The Greenhouse Effect: Common Misconceptions and Effective Instruction

By

Trevor W Strickland

Physics Department
California Polytechnic State University
San Luis Obispo, California
12/2009

Table of Contents

Section	Page
Acknowledgements	2
Introduction/Background	3
Methods	7
Results	20
Assumptions	51
Summary	57
Appendix	62
Bibliography	78

Acknowledgements

I'd like to thank the professors of California Polytechnic State University in San Luis Obispo for all of their instruction over the years. The knowledge they have shared and methods they have taught me to practice have given me a new perspective on the world. I'd also like to thank my friends and family for their steadfast support and confidence through times when my goals seemed insurmountable. To the instructors of the classes in this study: your cooperation and willingness to be a part of this research is greatly appreciated. To the students of Cal Poly and the University of Arizona, I'd like to express my thanks for your time and shared knowledge of the greenhouse effect. Your survey answers and participation has contributed to the betterment of future education. Finally, and most importantly, I'd like to thank Dr. John Keller for allowing me to share in his continued educational research. His time, efforts, and mentorship in my education and through this research have been invaluable to me. He has significantly contributed to the scientist, educator, and person that I am and will become.

Introduction/Background

As a Physics Bachelor of Arts major at Cal Poly, I had initially set out to explore this specific science and the universe that it strives to characterize. The phenomena described in my classes fed my curiosity and left me fascinated and desiring to share my excitement for the subject. This led me to a tutoring position in the Math and Science Learning Center at Cal Poly which birthed my interest in the teaching profession. In the Spring of 2007 and later, in the Fall of 2008, I studied astronomy under the instruction of Dr. John Keller. As it so happened, one of Dr. Keller's many roles at Cal Poly was to recruit science majors for careers in teaching. Through discussions with Dr. Keller about science teaching I learned of an educational research project he had been working on in recent years. Dr. Keller was kind enough to bring me on board with his research and mentor me through the investigation of a part of this research for my senior project at Cal Poly.

The educational research project that I engaged in with Dr. Keller involved identifying common student misconceptions regarding the greenhouse effect and investigating instructional methods of teaching. Over a period of time in which Dr. Keller was a graduate student at the University of Arizona, he developed a test designed to measure a student's understanding of the greenhouse effect. This test

is known as the Greenhouse Effect Concept Inventory (GECI). A central goal of this project is to investigate the effectiveness of this concept inventory at measuring differences in student knowledge of the greenhouse effect. In addition to the Cal Poly student population that participated in this research, we will be looking at the population of students at the University of Arizona to which Dr. Keller first administered the GECI in 2006. By comparing these populations we hope to identify differences in the pre-instruction populations and speculate as to the cause of the differences in pre-instruction GECI scores.

My goals in this project are to practice professional scientific research, develop a better understanding of effective instructional methods, and help to increase awareness and a proper understanding of the greenhouse effect. In my years at Cal Poly, I have studied physics and the scientific method. My hope is that this project will provide an opportunity to implement the ideas I have learned about and practiced in the classroom and that from this I will grow as a scientist. As a prospective teacher, it will be highly valuable to investigate and take note of the instructional methods which produce gains in knowledge. These methods will contribute to my personal development as a scientist with a duty to educate, whether the forum is a classroom or elsewhere. Finally, I believe a strong understanding and knowledge of the greenhouse effect is vital to the preservation of our Earth and our pursuit of sustainability. We must effectively educate future scientists and laypersons alike if we hope to have an impact on the way we as

humans understand and affect the environment in which we live. My desire is to contribute to this education and, though indirectly, help provide a comprehensive knowledge of the greenhouse effect that will allow people to make informed decisions about their actions toward, and contributions to, Earth.

The Greenhouse Effect Concept Inventory (GECI) was developed by Dr. John Keller at the University of Arizona in order to identify common misconceptions regarding the greenhouse effect and to measure changes in understanding of this concept through instruction. This was initiated through an initial process of collecting open-ended student supplied responses regarding the greenhouse effect. The current GECI is a product of the refinement of previous versions containing items of multiple choice as well as multiple choice along with explanation of reasoning. The survey now consists of 25 multiple choice items: 20 items pertaining to the Greenhouse Effect with 5 response choices each and 5 items which serve to report upon background information of the person surveyed. The GECI was constructed with this number of items in hopes that it could be completed within a 15 minute time period.

In order to most effectively highlight inaccurate student perceptions and areas of instruction where the greenhouse effect is not properly being addressed, the concept inventory has been assembled with as much care toward the key concepts of the greenhouse effect as allowed for in a 20 item, 15 minute survey.

Concepts covered within the GECI include types of greenhouse gases (3 items), types of electromagnetic energy associated with the Sun, atmosphere, and surface (7 items), energy equilibrium and balance (4 items), greenhouse effect mechanisms (3 items), and comparison of global warming to the greenhouse effect (3 items). For the purposes of this study it is believed that knowledge of each of these fundamental ideas is most indicative of an understanding of the greenhouse effect.

A final version of the GECI was administered to six undergraduate classes at the University of Arizona during the Spring of 2006. Later, in the Spring of 2008, the survey was also completed by a group of 6 undergraduate classes at the California Polytechnic State University in San Luis Obispo. Using data from these two populations, the following three research questions will be investigated: 1) Is the GECI sensitive to differences in student knowledge of the greenhouse effect after instruction; 2) Is there any difference in post-course mean GECI score based upon the instructional method used, and 3) Do the two populations (Arizona and Cal Poly students) show similar pre-instruction understanding of the greenhouse effect?

Methods

Due to the nature of the specific research conducted, it was of great importance to select science classes in which greenhouse effect discussions would be relevant to the subject matter in which the students were instructed. In addition, it was necessary to include control groups of classes that would not be instructed in the greenhouse effect. This would allow us to compare the gains of classes that were instructed in the greenhouse effect to those classes that were not in order to observe any difference resulting from the specific method of instruction. Both the degree to which the greenhouse effect was covered and the method of instruction varied from class to class.

University of Arizona Class Descriptions

In research conducted during Spring 2006 at the University of Arizona, Dr. John Keller developed the GECI through results gathered by administration to six classes. The details and results of this research can be found in his dissertation "Development of a Concept Inventory Addressing Students' Beliefs and Reasoning Difficulties Regarding the Greenhouse Effect" (Keller, 2006). The following is a summary of the six classes selected to be included in the study at Arizona.

At the University of Arizona, six introductory science classes for non-science majors were selected for research. The six classes surveyed were as follows: two astronomy classes covering identical material in the course of the semester, two atmospheric science classes again which covered identical course material, and two individual planetary science courses which each covered unique content.

A lecture tutorial activity was developed for, and tested in, three of these studied classes. The lecture tutorial activity is an exercise in which students gathered into small discussion groups and worked through a series of greenhouse effect worksheet questions as a team. The activity is designed to teach the greenhouse effect, contrast the greenhouse effect against global warming, and address common misconceptions about the greenhouse effect. Two different methods of delivering the greenhouse effect material were used for the purpose of studying whether either method was more effective at conveying the information and addressing areas of misconception.

The two astronomy courses each met twice a week for 75 minutes per class meeting. The first of these class sections participated in the lecture tutorial activity. This lecture tutorial activity involved a 22 minute lecture on the greenhouse effect, followed by a 22 minutes activity completed in small discussion groups. The class then finished with a 7 minutes debriefing and an 8 minute comparison of the greenhouse effect to global warming. The other astronomy class section was

delivered the lecture tutorial narrative. This narrative included no student discussion and involved the instructor reading each of the questions in the lecture tutorial and the corresponding correct answer while also addressing common misconceptions. They were given a 27 minute lecture on the greenhouse effect followed by a 16 minute lecture tutorial narrative which covered the same concepts that were discussed in the other section's small group activity. This second section also finished with a 9 minute comparison of the greenhouse effect to global warming. In total, the first astronomy section spent 59 minutes between instruction and activity, while the second section spent a similar 52 minutes in exclusively lecture. This would later provide for an analysis comparing two very similar classes: one taught by lecture with the supplement of an activity, the other taught purely via lecture. Four other classes were involved in the study at the University of Arizona, however, those classes are not referenced in this study.

Cal Poly Class Descriptions

In the research conducted at California Polytechnic State University in San Luis Obispo during the spring of 2008, six classes were chosen to be given the GECI. Each of these classes was taught by its own individual instructor. The GECI was given once at the beginning of the quarter-long course and once at the end. Five of the six classes were science classes which covered the greenhouse effect or related material to varying extents while one class was an anthropology class used as a

control group. Each of these classes described below was used to gauge the effectiveness of the GECI as an instrument for measuring gains in understanding from instruction on the greenhouse effect.

The first class involved was an introduction to earth science course. This class will be referred to as Class 1 in this study. This class spent one class period, for a total of two hours, covering greenhouse effect material. The information was delivered through a lecture using a combination of slides (some developed by the class's instructor and the others developed and provided by the instructor of the fourth Cal Poly class described below). Following this lecture, the Lecture Tutorial Activity developed at the University of Arizona was used in this class. This Lecture Tutorial Activity was incorporated within the same class period. Topics covered in class focused on sources of heat, atmospheric warming of Earth's surface, carbon dioxide and methane as greenhouse gases, and geological evidence that the greenhouse effect is occurring and increasing. In this treatment group there was no homework or reading assigned that related to the greenhouse effect. Out of 51 questions, the final exam included 5 multiple choice questions on the greenhouse effect. The final exam, however, was administered about a week after the post-instruction GECI. This being the case, students were likely not highly motivated to study for the final examination greenhouse effect questions before taking the post-instruction GECI.

The second course was a 100 level astronomy course. This course specifically focused on the solar system. This class will be referred to Class 2 for the study. Through the course of the quarter, the class covered very little material involving the greenhouse effect. In discussion of the atmosphere of the planet Venus, the instructor contrasted the atmosphere of Earth with that of Venus. It was mentioned that Earth's atmosphere produces a phenomena known as the greenhouse effect. Almost nothing directly relating to the greenhouse effect beyond this mention was discussed in class.

