4, no discussion 1.8, $p=0.021$). Student post-treatment opinions ($N= 80$) reinforce this with as a majority considering the “discussions” helpful and “no discussions” as unhelpful (whiteboards 64% helpful, discussions 86% helpful, and no discussion 77% unhelpful). Finally, individual student data show gains in their performance after a discussion session (Figure 2).

The whiteboard discussion works for several reasons. First, as indicated by student opinion in the preview survey, many students rely on someone else telling them what is important in the lab. Second, students are more involved in the discussion when they have to prepare a whiteboard. Their involvement increases the stimuli that may lead to increased mental processes, with subsequent gains in understanding and memory. The whiteboard discussion involves small group interaction, some kinesthetic process for the person writing on the board, oral presentation of the data to others, and visual cues for those reading the whiteboards. Moreover, the process involves increased attention associated with peer interactions. Some students may try harder because they want their poster to impress others or at least not embarrass themselves to their peers. Another student indicated that, “I pay attention more to the other groups because I want to see if they were correct.”

Because of the increased stimuli mentioned above the whiteboard is likely to affect more students than a discussion. During a discussion, especially in a group of almost 30, students were more passive. Only a few would ask questions or volunteer information. Others indicated that they were less attentive or were waiting for me, the teacher, to give them the information. Overall, the discussion format produced less stimulation that can lead to the mental processing necessary for learning.

Student opinion considered class discussion more helpful than doing a discussion with the whiteboards, something not supported by the performance data. Many students considered the discussion as the easier option because they had “less work” and the “teacher would make sure all the information was covered.” Another factor that might have influenced student opinion came from how they perceived the work of others. Several students commented that they couldn’t trust the poster session because the students in some of the other groups didn’t take the presentation seriously. I agree with this perception because some students only do high quality work when they know it will be graded. Another contributing factor is a tendency for students’ to gauge inaccurately their performance. This is likely with the study because they did not receive their quiz scores, and as a result, they may not have noticed doing better after the whiteboards.

The third research question asked if the discussion sessions provided meaningful feedback to students. The results to this question are less conclusive because most of the results depend on student opinion. As mentioned above, the students found the discussion sessions helpful. However, because the students did not receive their scored quizzes or get more than a general evaluation on their poster presentations, there is no quantifiable or objective data on whether this treatment provided enough feedback to enable students to improve over time.

The fourth research question addressed how the treatment affected teacher evaluation time. In this regard, the treatment was very successful. Completion of the lab and assessment of student work took much less time than I needed previously when students completed individual lab reports. The lab quizzes could be graded in less than 20 minutes per class, and the scores gave me an idea of how well students grasped the ideas from the lab. While not formally evaluating group presentations, I could tell what they did and did not understand from their oral summary. If necessary I could address their misconceptions or inaccuracies at the moment. The whiteboard treatment required more class time than the discussion, but I had more information about individual student understanding from the small group discussions and whiteboard presentations. Much of this came about because the students were the doing the work, and I was more evaluator than instructor.

Areas for Future Investigation

This research project changed the normal class dynamics in that it had four different treatments for the same lab. Instead of four classes using the same lesson plan, there was a different plan in each class. In addition, as a research project, student work was not evaluated in a way that could contribute to the student’s term grade. If continuing this study, I would like to see the results of the following actions. First, I would like to include student quiz grades as part of the students’ lab grades. I noted many times in my record book that students were not participating in the treatment or approaching the quiz with the same level of seriousness that I would see on a regular, graded assignment. Recording the quiz scores as grades would also allow me to measure if the use of whiteboards could improve scores even more. In this research, while the mean class score from whiteboard discussion was higher than the no discussion method, these mean scores were still lower than what I would consider as mastery level. Second, when I originally designed my treatment plan, I wanted to include a score for the students' whiteboard presentation. I envision the use of a rubric where the group is evaluated on clarity and completeness of its presentation to the class. I would hope to see if this would improve the overall quality of the students' work. Finally, I would like to see whether the discussion method brought about lasting learning with students' scores on the quizzes corresponding to similar success when asked about the concepts from the lab on a unit test.

VALUE

This research project will influence my teaching in several ways. First, the student surveys informed me of what parts of a lab experience they found to be helpful and unhelpful. The initial student insight about lab reports not being effective and how many did not learn from doing the lab has already changed how I use lab work. Over my teaching career, perhaps because my county has an expectation that 25% of the class time will be devoted to lab work, I had let myself think that by doing a lab students would understand the associated concept. This project changed my thinking to view labs not as the end but as the opening for student learning. Students need more involvement with the lab ideas, such as discussing ideas in groups and presenting information to others, in order to make this an effective learning experience.