An oceanography class was also included in the study at Cal Poly. This class will be referred to as Class 3 in this study. In this class 15 to 20 minutes was spent during one lecture to address the greenhouse effect. The instructor spent time lecturing on the composition and structure of Earth's atmosphere. Overheads, cartoons, diagrams, and simple pictures were used to describe and explain the processes and causes of the greenhouse effect. The instructor made a point to mention that water vapor is a large contributor to the greenhouse effect and that we have little control over this. It was also mentioned that the greenhouse effect is more pronounced at tropical latitudes than polar latitudes because of air moisture. Homework was given which required students to recall the gas composition of the atmosphere as well as which of these gases are involved in the greenhouse effect and which are not. Students in this class were tested on the greenhouse effect and related topics and were given testing review sheets that mentioned they would need

to understand the greenhouse effect. They were expected to understand the cause of the greenhouse effect, that carbon dioxide is a greenhouse gas that we have control over, and that ozone is not related to the greenhouse effect. About one tenth of the weekly quizzes covered the greenhouse effect and one to two percent of the questions on the course final were closely or directly related to the greenhouse effect. It is likely that the weekly quizzes were at least a small motivating factor from studying the greenhouse effect material. Since the final exam took place about a week after the post-instruction GECI administration, however, it is likely that this did not highly motivate students to study the greenhouse effect material.

The fourth class in the Cal Poly study was a 100 level astronomy course that primarily focused on stars and galaxies. This class will be referred to as Class 4 in this study. The class spent approximately two hours covering material relating directly to the greenhouse effect. Two separate one hour lectures were spent on the solar cycle and greenhouse effect where the sun's intensity cycle and the idea of variability were discussed. Also included in these lectures was an animation of the greenhouse effect. In the second 50 minute time period the students engaged in the Lecture Tutorial Activity that the students at the University of Arizona had been administered. In this 50 minute session, students divided themselves into groups of three and completed the tutorial handout through discussion and collaboration. After completing the tutorial the class was given a 15 minute debriefing which

included a summary of the tutorial. There was no homework or reading assigned which involved the greenhouse effect, however, the classes were expected to know the key concepts for the course midterm and final. In the second midterm students were asked in an open-ended question to explain the greenhouse effect and how it works. This midterm also included three to five multiple choice questions which covered the characteristics of greenhouse gases, identification of aspects of the natural greenhouse effect, and blackbody curves. The course final included a question which asked students to identify the main type of energy given off by the sun. Students were expected to learn and understand radiative balance, greenhouse gases and their behaviors, solar and earth spectrums, absorption and reemission of radiation, and the difference between the greenhouse effect and global warming.

The final science course included in the study at Cal Poly was a 100 level physical science course whose subject material was centered on Earth science for liberal studies majors. This class will be referred to as Class 5 in this study. Students from within this course were split into two supplemental lab sections in addition to the lecture portion of the class. About 25 minutes of one class period was spent in lecture on the greenhouse effect. The instructor made mention of the misconception that the effect itself is harmful and highlighted the fact that it is simply an increase in the effect that causes concern. The greenhouse effect was defined and common greenhouse gases were identified. There was specific mention

of carbon dioxide levels from humans and the effect of water vapor. The course instructor noted that positive feedback of more heat is dangerous because is causes more infrared absorption. There may have been some small mention of the greenhouse effect and related experiments in the lab sections of this class. There was minimal, if any, homework and reading assigned by the instructor to students. One or perhaps two questions on the final directly addressed the greenhouse effect and students were aware that they were expected to know what the greenhouse effect is, what positive feedback is, and the names of common greenhouse gases.

The sixth course in the Cal Poly group of classes was an anthropology class. This class will be referred to as Class 6 for this study. This course involved absolutely no subject material which directly or indirectly related to the greenhouse effect. This course was used as a control group. It was predicted that this class would not show significant gains over the course of the quarter.

Cal Poly Class Description Table

Class	Number of Students	Instructor	Method of Classroom Instruction	Length of Classroom Instruction	Out of Class Instruction
Class 1 Introduction to Earth Science	46	Instructor 5	Lecture + Lecture Tutorial Activity	120 minutes	None
Class 2 Astronomy: Solar System	39	Instructor 6	None	0 minutes	None
Class 3 Oceanography	23	Instructor 7	Lecture	15-20 minutes	Homework Only
Class 4 Astronomy: Stars and Galaxies	58	Instructor 8	Lecture + Lecture Tutorial Activity	2 x 50 minutes	None
Class 5 Physical Science	32	Instructor 9	Lecture	25 minutes	Minimal Homework & Reading
Class 6 Anthropology (Control Group)	39	Instructor 10	None	0 minutes	None

The GECI was designed with five demographic questions at the end in order to better characterize the population for any further study. In each of the above courses, it was known which students were involved in any other course which covered the greenhouse effect to any extent. This allowed us to filter out these students from the data to be analyzed. Isolated effects of the lecture or Lecture Tutorial Activity could not be known for these students as they were subject to

known external influence. Other than inquiring into whether students were involved in courses that covered the greenhouse effect during the course of the study, none of the other demographic questions were used in this specific study.

Data Cleaning

Scantron data from the six Cal Poly classes were scanned into Excel files prior to my involvement in this study. Upon my involvement, I began the cleaning process by spot checking the data in these files with a random set of corresponding scantron forms. This allowed me to make sure that the survey hard copies were aligned in order with the listing order in the excel files of the surveys and check the reliability of the scoring done by the scantron machine.

Next I cleaned the Excel files themselves. Some of the survey hard copies contained answers that had been changed. When students decided to change their bubbled-in answer on a scantron but did not erase their original answer bubble adequately, this would be transposed to the hard copy files in excel as an asterisk. So this asterisk would signify that the scantron was interpreted as being marked twice. Each time I found an asterisk in the excel files I would sort through the ordered hard copy files in order to find the scantron corresponding to the excel row that contained the asterisk. I would then correct the excel file based on the more obviously marked answer from the scantron. Following this, students' pre and post surveys had to be matched, as this was essential for studying gains in understanding.

Students who did not submit both a pre- and post- survey were eliminated from the data set.

After cleaning and matching the Excel files I merged the dataset so that all of the surveyed classes from Cal Poly were in one Excel file. I added a class number to each survey, an ID within that class, and a survey tag to indicate whether the survey was taken before or after class instruction on the Greenhouse Effect. I used PRE and POST as indicators. In my final excel file I put both the PRE and POST survey data for each individual class and student ID on a single line and labeled the columns of questions as PRE or POST along with the question number. The class student ID is a three digit number made up of one digit for the specific class the survey came from followed by two more digits indicating the form number within that class (for example a class student ID of 325 would be the 25^{th} form from class number 3). The class student ID column allows me to trace back a survey to the original files in Excel in order to track down a specific form should the need arise. Students who indicated in their survey that they had taken any additional courses which covered the greenhouse effect during the academic quarter that this study was taking place were removed from the sample before the results were calculated. This helped to limit outside instructional influence and better ensure that student GECI scores were a result of instructional methods used.

Finally, I imported this cleaned and completed Excel file into a statistics program called Statistical Package for the Social Sciences (SPSS) for statistical analysis.

Scoring

In SPSS I was able to score each individual PRE and POST survey. I created new PRE and POST answer columns which correspond to a specific numbered question on the survey. Using the "Transform Variable" function in SPSS, each cell corresponding to a specific PRE or POST survey question was recorded as a value of 1 for a correct answer or a value of 0 for an incorrect answer in a new column. Next I created a column to sum the total number of correct answers from the PRE survey for that student and another column to do the same for the POST survey. Using these scores for each student's PRE and POST survey, I created two more columns to calculate the percentage correct from the PRE and POST surveys for each student. Forming one more column, I calculated the normalized gain in order to gauge the improvement of each student. The normalized gain is designed to report the improvement of a student based upon how difficult it is to improve on their original pre-instruction score. The more questions they answer correctly the first time through the survey the more difficult it is to get additional questions correct the second time taking the survey. I calculated this normalized gain as a proportion by taking the difference of the POST percentage correct and the PRE percentage

correct, divided by the result of the PRE percentage correct subtracted from 100 percent. Multiplying this by 100 gave us the normalized gain as a percentage.

$$\text{Normalized Gain} = (\text{Post\%}-\text{Pre\%})/(100-\text{Pre\%})\text{x}100\%$$

Results

This section presents the findings of the greenhouse effect educational study conducted at California Polytechnic State University, San Luis Obispo, during Spring 2008. Three specific questions will be investigated following these results: 1) Is the Greenhouse Effect Concept Inventory sensitive to a difference in student knowledge of the greenhouse effect after instruction; 2) Do students who are engaged in learning activities, students who receive only lecture, and students who are not taught the greenhouse effect have the same mean post-course score; and 3) How do Arizona students from Spring 2006 and Cal Poly Students from Spring 2008 compare in terms of understanding of the greenhouse effect pre-instruction?

The outcome of this analysis could be affected if students in the study were involved in any additional science courses that taught concepts from the greenhouse effect during the quarter in which the survey took place. For this reason, the students who reported that they had taken another course which covered the greenhouse effect during the academic quarter that this study was taking place were removed from the sample before the results were calculated. Students may have also gained knowledge about the greenhouse effect from outside sources such as the media, family and friends, and books. This was not accounted for. Also, some students may have taken more time on each question than other students, causing them to rush through the end of the concept inventory. Surveys missing 3 or more

questions were removed from the sample as it was likely they ran out of time to finish the survey. Some may have taken courses on greenhouse material prior to the pre-course test. Finally, the time spent covering the greenhouse effect and the depth of investigation into that concept varied between classes based on instructor. These are all additional factors which contribute to pre- and post-instruction GECI scores.

Basic Descriptive Statistics

I began my analysis of the pre- and post-instruction scores by computing the basic descriptive statistics for each of the six Cal Poly classes and their score out of 20 on the test before taking their course ("Pre Score") and after taking their course ("Post Score"). These basic descriptive statistics can be found in Table 1 below. The pre-course mean GECI scores are all quite similar ranging from 5.103 to 6.609 out of a possible 20. These scores were statistically analyzed for significance and the results of this analysis are discussed in what follows. Table 1 also shows us that the scores from the post-course GECI administration have changed. The range in mean score for the post-course GECI is 5.769 to 13.379 out of a possible 20. The variances between classes differ as well. We will examine these changes for statistical significance in order to find whether or not the classes in the study began at a similar level of knowledge of the greenhouse effect and how their knowledge may have changed after instruction and based upon the method of instruction.

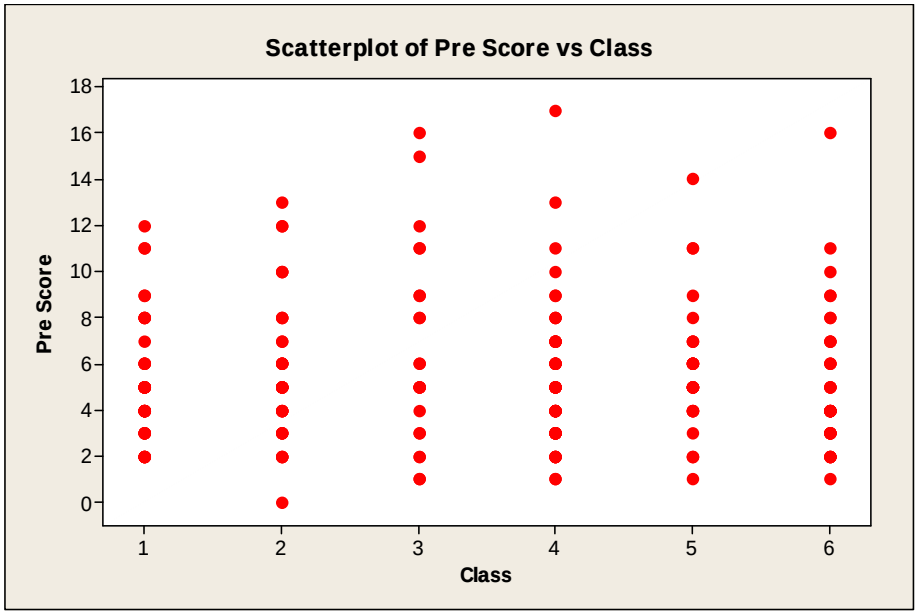
Table I: Descriptive Statistics for Pre Score and Post Score

		TOTAL						
Variable	Class	Count	N	N*	Mean	SE Mean	StDev	Variance
Pre Score	1	46	46	0	5.130	0.369	2.500	6.249
	2	39	39	0	5.641	0.481	3.004	9.026
	3	23	23	0	6.609	0.882	4.229	17.885
	4	58	58	0	5.155	0.386	2.943	8.660
	5	32	32	0	5.625	0.487	2.756	7.597
	6	39	39	0	5.103	0.480	2.998	8.989
Post Score	1	46	46	0	12.870	0.396	2.688	7.227
	2	39	39	0	6.462	0.477	2.981	8.887
	3	23	23	0	8.609	0.829	3.974	15.794
	4	58	58	0	13.379	0.403	3.066	9.397
	5	32	32	0	9.094	0.450	2.545	6.475
	6	39	39	0	5.769	0.560	3.498	12.235
Variable	Class	Minimum	Q1	Median	Q3	Maximum	IQR	
Pre Score	1	2.000	3.750	5.000	6.000	12.000	2.250	
	2	0.000	4.000	5.000	7.000	13.000	3.000	
	3	1.000	3.000	5.000	9.000	16.000	6.000	
	4	1.000	3.000	4.500	7.000	17.000	4.000	
	5	1.000	4.000	5.000	6.750	14.000	2.750	
	6	1.000	3.000	4.000	7.000	16.000	4.000	
Post Score	1	7.000	11.750	13.000	15.000	17.000	3.250	
	2	1.000	4.000	6.000	9.000	13.000	5.000	
	3	2.000	6.000	7.000	11.000	18.000	5.000	
	4	5.000	11.000	13.500	16.000	20.000	5.000	
	5	2.000	8.000	9.000	11.000	15.000	3.000	
	6	0.000	3.000	5.000	7.000	15.000	4.000	

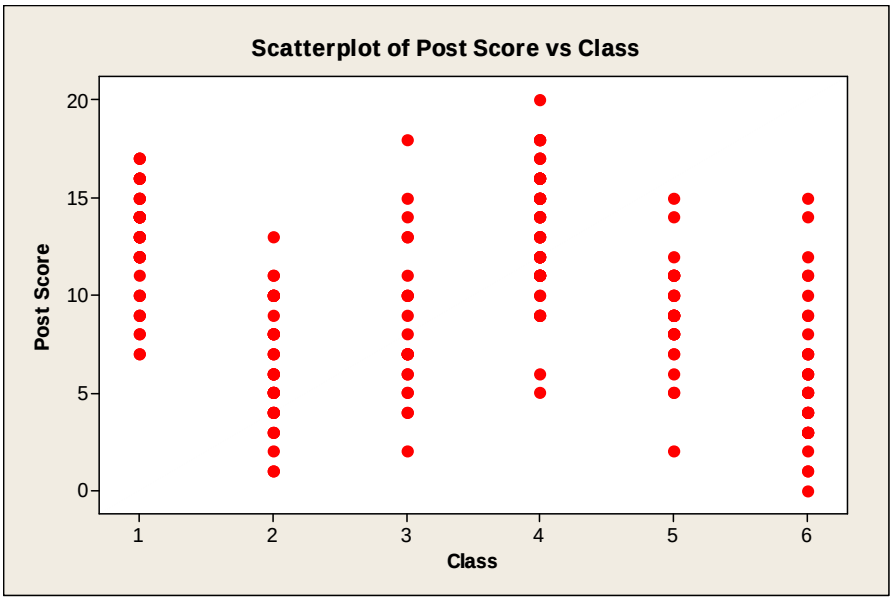
First, we notice that the means of all the class'-pre-instruction-scores are quite similar. Also, we see that the post score means have changed. We will investigate if these changes are statistically significant below. The variances differ in both pre and post scores, and again we will need to test this to see if their differences are significant.

Below we provide scatter plots showing the score data for each class, first for the pre-course scores and then for the post-course scores:

Scatterplot 1



Scatterplot 2



In Scatterplot 1 for the pre-instruction scores we can see that most of the scores are below 11 items out of twenty and no class clearly stands out. In Scatterplot 2, we see much more stratified groups, with some classes displaying scores in a range that is noticeably higher than others.

Pre-Instruction ANOVA

In order to gauge whether the GECI (Greenhouse Effect Concept Inventory) is an effective tool for measuring learning we need to know if we begin with similar sample populations. I used a one-way ANOVA to compare the mean pre-course scores for each of the six Cal Poly classes. I did this in order to find whether or not there was a significant difference in the mean scores between the six classes before instruction in the greenhouse effect. This would allow us to know if there is sufficient evidence to state that at least one class's mean score on the pre-course GECI varied significantly from the rest. In this case we would know that the sample populations are not similar before instruction as this could affect our interpretations of post-course gains significantly.

We expected that before instruction the six classes would have mean scores that are not significantly different since they have not been instructed in the greenhouse effect as far as we can tell. At this point in the study, the mean scores of the control groups should show little or no difference from mean scores of the other four classes.

A 95% confidence interval was used to analyze these results. This means that we can say with 95% confidence that the difference in the mean score between each class when compared with each of the other classes lies within the range given in the analysis. If this range includes zero then we are left with the chance that there is no difference in mean score between classes.

One-way ANOVA: Pre Score versus Class

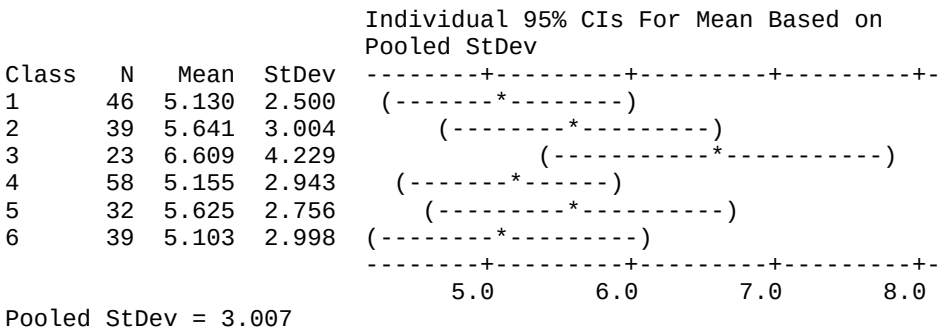
SOURCE	DF	SS	MS	F	P
Class	5	47.59	9.52	1.05	0.387
Error	231	2088.36	9.04		
Total	236	2135.96			

S = 3.007 R-Sq = 2.23% R-Sq(adj) = 0.11%

From the results of our one-way analysis of variance we can see that the p-value we obtain for the class by class comparison for the pre-course GECI scores is 0.387. This is the probability that, given that the null hypothesis is true, a test statistic will be the same or more extreme than that which we received from the sample data. Since our p-value is 0.387, which is larger than our alpha value of 0.05 for a 95% confidence interval, we fail to reject the null hypothesis that the mean scores between classes for the pre-course test are all equal. There is not sufficient evidence to show that at least one mean score for the pre-course scores differs from the rest of the mean scores. Our results from this ANOVA support the claim that our sample populations were similar before taking their respective courses. This is valuable to us because we have populations that do not show sufficient evidence to prove a significant difference in their means and we have students that seem to be

at a similar level of knowledge of the concepts covered in the GECI before instruction.

Pre-Instruction Mean Confidence Intervals



In the Pre-Instruction Mean Confidence Intervals Graph of the 95% confidence intervals for the mean we can see where the individual confidence interval for each class mean lies along the scale from 0 to 20, zero being the lowest score possible on the GECI and 20 being the highest.