As both the student opinions and collected data make apparent, students benefit from discussing the ideas in a lab after it is completed. Including more discussion time will cause some change in the distribution of class time, especially if I will use the whiteboard discussions. The use of the whiteboards for a discussion can add almost one full period to the time necessary for a lab. I will need to look more carefully at my course plan to build in time for the discussion afterwards. I anticipate that I will find the increased discussion time by cutting some lecture time and putting more emphasis on students learning certain topics through guided readings or the use of podcasts that they watch outside of class.

The use of the post-lab quiz has helped me see the value and relative ease of using short assignments like this as formative assessments. Doing a quiz at the end of a period provides me with some quantifiable data as to how well the students understood what happened in class. Perhaps the biggest value from this is helping me to see that many students are comfortable leaving class not understanding the material. They don't ask questions, something that I may have associated with comprehension, but which may in fact more accurately reflect lack of engagement. This should be helpful as my county moves to a system where some measure of teacher performance is student improvement. Regular feedback may help me redefine how I use my time, such as spending more time at the beginning of the next class going over misunderstood ideas and incorporating more practice time in the lesson plan so students are working with the ideas more than just hearing about them.

Perhaps the greatest overall value of this research project was the insight I gained about students and their performance. In the past, I used student surveys frequently, but this was the first time in a long time where I compared their opinions to more objective data. This disagreement between student opinions of discussion and the whiteboard discussions shows me strongly how I need to balance student ideas with quantifiable data. In this case, seeing the disconnect between opinion and objective data led me to think more about why the whiteboards produced better results, something missed by the students.

In a related manner, this project helped me to learn more about the students I am currently teaching. For twenty years of my teaching career, I taught middle school students at two different, but academically similar independent schools. I did the research project with public school seniors, only my second year working with predominantly twelfth-grade students. Between the surveys, their work with the whiteboard and class discussions, and their performance on post-lab quizzes, I developed a greater understanding of them as learners. First, as the difference between student opinions of the helpfulness of discussions compared with the whiteboards show, many of them are poor judges of their abilities and what helps them to learn. This inaccuracy surprised me because as middle to high-level students who will be attending four-year colleges next year, they have a poor understanding of what is necessary for learning. I realized I expected more because they seem so much more mature than my previous independent school eighth graders who, perhaps because of the schools, actually were much more aware of what they needed to do to learn. This project helped me to see how my current students approach lab work. Perhaps because of the difference in school cultures, my current students don't have much experience with shaping their own understanding through experimentation. Many students anticipate the lab will confirm something they were told in class or expect that writing down a few observations would earn them a high grade. This indicates a need for me to spend more time at the beginning of the year developing expectations for how learning from labs and learning in general is what is important during this class.

Overall, the action research process should help me improve how I help students to learn. The project showed the value of data triangulation in guiding teacher understanding of the student learning experiences. I saw the value that can come from trying different approaches in similar classes and then measuring the results.   Most importantly, I learned how my students approach lab work, and through the research, determined how I could be both more effective in helping students learn and in reducing the time needed to evaluate student lab work. From these measures, both the students and I should benefit from this project.

REFERENCES CITED

Atkin, J.M. (2002). Using assessment to help students learn. In Bybee, R. W. (Ed.). *Learning science and the science of learning* (pp. 97-104). Arlington, VA: NSTA Press.

Bell, P. (2004). *The school science laboratory: Considerations of learning, technology, and scientific practice*. Paper prepared for the Committee on High School Science Laboratories: Role and Vision, July 12-13, National Research Council, Washington, DC. Retrieved May 1, 2011 from http://www7.nationalacademies.org/bose/Philip_Bell_%20Final_%20Paper.pdf

Bleicher, R. (1994). High school students presenting science: An interactional sociolinguistic analysis. *Journal of Research in Science Teaching*. *31*, 697-719.

Clark, D. B., & Chaudhury, S. R. (2008). Creating coherent inquiry projects to support student cognition and collaboration in physics. In Luft, J, Bell, R.D,, & Gess-Newsome, J. (Ed.). *Science as inquiry in the secondary setting (pp. 65-78).* Arlington, VA: NSTA Press.

Clough, M. P. (2002). Using the laboratory to enhance student learning. In Bybee, R. W. (Ed.). *Learning science and the science of learning* (pp. 85-96). Arlington, VA: NSTA Press.

Coan, A. (2005). Idea bank: Interview assessment. *The Science Teacher, 72*(3*),* 64-69.

Deters, K. (2004). Inquiry in the chemistry classroom. *The Science Teacher, 71*(12), 42-45.

Duschl, R. A. (2003). Assessment of inquiry. In Atkin, J. M., & Coffey, J. E. (Ed.). *Everyday assessment in the science classroom* (pp. 41-60). Arlington, VA: NSTA Press.

Eisenkraft, A. (2007). Personal testimony. Committee on Science and Technology House of Representatives. (2007). *Improving the laboratory experience for America* (serial number 110-9) (pp. 17-18). Washington, D.C: U. S. Government Printing Office.