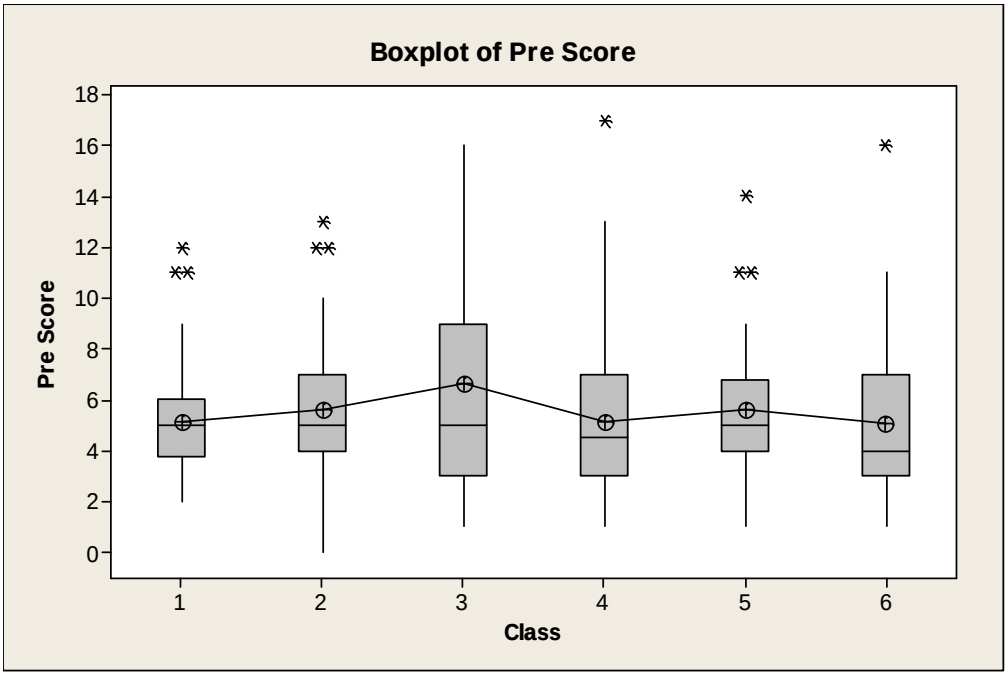
This finding is significant because it supports our hypothesis that the students in our study had a relatively equal level of understanding of the greenhouse effect before they were instructed. No one particular class seems to have a better understanding of the greenhouse effect than another, therefore, we have comparable classes with which to experiment and measure gains in understanding. Any effect that a specific instructional treatment may have on a class in the study when compared with another class that was received a different instructional treatment should be more clearly defined since our sample populations

are all starting at a similar level of understanding. It is also important to note that all of the classes had mean scores in the range of 5 to 6 out of a possible 20. This leaves plenty of room for gains when it comes time for the post-instruction survey.

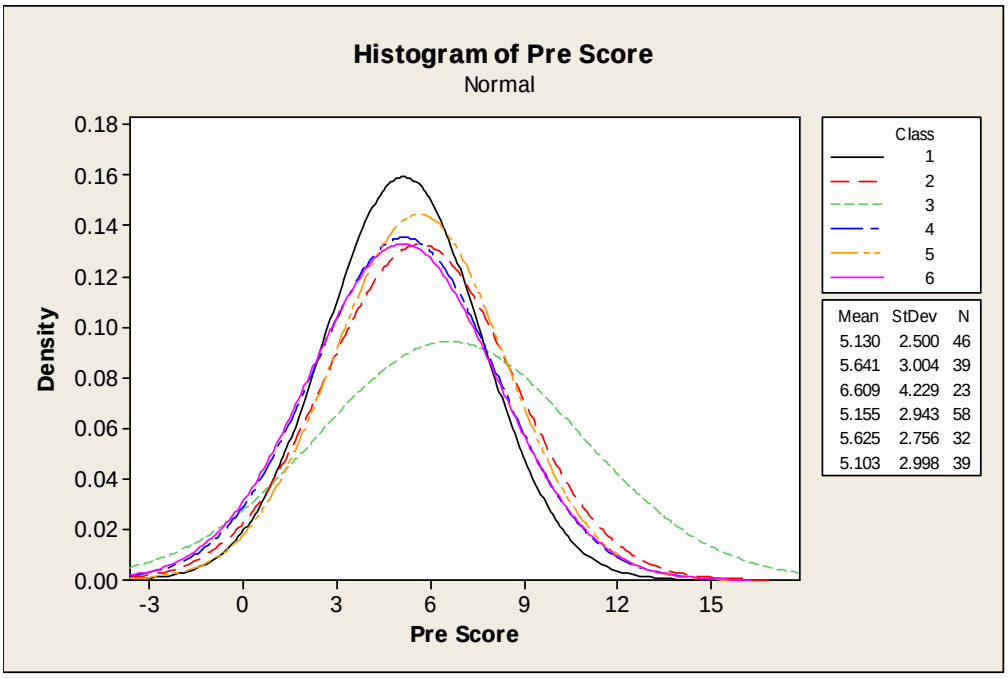
Below we have provided a boxplot which presents the five number summaries of the pre-course scores for each class. The means are represented by the circle inside the box. The medians are indicated by the horizontal line in the center of the box. The highest and lowest values can be found at the top and bottom of the vertical line, respectively (this line is known as the whiskers). The first quartile is indicated by the lower boundary of the box, while the third quartile is indicated by the upper boundary of the box. The asterisks mark the outliers.

Pre-Instruction Mean

Boxplot



Below we have a histogram displaying the normal distribution for each class involved in the Cal Poly study based upon their individual student scores.



We see here that Class 3 had the highest mean and a widespread normal distribution which signifies a wide range of scores. The peak density for Class 3 is lesser than that of the other classes while the density at higher scores is greater than those of the other classes. We can attribute the particularly large standard deviation of Class 3 to the small sample size and the education level of the students. The class was an upper division general education course comprised of students which were further along in their educational coursework and more varied in their educational background. The rest of the classes showed means which were lower than that of Class 3 but more concentrated around 5 to 6 correct survey answers.

Post-Instruction ANOVA

Another One-way Analysis of Variance Test was used to evaluate the post-instruction scores. This ANOVA was performed in order to evaluate whether or not the mean GECI scores of the six classes at Cal Poly differed by a statistically significant amount in the post-instruction administration of the GECI. This would specifically tell us if the six classes in the study had mean post-instruction scores that differed enough that we could say with 95% confidence that the level of understanding of the greenhouse effect differed in a significant way from class to class after instruction.

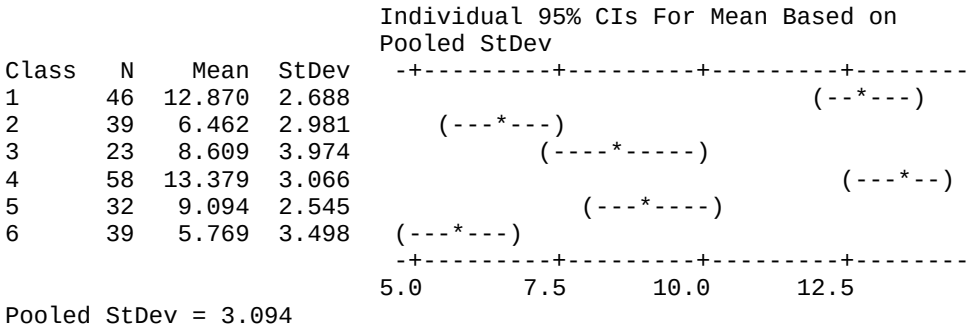
One-way ANOVA: Post Score versus Class

SOURCE	DF	SS	MS	F	P
Class	5	2292.85	458.57	47.90	0.000
Error	231	2211.68	9.57		
Total	236	4504.53			
S = 3.094	R-Sq = 50.90%		R-Sq(adj) = 49.84%		

From our ANOVA we obtained a p-value of 0.000. Because this p-value is less than our alpha value of 0.05 for a 95% confidence interval, we reject the null hypothesis. Our null hypothesis in this case is that all of the post-instruction scores are equal. This means that we can say with 95% confidence that the mean post-instruction GECI score for at least one of the six Cal Poly classes differed from the mean post-instruction score of the other classes by a statistically significant amount. We have therefore found sufficient evidence that the mean post-instruction GECI scores for the six Cal Poly classes are not all equal. This finding is significant in that it tells us that students from at least one class had a better understanding than

students in at least one of the other classes. Since these classes were found to be at a relatively equal level of understanding before instruction, we have reason to look into how the post-instruction scores differed and why.

Post-Instruction Mean Confidence Intervals



In the Post-Instruction Mean Confidence Intervals graph above we see the confidence intervals for the post-instruction mean of each class. There are similarities in the confidence intervals of classes 1 and 4, 3 and 5, and 2 and 6.

Following this one-way analysis of variance, we performed a Tukey Test for pairwise comparisons of classes. In each dataset from the Tukey Test below, we are presented with a 95% simultaneous confidence interval for each class. The Tukey then subtracts the mean of each class, one at a time, from that of the remaining classes in order to see if their 95% confidence intervals include zero, which would represent a chance of the means being equal. The results of this Tukey Test are provided below.

Tukey 95% Simultaneous Confidence Intervals

All Pairwise Comparisons among Levels of Class
Individual confidence level = 99.55%

Post-Instruction Mean Simultaneous Confidence Interval for Class 1

Class = 1 subtracted from:

Class	Lower	Center	Upper	--------+---------+---------+---------+
2	-8.342	-6.408	-4.474	(---*---)
3	-6.529	-4.261	-1.992	(---*----)
4	-1.244	0.510	2.264	(--*---)
5	-5.821	-3.776	-1.731	(---*----)
6	-9.034	-7.100	-5.167	(---*---)
				--------+---------+---------+---------+
				-5.0 0.0 5.0 10.0

Post-Instruction Mean Simultaneous Confidence Interval for Class 2

Class = 2 subtracted from:

Class	Lower	Center	Upper	--------+---------+---------+---------+
3	-0.188	2.147	4.483	(---*----)
4	5.078	6.918	8.757	(---*---)
5	0.513	2.632	4.751	(---*----)
6	-2.704	-0.692	1.319	(---*---)
				--------+---------+---------+---------+
				-5.0 0.0 5.0 10.0

Post-Instruction Mean Simultaneous Confidence Interval for Class 3

Class = 3 subtracted from:

Class	Lower	Center	Upper	--------+---------+---------+---------+
4	2.582	4.771	6.960	(----*---)
5	-1.943	0.485	2.913	(----*----)
6	-5.175	-2.839	-0.504	(---*----)
				--------+---------+---------+---------+
				-5.0 0.0 5.0 10.0

Post-Instruction Mean Simultaneous Confidence Interval for Class 4

Class = 4 subtracted from:

Class	Lower	Center	Upper	--------+---------+---------+---------+
5	-6.242	-4.286	-2.329	(--*---)
6	-9.450	-7.610	-5.771	(---*--)
				--------+---------+---------+---------+
				-5.0 0.0 5.0 10.0

Post-Instruction Mean Simultaneous Confidence Interval for Class 5

Class = 5 subtracted from:

Class	Lower	Center	Upper	--------+---------+---------+---------+
6	-5.443	-3.325	-1.206	(---*----)
				--------+---------+---------+---------+
				-5.0 0.0 5.0 10.0

The Post-Instruction Mean Simultaneous Confidence Interval graphs above show the difference for the mean of one individual class population compared to the mean and confidence interval of the rest of the class populations in our sample. When Class 1 was subtracted from the other classes, we found that only the confidence interval of Class 4 included zero. This means that only for Class 4 can we say that there is not sufficient evidence to show a significant difference in the post-instruction mean scores of Class 1 and Class 4. All other classes, when compared with Class 1 showed a statistically significant difference in post-instruction mean score. More specifically, the confidence intervals for the classes, other than Class 4, were lower than that of Class 4. These other classes had a negative value when the confidence interval for Class 1 was subtracted from that of classes 2, 3, 5, and 6. Therefore, Class 1 and Class 4 had a higher mean post-instruction score in order to produce this result.

When the post-instruction mean score of Class 2 was subtracted from the score of the other classes, only Class 6 showed insufficient evidence to determine that the mean scores differed. Classes 2 and 6 showed confidence intervals which were lesser than those of the other classes. Classes 3, 4, and 5 had positive values when subtracting the confidence interval of class 2. This tells us that classes 3, 4, and 5 had higher mean scores than classes 2 and 6.

Next, as the confidence interval from Class 3 was subtracted from that of the remaining classes, Class 5 was the only class that lacked sufficient evidence to show a difference in score when compared to Class 3 at the 95% confidence level. When the confidence interval for Class 3 was subtracted from that of classes 4, 5, and 6, Class 4 showed a positive interval while Class 6 showed a negative value. Therefore, we conclude that Class 6 had a lower mean than that of classes 3 and 5, while Class 4 had a higher mean than that of classes 3 and 5. Class 4 showed a sufficient evidence to say that its post-instruction mean score differs significantly from that of classes 5 and 6.

Finally, Class 5 also showed sufficient evidence for us to state that the mean score differs in a significant way from that of Class 6. When compared with the confidence interval of Class 5, Class 6 showed a negative interval which tells us that the mean post-instruction GECI score was lesser in the case of Class 6 than that of Class 5.

In summary, the Tukey-Kramer Test shows us that, for the post-instruction mean GECI scores, classes 1 and 4 were similar, classes 2 and 6 were similar, and classes 3 and 5 were similar. In this case, when we say similar it is meant that there is not sufficient evidence to show that the mean scores differ by an amount that is statistically significant.

Classes 1 and 4 were the classes which were instructionally treated with the Lecture Tutorial and Activity which involved discussion groups. The Tukey-Kramer Test for the class by class post-instruction mean scores contributes supporting evidence to the idea that classes 1 and 4 had post-instruction mean GECI scores that were similar and that these scores were greater than those of the other classes at Cal Poly to which they were compared. We have, therefore, been presented with support for the idea that the Lecture Tutorial Activity produced higher gains in student score on the GECI than the instructional methods used for the other Cal Poly classes.

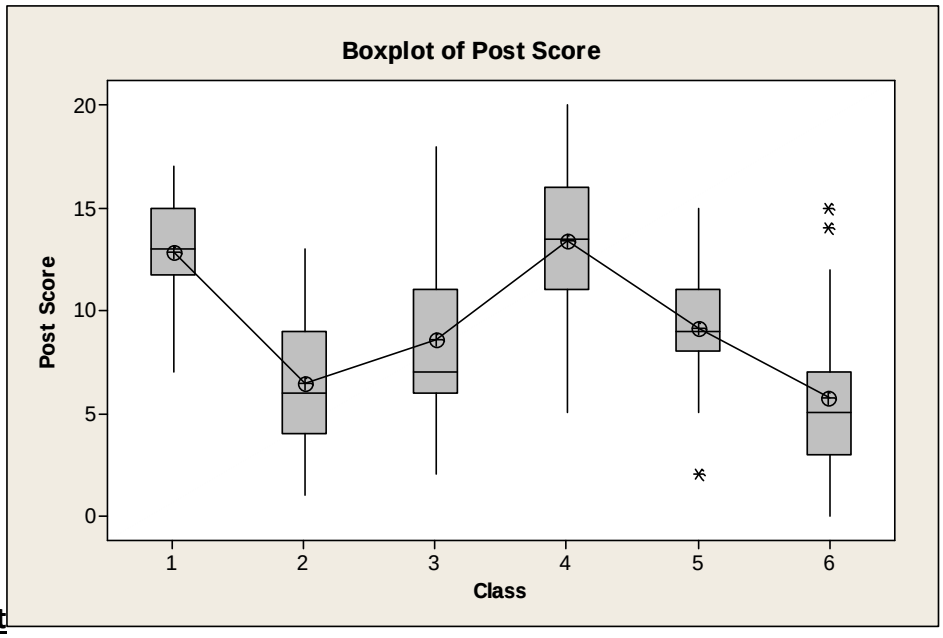
Classes 2 and 6 were the Cal Poly classes that were not delivered any material on the greenhouse effect. Class 6 was used as control group while Class 2 cannot be counted as such since lectures included material related to the greenhouse effect. These classes performed similarly according to the Tukey-Kramer Test. Their scores were lesser than those of all the other classes in the study. This is as we would expect after not having been delivered any material in class to improve their knowledge of the greenhouse effect. Our significant result from classes 2 and 6 is that they did not receive instruction on the greenhouse effect and presumably consequently showed the lowest mean post-instruction GECI scores.

Finally, classes 3 and 5 were the classes in the study which were lectured on the greenhouse effect and which did not engage in any sort of activity or discussion. The Tukey-Kramer Test showed that classes 3 and 5 had confidence intervals for their post-instruction mean GECI score which were higher than those of classes 2 and 6 but lower than those of classes 1 and 4. This suggests to us that lecture was more effective than no instruction but less effective than the Lecture Tutorial Activity.

These results of the ANOVA and Tukey Test also support the idea that the GECI is a successful tool in measuring student learning of the greenhouse effect. These tests show mean GECI scores for classes instructed in the greenhouse effect which were significantly elevated when compared to those classes that received no instruction on the greenhouse effect. We must remember that earlier in testing these classes had scores which were not significantly different from each other before instruction. This suggests that the Greenhouse Effect Concept Inventory is in fact sensitive to a difference in student knowledge after instruction on the greenhouse effect.

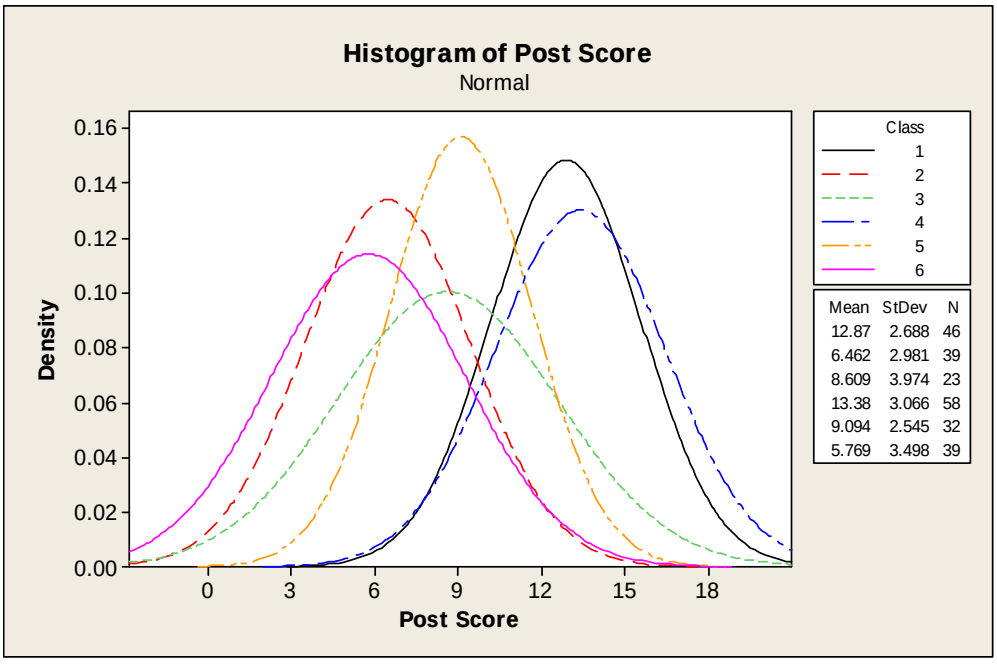
Below we have a boxplot showing the five number summaries of the post-course GECI scores for each class.

Post-Instruction Mean Score



Boxplot

The following histogram shows the normal distribution for the post-course scores of each class. The peaks indicate the mean value while the width of the distribution indicates the concentration of the scores around the mean. A wider distribution indicates more scores spread away from the mean. The height represents the number density of student who are expected have the corresponding score.