Ferzli, M., Carter, M., & Wiebe, E. (2005). LabWrite: Transforming lab reports from busy work to meaningful learning opportunities. *Journal of College Science Teaching*. *35*(6), 31-33.

Georgia Department of Education. (2006). Chemistry performance standards. Retrieved March 1, 2009, from Georgia Performance Standards web site: https://www.georgiastandards.org/standards/Georgia%20Performance%20Standards/Chemistryrevised2006.pdf

Lewis, S. E., & Lewis, J. E. (2008). Seeking effectiveness and equity in a large college chemistry course: an HLM investigation of Peer-Led Guided Inquiry. *Journal of Research in Science Teaching*. *45(7)*, 794-811.

Lynch, S. (2004). *What are the effects of highly rated, lab-based curriculum materials on diverse learners?* Presentation to the Committee on High School Science Laboratories: Role and Vision. July 12. Retrieved February 28, 2009 from http://www7.nlationalacademies.org/bose/July_12-13_2004_High_School_Labs_Meeting_Agenda.htm

McKee, E.,V. M. Williamson, & L. E Ruebush. (2007). Effects of a demonstration laboratory on student learning. *Journal of Science Education and Technology*. *16(5)*, 395- 400.

Mundell, J. (2007). Personal testimony. Committee on Science and Technology House of Representatives. (2007). *Improving the laboratory experience for America* (serial number 110-9) (pp. 24-30). Washington, D.C: U. S. Government Printing Office.

Singer, S. R., Hilton, M.R., & Schweingruber, H.A., (Ed.). (2005). *America's lab report: Investigations in high school science*. Washington, D.C: National Academies Press.

Van Scotter, P., & Pinkerton, K. D. (2008). Assessing science as inquiry in the classroom. In Luft, J, Bell, R.D., & Gess-Newsome, J. (Ed.). *Science as inquiry in the secondary setting (pp. 107-120)*. Arlington, VA: NSTA Press.

APPENDICES

APPENDIX A

SUMMARY OF RESPONSES ON 2009 SURVEY

Appendix A

Summary of Responses on 2009 Survey

Ranked 1 or 2 Most helpful (class data) (*N*=48)	Ranked 1 or 2 Least helpful (class data) (*N*=48)	Category of questions	Ranked 1 or 2 Most helpful (Interviews)	Ranked 1 or 2 Least helpful (Interviews)
56%	17%	Read the textbook	7	1
54%	8%	Listen to teacher	3	1
23%	13%	Complete problems	4	0
8%	25%	Complete activity	1	1
2%	90%	Write lab report	0	8
25%	25%	Make notes	3	3
29%	21%	Talk with classmates	1	3

APPENDIX B

2010 PREVIEW SURVEY

This is an anonymous survey about lab investigations and their role in helping you to learn science. Please make up a number and letter code that can be used to follow you anonymously. Write it here _______________ and in your agenda, or notebook, or someplace where you will not lose it.

Please complete the questions honestly with as many supporting details as you can. Your answers will not affect your grade. While complete sentences are not necessary, the more details you can provide the better. For questions involving the choice of an opinion, please write the number of your response in the left margin.

1. Do you like to do lab work? Always (1) Sometimes (2) Never (3)
   Please explain why or why not..

For the following ranking questions, #2-11, please use the following scale and write the number in the left margin:
very helpful (1) helpful (2) unhelpful (3) very unhelpful (4)

2. Based upon your previous lab work in high school science classes, pick the phrase that best describes how you consider doing a lab as a learning tool?
   Briefly explain your choice.

3. After completing the lab, how effective in increasing your learning were questions about the lab?
   Briefly explain your choice.

4. After completing the lab, how effective in increasing your learning was the completion of a lab report following the Fulton county report?

Please consider the following and describe how helpful each was in understanding the lab.

5. working by yourself, like on a performance assessment.

6. working with one other person randomly assigned to you.

7. working with one other person of your own choice

8. working in a team of three to five people

9. preparing a pre-lab report where you wrote procedure and data tables in advance

10. teacher and class discussion of the concepts behind the lab before doing the lab

11. teacher and class discussion of the concepts behind the lab after doing the lab

Please elaborate on the best method for you from questions 5-11 and why this was better.

12. What is the hardest part of a lab assignment? Why was this hard? What could be done to make it easier?

Lab work involves several skills. Please use the following scale to evaluate your confidence in each area and write the number in the left margin:
Not very confident (1) I can get by (2) confident (3) highly confident (4)

13. Asking a testable question.

14. Identifying experimental variables, constants, and controls.

15. Correctly preparing and identifying graphs of experimental data.

16. Interpreting results to determine the outcome of the experiment.

17. Writing a conclusion that addresses how the experiment could have been done better.

So far this year in AP Environmental Science, you've done the following lab activities. For each one, please describe what was helpful about it as a learning exercise and what could have made it more helpful.