Pre-Post Comparison

A table has been constructed below that summarizes the results that we have presented and discussed from the pre- and post-instruction scores of each class from the Cal Poly study. In this Pre- and Post-Instruction GECI Score Summary the mean scores, number of students surveyed, and standard deviation can be found. Also, the normalized gain is now displayed for each class. This gives an idea of how much each class improved in its GECI score from pre- to post-instruction while taking into account the varying difficulty in achieving a higher score than that of the pre-instruction score. This is based upon the number of items correctly answered in the pre-instruction administration which then leaves only a limited

number of items to be answered correctly in order to improve one's score. The normalized gain is calculated as follows:

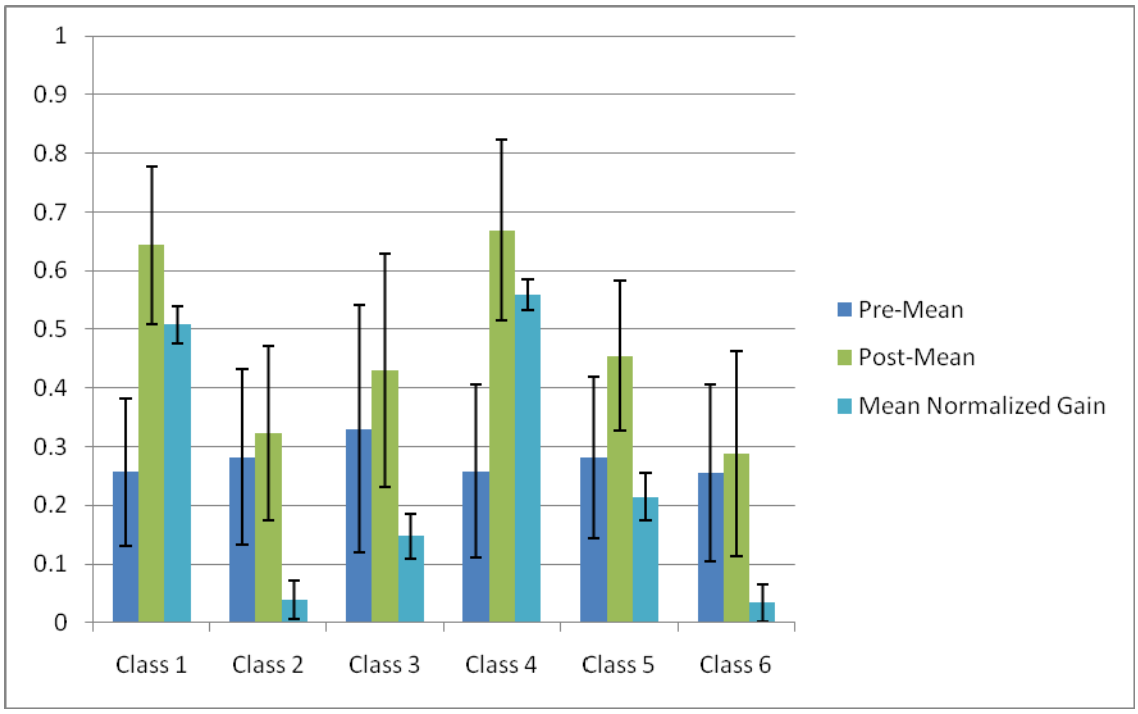
Percent Normalized Gain = [(Post%-Pre%)/(100-Pre%)]x100%.

Table II: Cal Poly Gain By Class

Class	Number of Students	Mean Pre-Percentage	Pre-Instruction Standard Deviation	Mean Post-Percentage	Post-Instruction Standard Deviation	Mean Normalized Gain
4	58	0.2578	0.14714	0.6690	0.15328	55.8548
1	46	0.2565	0.12499	0.6435	0.13442	50.7560
5	32	0.2812	0.13781	0.4547	0.12723	21.4439
3	23	0.3304	0.21146	0.4304	0.19871	14.7059
2	39	0.2821	0.15021	0.3231	0.14905	3.8863
6	39	0.2551	0.14991	0.2885	0.17489	3.3438
Total	237	0.2713	0.15042	0.4924	0.21844	29.0328

The table compares the mean percentage correct in each class before instruction and after instruction. Classes 1 and 4 showed the largest gain, followed by classes 3 and 5, while classes 2 and 6 showed very small gains from pre to post-score. Also shown are the normalized gains calculated from these pre- and post-instruction scores.

Cal Poly Gain By Class



Here we see the pre-instruction and post-instruction means compared along with their corresponding standard deviations. The normalized gain from pre to post is also given. It is clear that Class 1 and Class 4 showed the largest gains. Classes 3 and 5 also showed some moderate gains. Classes 2 and 6 showed very little gain.

In order to further test the effectiveness of the specific instructional method used and the Greenhouse Effect Concept Inventory for its sensitivity to differences in student knowledge after instruction, we perform a paired T-test in hopes of finding sufficient evidence to say whether there were significant changes between the pre-course and post-course scores on a class by class basis. We then look at the

95% confidence intervals of these classes in order to identify whether the mean score for each class increased or decreased.

In table below, which was generated from our paired T-tests, we subtracted the post-course score from the pre-course score for each student in each class. The result of this subtraction was the new sample data for each class. Our null hypothesis for all of these paired T-tests is that the mean of this new data equals zero. If the null hypothesis were true there would be no difference between the pre-course score and the post-course score for that class. If the mean for the class is not zero then we choose the alternative hypothesis and say that we have found sufficient evidence to say that there is a statistically significant difference between the pre-course and post-course mean for that class.

Class	N	Mean	Standard Deviation	95% Confidence Interval	P-Value
Class 1	46	-7.739	3.448	(-8.763, -6.715)	0.000
Class 2	39	-0.821	2.480	(-1.624, -0.017)	0.046
Class 3	23	-2.000	2.045	(-2.884, -1.116)	0.000
Class 4	58	-8.224	3.021	(-9.018, -7.430)	0.000
Class 5	32	-3.469	3.005	(-4.552, -2.385)	0.000
Class 6	39	-0.667	2.669	(-1.532, 0.198)	0.127

In conclusion, all of the classes except for class 6 showed a statistically significant difference between their pre-course mean score and their post-course mean score. Even more than this, the classes that showed a difference had 95% confidence intervals showing increases in the mean score from pre- to post-course survey. It may be worth noting that class 2 had a greater p-value than every class other than class 6 and the 95% confidence interval showed only a very small gain when compared with the gains of the other classes. Class 2 had some instruction on topics related to the greenhouse effect which accounts for this marginal increase in mean post-instruction score. The 95% confidence intervals for classes 1 and 4 showed the greatest increases in mean score and classes 3 and 5 showed lesser increases but increases nonetheless. This supports the data we received from our analysis of variance test for the post-course means from which we were also able to obtain confidence intervals. These confidence intervals allowed us to rank the means from highest to lowest and differentiate between means that differed versus those which showed a chance of being the same.

Summary of Results

In summary of the results seen within the six classes in the study at Cal Poly, we have found significant results pertaining to the effectiveness of the GECI at measuring differences in student understanding and differences in student understanding based upon the mode of instruction. Using ANOVA Testing, we have

shown that there was not a significant difference in the pre-instruction mean scores of all six classes involved in the study. Using another ANOVA we showed that after instruction the means for post-instruction GECI scores were no longer all similar to each other. This is early evidence that the GECI is an effective tool for measuring student understanding. Before instruction all classes seem to be at a relatively equally low scoring level. After instruction their scores have changed in a significant way. Tukey-Kramer simultaneous confidence intervals then gave us insight into the differences in post-instruction mean scores. We were able to see that the post-instruction scores of classes 1 and 4 were similar, classes 2 and 6 were similar, and classes 3 and 5 were similar. Based upon the location of the confidence intervals in the Tukey-Kramer display and the mean scores for pre and post-instruction from Table 1, we can see that classes 1 and 4 had the greatest pre to post-instruction gains followed by a lesser gain from the similar classes 3 and 5, and almost no gains shown by classes 2 and 6. Using a T-Test we were able to determine that the pre and post-instruction mean scores of all classes, with the exception of class 6, differed in a statistically significant way. Table 1 shows us that these differences were gains in mean score. The greatest gains were from classes 1 and 4 which both received the Lecture Tutorial Activity on the greenhouse effect. The lesser gains of classes 3 and 5 were discovered after delivering them the greenhouse effect information by lecture. Classes 2 showed by far the smallest gain and class 6 was shown to have an insignificant gain. This was as we had expected since Class 2

received only limited instruction on greenhouse effect related topics. And Class 6 was the control groups which received no instruction on the greenhouse effect. We have now seen that the GECI has measured the pre-instruction understanding in a way that shows all classes to have insignificant differences in mean score but it has also shown increases in student scores after instruction. The GECI was also sensitive to the different modes of instruction. We saw clear segregation in the classes based upon their modes of instruction. Not only was the GECI sensitive to a change in student understanding but it was able to measure understanding precisely enough to show class segregation based upon the type of instruction each class received. These result support the idea that the Greenhouse Effect Concept Inventory is in fact a useful tool for measuring student understanding and that a lecture in combination with a discussion activity is more effective at conveying material than simply delivering a lecture narrative which in turn is more effective than an absence of instruction.

Cal Poly and University of Arizona Comparison

Now that our inspection of the study results of the six classes at Cal Poly has been reviewed, we should compare the Cal Poly population results to that of the study done by Dr. John Keller in 2006 at the University of Arizona. We began this comparison by using a Two-Sample T-Test to compare the mean pre-instruction GECI score of the University of Arizona population to the Cal Poly population. This was in

order to give us an idea of how similar the populations were in their understanding of the greenhouse effect before receiving any lecture on the topic. By comparing these pre-instruction means we can see if any number of external factors such as the media, the specific students attending the particular university, or increased awareness over the two year span between studies may have had an effect on the understanding of students.

University of Arizona vs. Cal Poly Pre-Score

Two-Sample T-Test and CI

Sample	N	Mean	Standard Deviation	SE Mean
Univeristy of Arizona	273	0.276	0.141	0.0085
Cal Poly	159	0.273	0.151	0.012

Difference = mu (1) - mu (2)
Estimate for difference: 0.0034
95% CI for difference: (-0.0256, 0.0324)
T-Test of difference = 0 (vs not =): T-Value = 0.23 P-Value = 0.817 DF = 312

Our Two-Sample T-Test resulted in a p-value of 0.817. This is greater than our alpha value of 0.05 for a 95% confidence interval so we fail to reject the null hypothesis that the two populations have equal mean pre-instruction scores. This means that there is not sufficient evidence to be 95% confident that the pre-instruction mean GECI scores of the population at the University of Arizona and Cal Poly are not equal. This supports the idea that the two populations were very

similar in their understanding of the greenhouse effect before receiving instruction on the topic. There is some evidence here to suggest that despite any change in media coverage or rising awareness of the greenhouse effect, student populations continue to have a similar level of understanding. We seem to have a conclusive result that the student populations have remained relatively unchanged in their understanding of the greenhouse effect, at least in the case of these two populations, over the period of time from 2006 to 2008.