18. Tragedy of the Commons (fishing for goldfish)

19. Completion of guided learning about earthquakes

20. Preparation and use of suncharts

21. Preparation and use of population pyramids.

22. Graphing moose, wolf, and balsam fir numbers

Is there anything else you would like me to know about your work in labs? Thank you for your attention to this.

APPENDIX C

PREVIEW MATERIAL EXAMPLE

Chapter 1

BIOLOGICAL MONITORING

- Biological Monitoring
- Why Monitor for Macroinvertebrates
- Determining Stream Type and Sampling Location
- Begin Sampling For: Rocky Bottom Streams
- Begin Sampling For: Muddy Bottom Streams
- *Calculate Your Results*

Biological monitoring involves identifying and counting macroinvertebrates. The purpose of biological monitoring is to quickly assess both **water quality and habitat**. The abundance and diversity of macroinvertebrates found is an indication of overall stream quality. Macroinvertebrates include aquatic insects, crustaceans, worms, and mollusks that live in various stream habitats and derive their oxygen from water. They are used as indicators of stream quality. These insects and crustaceans are impacted by all the stresses that occur in a stream environment, both man-made and naturally occurring.

Aquatic macroinvertebrates are good indicators of stream quality because:

- They are affected by the physical, chemical and biological conditions of the stream.
- They can't escape pollution and show effects of short- and long-term pollution events.
- They are relatively long lived – the life cycles of some sensitive macroinvertebrates range from one to several years.
- They are an important part of the food web, representing a broad range of trophic levels.
- They are abundant in most streams. Some 1st and 2nd order streams may lack fish, but they generally have macroinvertebrates.
- They are a food source for many recreationally and commercially important fish.
- They are relatively easy to collect and identify with inexpensive materials.

Macroinvertebrates are present during all kinds of stream conditions from drought to floods. Macroinvertebrates are adaptable to extremes of water flow. Some may burrow when it is raining and flow increases. However, heavy rain in areas with a high percentage of impervious surface (most urban areas) can cause flash floods and carry macroinvertebrates downstream.

Populations of macroinvertebrates may differ in North and South Georgia. For example, since the Adopt-A-Stream biological index is based on dissolved oxygen, the “sensitive” organisms that require a lot of oxygen, such as the stonefly, may not be found in warm, slow-moving streams in South Georgia. That does not mean that the stream has bad water quality or habitat, just that streams in North and South Georgia support different populations of macros. If you are monitoring in South or Coastal Georgia, it is important for you to conduct monitoring each season for several years. Doing this will help you recognize biological trends in your stream so that you can determine which changes are natural and which may be induced by human impact.

Populations of macroinvertebrates may vary from headwater streams to the river mouth. For more information, please review “The River Continuum Concept,” Chapter 1, *Visual Stream Survey* manual.

Seasonal cycles can also affect the number and kinds of macroinvertebrates collected. Organisms such as immature stoneflies and mayflies will gain weight and size primarily during the fall and winter. During the spring and summer they may reach maturity and begin to metamorphose into their adult (non-aquatic) stage. Therefore, the presence of aquatic macroinvertebrates will tend to be more evident during winter and spring just before metamorphosis. After adults emerge, females lay eggs near or in the water. Soon after, the larvae and nymphs hatch and begin to grow, feeding on leaf litter, detritus and other organic matter that might be present. For more information on macroinvertebrates and their life cycles, please turn to “Some Background On Aquatic Insects” in Index A. If conditions are unsafe for any reason, including high water or slippery rocks, **DO NOT SAMPLE**.

Why Monitor for Macroinvertebrates

The basic principle behind the study of macroinvertebrates is that some species are more sensitive to pollution than others. Therefore, if a stream site is inhabited by organisms that can tolerate pollution, and the pollution-sensitive organisms are missing, a pollution problem is likely.

For example, stonefly nymphs, which are very sensitive to most pollutants, cannot survive if a stream's dissolved oxygen falls below a certain level. If a biosurvey shows that no stoneflies are present in a stream that used to support them, a hypothesis might be that dissolved oxygen has fallen to a point that keeps stoneflies from reproducing or has killed them outright.

This brings up both the advantage and disadvantage of the biosurvey. The advantage of the biosurvey is it tells us very clearly when the stream ecosystem is impaired, or "sick," due to pollution or habitat loss. It is not difficult to realize that a stream full of many kinds of crawling and swimming "critters" is healthier than one without much life. Different macros occupy different ecological niches within the aquatic environment, so diversity of species generally means a healthy, balanced ecosystem. The disadvantage of the biosurvey, on the other hand, is it cannot definitively tell us why certain types of creatures are present or absent.