For the purpose of further comparing the population at Cal Poly to the population at the University of Arizona, we used a Two-Sample T-Test to compare the gain in mean GECI score from pre to post-instruction. Both classes were astronomy classes and were both instructed on the greenhouse effect using the same Lecture Tutorial Activity. In terms of treatment and initial understanding these classes proved to be highly similar so for this study we were interested in testing to see if the two populations responded similarly to the Lecture Tutorial Activity.

University of Arizona Gain Versus Cal Poly Gain (ASTR-Lecture+LTA vs. Class 4)

Two-Sample T-Test and CI

Sample	N	Mean	Standard Deviation	SE Mean
Arizona ASTR-Lecture+LTA	69	0.577	0.224	0.027
Cal Poly Class 4	58	0.559	0.201	0.026

Difference = mu (1) - mu (2)
Estimate for difference: 0.0185
95% CI for difference: (-0.0564, 0.0933)
T-Test of difference = 0 (vs not =): T-Value = 0.49 P-Value = 0.626 DF = 124

The p-value returned by the T-Test in this case was 0.626 which is greater than our alpha value of 0.05 for a 95% confidence interval. For this reason we fail to reject the null hypothesis, which in this case is that the gain shown by the ASTR-Lecture+LTA class at the University of Arizona class is equal to that of Class 4 from the Cal Poly study. There is not sufficient evidence to show that there is a difference in the normalized gain shown by these two classes. This result supports the idea that both populations responded similarly to the Lecture Tutorial Activity in terms of their gain in understanding of the greenhouse effect as measured by the Greenhouse Effect Concept Inventory.

Our final investigation in our Cal Poly and University of Arizona comparison focused on the item by item percentage of students answering correctly on the pre-instruction GECI. The table below allows us to compare the performance of both

populations, on an item by item basis, before receiving instruction on the greenhouse effect in order to find how similar our populations were initially and to a more specific extent than simply their overall mean score on the GECI as a whole. This is in hopes that the GECI can be improved further by looking at how the two populations answered individual GECI items.

The percentage difference was calculated as:

(Cal Poly % correct – U of A % correct) / [(Cal Poly % correct + U of A % correct)/2] x 100%.

Table III: Correct Pre Answer Percentage of Cal Poly versus University of Arizona

GECI Question Number	Cal Poly Correct Pre Answer Percentage	University Of Arizona Correct Pre Answer Percentage	Absolute Percent Difference	Relative Percentage Difference
18	0.114	0.253	-0.1390	-75.7493
14	0.059	0.108	-0.0489	-58.5629
16	0.143	0.254	-0.1105	-55.6675
3	0.131	0.218	-0.0871	-49.9140
10	0.059	0.090	-0.0306	-41.0738
12	0.097	0.142	-0.0445	-37.2385
19	0.278	0.401	-0.1225	-36.0825
4	0.093	0.066	0.0268	33.7107
1	0.076	0.105	-0.0290	-32.0442
15	0.487	0.380	0.1072	24.7290
20	0.371	0.296	0.0753	22.5787
9	0.835	0.699	0.1364	17.7836
7	0.181	0.215	-0.0335	-16.9192
17	0.436	0.373	0.0634	15.6737
11	0.325	0.370	-0.0451	-12.9784
6	0.612	0.548	0.0638	11.0000
5	0.325	0.291	0.0338	10.9740
8	0.215	0.238	-0.0228	-10.0662
2	0.228	0.211	0.0168	7.6538
13	0.363	0.358	0.0048	1.3315

As shown in the table, students from Cal Poly and students from Arizona differed by 75.75% at most on a single question. On all but three questions, mean score differed by less than 50%. The majority of questions showed a difference of below 40% and almost half of the questions differed by less than 20%. Although the mean gain in score for the Lecture Tutorial Activity classes at Cal Poly and the University of Arizona did not prove to be significantly different, we do have large

differences in mean score when comparing the classes on a question by question basis. This should be looked into in future study.

GECI Revision and Development

Within the GECI there are 4 questions that need further investigation. There was question as to the effectiveness of questions 8, 9 and 15 in measuring student understanding. In addition, changes had been made to question 12 before the study was brought to Cal Poly but after the study performed at the University of Arizona. Table 3 gives the pre- and post-instruction percentage correct for the questions needing investigation. They have been placed in the table for comparison between the Cal Poly and University of Arizona results in order to see if students at both universities answered the questions in a similar manner. Based on the results decisions may be made about whether these questions should be kept, further studied, revised, or eliminated.

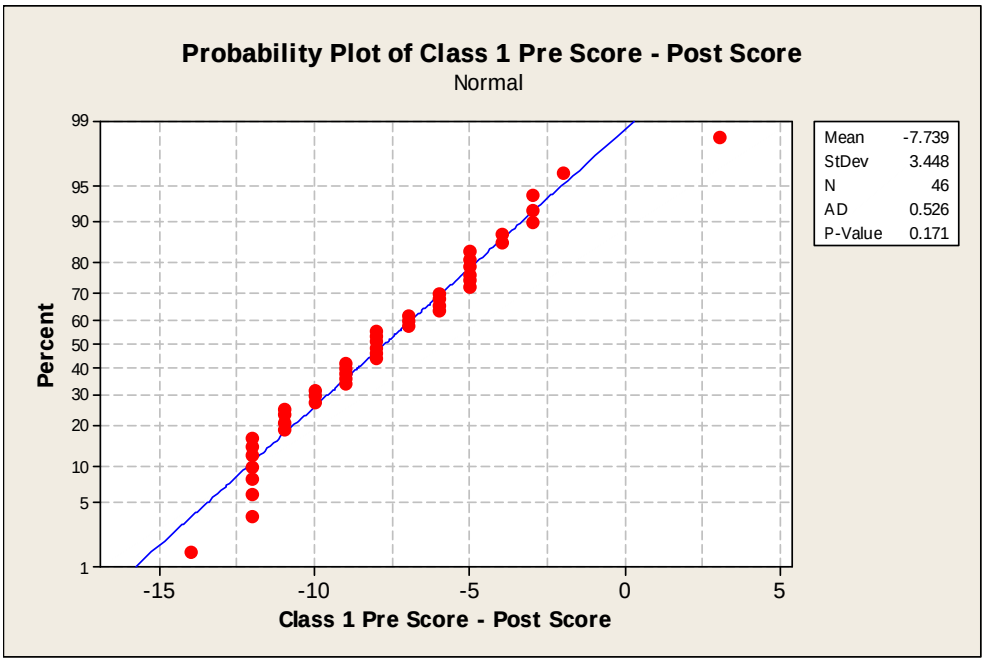
Table IV: Pre versus Post Comparison for GECI Questions 8, 9, 12, 15

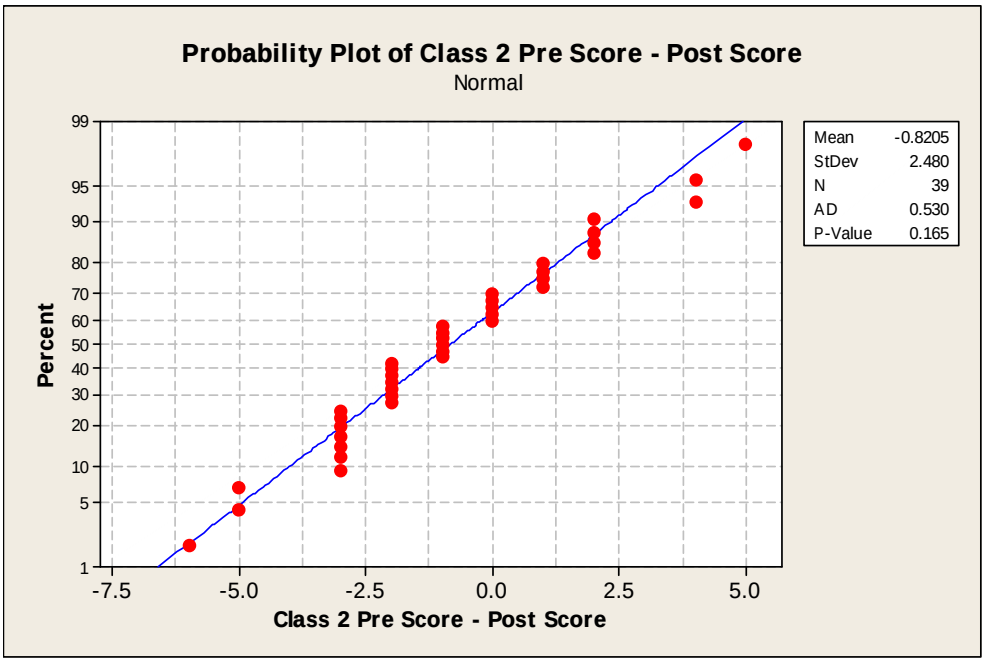
	Pre % Correct #8	Post % Correct #8	Post % Correct #9	Post % Correct #9	Post % Correct #12	Post % Correct #12	Post % Correct #15	Post % Correct #15
Cal Poly	0.215	0.245	0.835	0.814	0.097	0.530	0.487	0.532
University Of Arizona	0.238	0.244	0.699	0.648	0.142	0.548	0.380	0.581