In this case, the absence of stoneflies might indeed be due to low dissolved oxygen. But is the stream under-oxygenated because it flows too sluggishly, or because pollutants in the stream are damaging water quality by using up the oxygen? The absence of stoneflies might also be due to other pollutants discharged by factories or run off from farmland, water temperatures that are too high, habitat degradation such as excess sand or silt on the stream bottom has ruined stonefly sheltering areas, or other conditions. Thus a biosurvey should be accompanied by an assessment of *habitat and water quality* conditions in order to help explain biosurvey results.

APPENDIX D

POST-LAB QUIZ EXAMPLE

Biological Monitoring
ID #   same as you have used before , with class period in

front.__________________________

1.   Why are aquatic macroinvertebrates good indicators of stream quality? (list 2 reasons)

2. Most of the biological markers are in what phase of their life cycle?

3. Name one pollution sensitive (class 1) organism.

4. Name one pollution tolerant (class 3) organism.

5. Identify this organism using the key. Record the box sequence you follow.

Box 1 → then → then →

Add more as needed..

6. What qualities might affect biological markers that aren't measured on a chemical test?

* People who participate in what type of recreational human activity would find the information about macroinvertebrates to be of value to them?

APPENDIX E

TREATMENT DETAILED DATA

	# on quiz/ **lab #**	period 1 mean	per. 1 mode/ type	period 2 mean	per. 2 mode/ type	period 5 mean	per. 5 mode/ type	period 6 mean	per. 6 mode/ type	All data mean
# quiz	10	3.3	2, 4	3.5	3, 4	3.2	3, 4	3.2	3	3.13
N =	**1**	23	D	27	P	26	A	24	N	
# quiz	9	3.9	3	3.6	4	2.8	2	4.5	3	3.57
N =	**2**	22	A	26	D	21	N	26	P	
# quiz	9	3.0	2, 4	3.9	1, 5	4.7	3, 5	4.9	5	4.21
N =	**3**	25	N	22	A	27	P	25	D	
# quiz	7	5.7	7	5.5	7	5.4	6	5.3	6	6.08
N =	**4**	26	P	28	D	25	A	22	N	
# quiz	10	5.7	7	7.3	9	6.7	4	5.4	5	6.08
N =	**5**	26	N	31	A	25	D	16	P	
# quiz	10	5.5	3, 9	3.9	2, 3	5.2	6	2.1	none	4.21
N =	**6**	23	P	27	D	18	A	13	N	
# quiz	11	5.0	5, 6	4.9	5	6.4	5, 7	4.2	3, 4, 5	4.7
N =	**7**	21	D	24	P	21	N	15	A	
# quiz	9	5.5	4	4.3	5	4.0	4	4.5	7	4.56
N =	**8**	22	A	21	N	25	D	20	P	
# quiz	10	4.5	2, 6	7.5	8	5.4	6	8.0	7	6.58
N =	**9**	8	D	29	P	17	D	18	A	
# quiz	9	5.0	5, 6	5.1	4, 5	5.7	7, 8	4.2	4	4.99
N =	**10**	20	P	21	N	17	A	18	D	
# quiz	10	2.6	4	3.3	1	4.5	4	3.7	4	3.66

	# on quiz/ **lab #**	period 1 mean	per. 1 mode/ type	period 2 mean	per. 2 mode/ type	period 5 mean	per. 5 mode/ type	period 6 mean	per. 6 mode/ type	All data mean
N =	**11**	12	N	14	A	17	P	10	D	
# quiz	11	4.8	5,	4.8	6	5.4	5	6.7	7	5.23
N =	**12**	13	A	25	N	21	D	19	P	

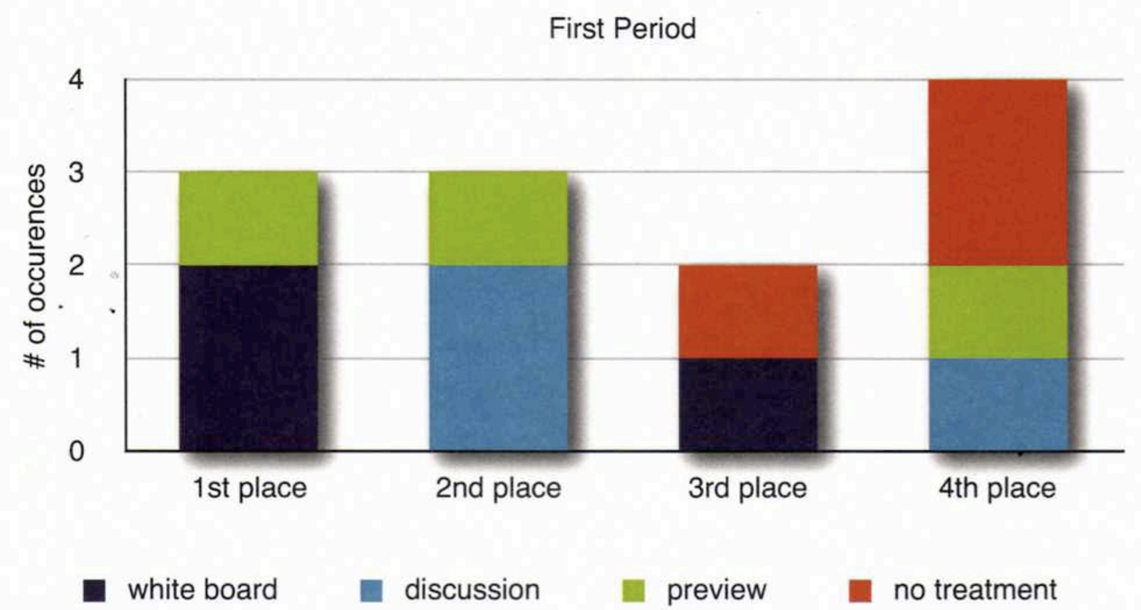
Table summarizes which class periods did whiteboard discussion (P), class discussion (D), preview exercise (A), or no treatment (N) for each of the twelve labs.
Table also includes the class mean and mode(s) for the post lab assessment quiz.
Right hand column is the mean for all students from the entire 4 classes

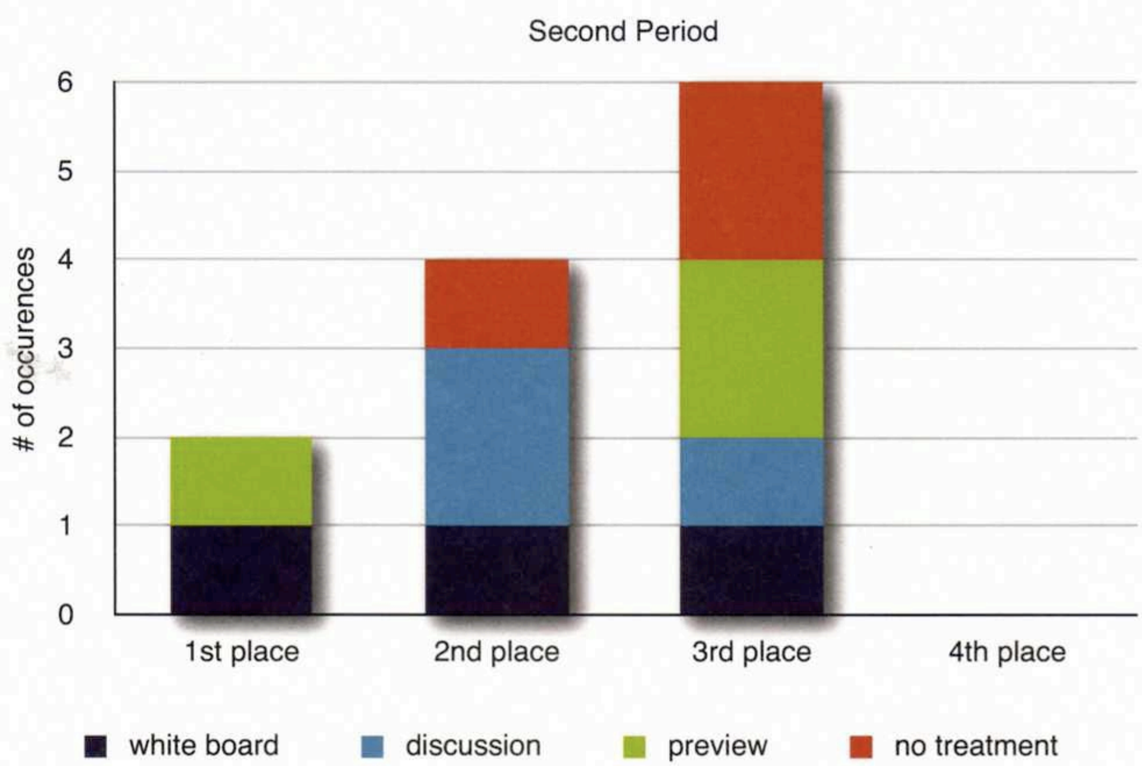
APPENDIX F

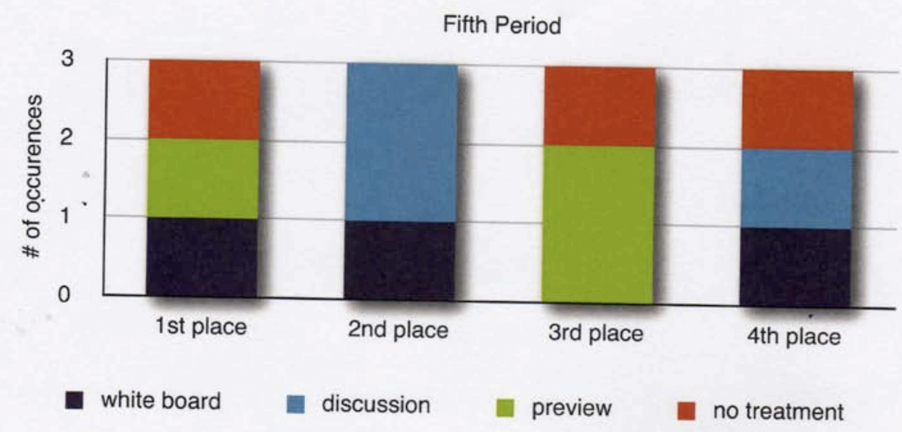
BREAKDOWN OF RESULTS BY CLASS

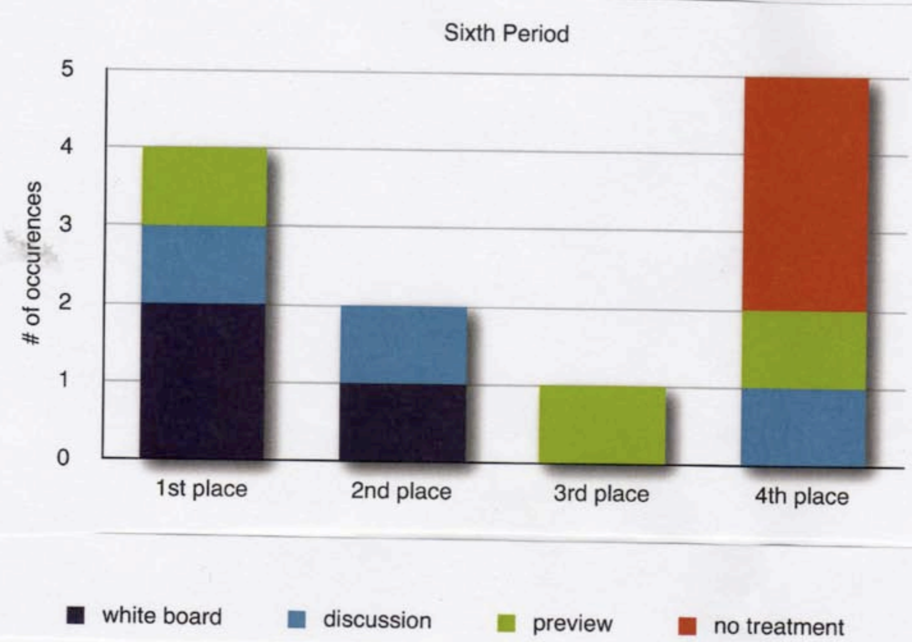
Appendix F

Rankings of all treatments for each class.






Fifth Period
# of occurences
0
1
2
3
1st place
2nd place
3rd place
4th place
white board
discussion
preview
no treatment


Sixth Period
# of occurences
0
1
2
3
4
5
1st place
2nd place
3rd place
4th place
white board
discussion
preview
no treatment

APPENDIX G

COMPARISON OF GROUP MEAN TO TREATMENT MEAN

APPENDIX G

Comparison of mean quiz score for all students with mean quiz score by treatment type

Treat-ment #	Entire sample mean (*N*=)	Whiteboard mean (class period)	Discussion mean (class period)	Preview mean (class period)	No discussion mean (class period)
1	3.1 (*N*= 100)	3.5 (2)	3.3 (1)	3.2 (5)	3.2 (6)
2	3.6 (*N*= 95)	4.5 (6)	3.6 (2)	3.9 (1)	2.8 (5)
3	4.2 (*N*= 99)	4.7 (5)	4.9 (6)	3.9 (2)	3.0 (1)
4	6.1 (*N*= 99)	5.7 (4)	5.5 (2)	5.4 (5)	5.3 (6)
5	6.1 (*N*= 99)	5.4 (6)	6.7 (5)	7.3 (2)	5.7 (1)
6	4.2 (*N*= 81)	5.5 (1)	3.9 (2)	5.2 (5)	2.1 (6)
7	4.7 (*N* = 81)	4.9 (2)	5.0 (1)	4.2 (6)	6.4 (5)
8	4.6 (*N* = 88)	4.6 (6)	4.0 (5)	5.5 (1)	4.3 (2)
9	6.6 (*N*= 66)	7.5 (2)	4.5 (1)	8.0 (6)	5.4 (5)
10	5.0 (*N*= 76)	5.0 (1)	4.2 (6)	5.7 (5)	5.1 (2)
11	3.7 (*N*= 53)	4.5 (5)	3.7 (6)	3.3 (2)	2.6 (1)
12	5.2 (*N* = 78)	6.7 (6)	5.4 (5)	4.8 (1)	4.8 (2)

APPENDIX H

POST-TREATMENT SURVEY

Follow up survey dealing with lab discussions.

Number as you have used before ______________________________________

Please complete the questions honestly with as many supporting details as you can. Your answers will not affect your grade.. While complete sentences are not necessary, the more details you can provide, the better. For questions involving the choice of an opinion, please write the number of your response in the left margin.

Duirng the last part of the fall semester and first part of the spring semester, you participated in an experiment involving four styles of lab discussion. In one style, you worked in a group to prepare answers for a white board presentation. For another style the lab was discussed in class. A third style involved you reading some information before starting the lab. The last variant involved no treatment, that is no discussion nor advance readings.

For each method, what do you think of the strategy as a way of helping you learn material.

Please use the following scale and write the number of your choice in the left margin.
Very helpful (1) helpful (2) unhelpful (3) very unhelpful (4)

1. How helpful were preparing and watching white board presentations?

2. How helpful was discussing the lab as a class after the lab?

3. How helpful was doing advanced reading explaining the lab before the lab?

4. How helpful was doing the lab but not discussing it anyway before taking a lab quiz?

The study involved two styles of post lab discussion, whiteboards and group discussion. In terms of helping you learn, circle the statement that applies best to you.

Making and watching whiteboard presentations were more helpful than doing just a class discussion.
Or

Conducting a class discussion was more useful than making and watching whiteboard presenations.
Or

Whiteboard presentations and discussions were equally effective.

For all the following, please say as many statements that apply and provide reasons for why you say what you do.

When using the whiteboards, what did you like about them?

What was helpful about using the whiteboards?

What was the biggest drawback to using the whiteboards?

With regards to class discussion, what did you like about them?

With regards to class discussion, what was helpful about the class discussion?

What was the biggest drawback to having a class discussion?

Which Environmental Science labs from this year did you enjoy the most? Try to include supporting reasons.

If you were to include more labs in this environmental science course, what would you like to do?

On what topics would you like to spend more lab time?

If the course included a longer term lab done either with a group or individually, do you think this would improve your learning?

How could lab activities be more helpful to your learning in AP Environmental Science?

Is this anything else you would like to say about AP Environmental Lab experiences?

APPENDIX I

INDIVIDUAL STUDENT DETAILED DATA

ID #

ID #	per	#	> SP	> SD	Disc. > Avg./ WB	Disc. > 2x avg.	Disc. < avg.	No disc. > avg.	No Disc.> 2x avg.	No disc. < avg.	Water pre-test	Water post test
333	1	4	2	2	0		5	3		3	4	8
3139	1	4	4 >	2	1		3	4		2	1	10
8591	1	3	1	1	3 / 1	1	5	5	1	1	5	10
1209	2	4	2	1	3 / 1		3	4		1	7	10
1212	2	3	2	1	3 / 1	1	1	3	1	1	7	12
3584	2	3	2	1	3 / 1		3	4	2	1	3	11
5816	2	4	2	1	3 / 2	1	1	3	1	1	X	X
7461	2	4	3	1	4 / 2	1	1	2	1	4	2	11
203...	2	3	na	na	1		3	1		1	1	10
418...	2	3	na	na	2 / 1	1	3	4	1	1	8	10
417	5	3	3	1	5 / 2	2 wb*	0	4		1	6	10
714	5	4	3 =	2	2 / 2		2	1		4	1	7
18531	5	3	na	na	1		1	2		1	0	6
324...	5	3	2	2	0		1	2		1	X	X
1825	6	4	2 >	2	5 / 3		1	1		2	6	5
6707	6	3	2 =	1	4 / 2	1 wb*	2	2		2	4	5
8558	6	3	na	na	3 / 3		3	0		1	5	9
765...	6	4	2	1	5 / 3	1 wb*	0	4	1	0	4	9

Notes on Table

Disc. > Avg./ wb means number of times the student earned a score higher than the average when doing whiteboard or class discussion. Number after slash refers to number for the whiteboard trial.

Disc. > 2x avg. means number of time the quiz score at least twice the average.

* means the score that was at least twice the average obtained on a white board treatment.

in column 3 refers to unhelpful (3) or very unhelpful (4) answer to survey question.

Columns 4 and 5 refer to post treatment survey and helpfulness of whiteboard (P) and discussion (D). 1 is very helpful, 2 helpful, 3, unhelpful, 4 very unhelpful. na -- no survey from this person. > ranked whiteboard as more helpful than discussion in later question, = said whiteboard and discussion equally helpful.

Filled background of identifying number silver indicates that data suggest discussion was helpful. Honeydew indicates students who on final survey indicated that post lab discussions were very helpful.