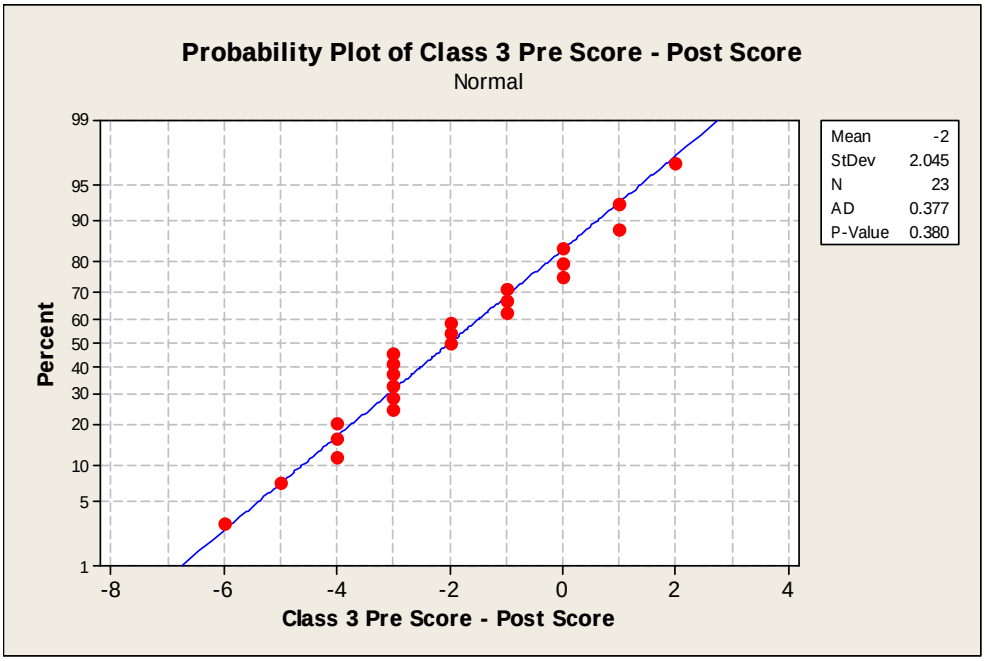
At both Cal Poly and the University of Arizona, Question 8 showed very little gain. This is a question that should be further investigated or better addressed in instruction since both populations showed very small improvement. Question 9 showed high pre-instruction scores and a decrease in the percentage of students answering correctly in the post-instruction administration both at the University of Arizona study in 2006 and also at Cal Poly in 2008. This suggests that the question itself maybe poor and should be removed. Question 12 was changed slightly in its wording in the time between the administration of the GECI and the University of Arizona and Cal Poly. The gains in this question from pre to post-instruction on this particular question were quite noticeable. It seems that the slight wording change has not affected the questions ability to measure a change in student understanding. This question does not seem to need further revision. Finally, Question 15 showed small gains. These gains were more significant at Cal Poly, however, both populations did show improvement on this question.

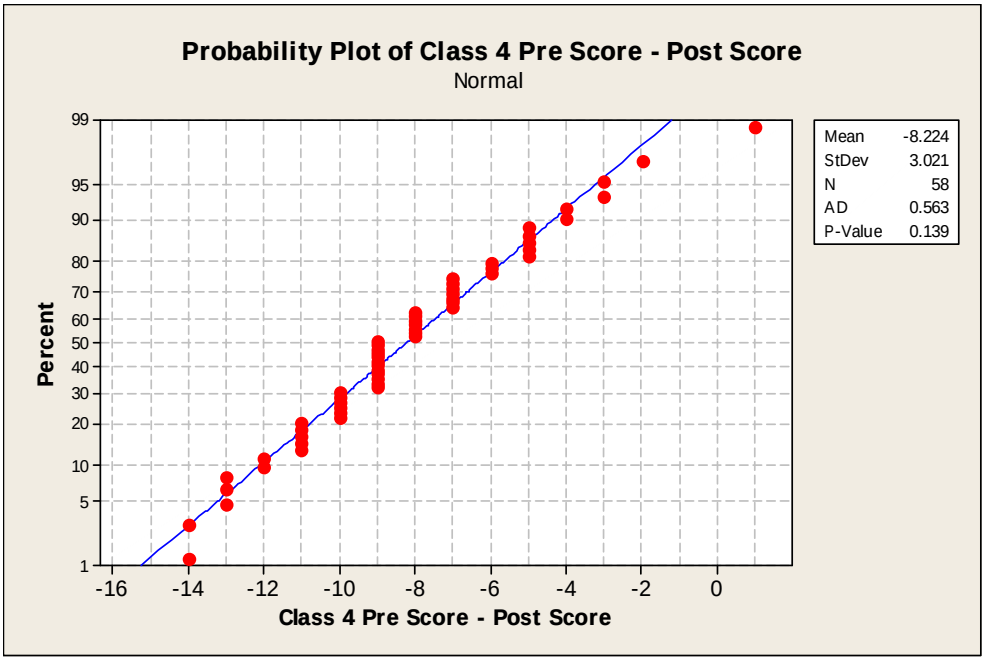
Assumptions

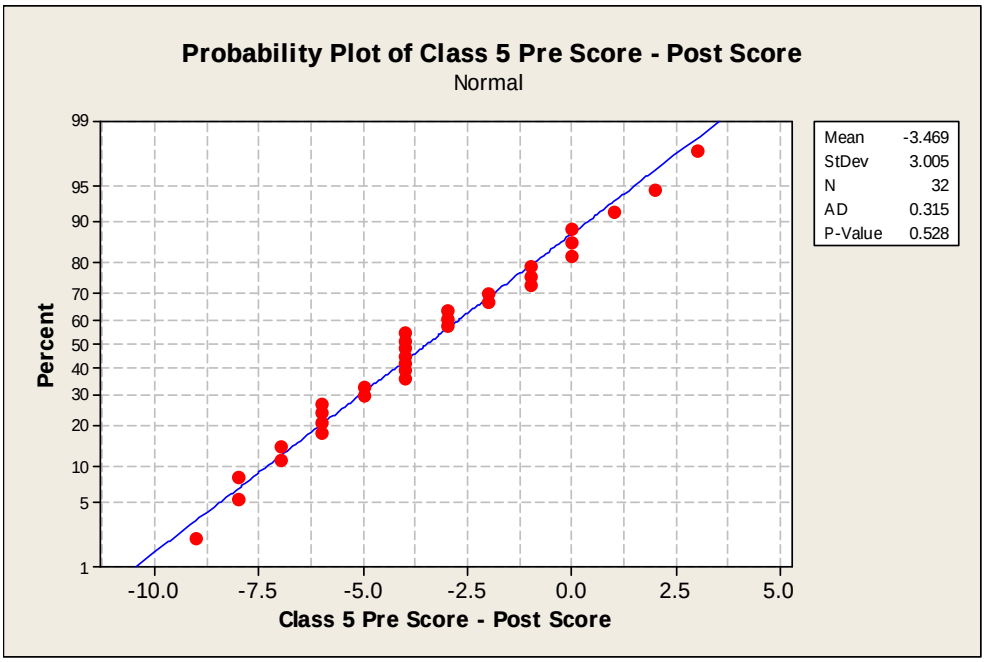
In order to ensure that our test results can be trusted we must check the assumptions for each test. For our paired T-test we assume we are sampling from normally distributed populations having equal variances. In our ANOVA test our assumptions are randomness and independence, normality, and homogeneity of variance. We know that we are sampling randomly and that our analysis of variance is for independent samples so we are left to test for normality and homogeneity of variance for both our ANOVAs and Paired T-test. Since both our test procedures involve the assumption that we are sampling from a normally distributed population we should first verify this using an Anderson-Darling Test for Normality. Below is a set of Anderson-Darling Tests for Normality for each of the class samples tested earlier.

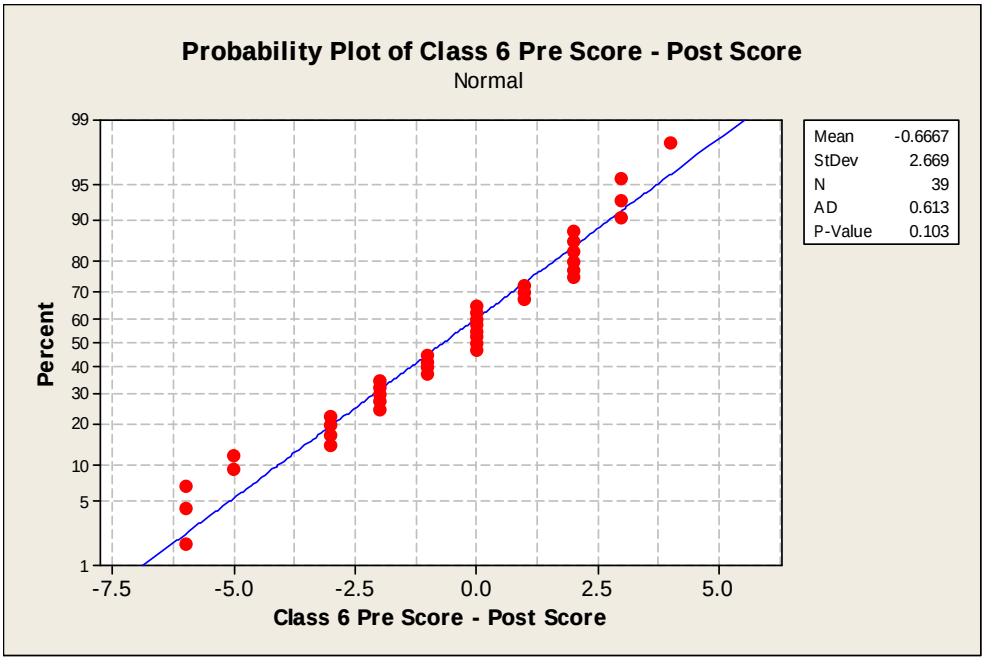

Probability Plot of Class 1 Pre Score - Post Score
Normal
Percent
99
95
90
80
70
60
50
40
30
20
10
5
1
-15
-10
-5
0
5
Class 1 Pre Score - Post Score
Mean -7.739
StDev 3.448
N 46
AD 0.526
P-Value 0.171


Probability Plot of Class 2 Pre Score - Post Score
Normal
Percent
99
95
90
80
70
60
50
40
30
20
10
5
1
-7.5
-5.0
-2.5
0.0
2.5
5.0
Class 2 Pre Score - Post Score
Mean -0.8205
StDev 2.480
N 39
AD 0.530
P-Value 0.165


Probability Plot of Class 3 Pre Score - Post Score
Normal
Percent
99
95
90
80
70
60
50
40
30
20
10
5
1
-8
-6
-4
-2
0
2
4
Class 3 Pre Score - Post Score
Mean -2
StDev 2.045
N 23
AD 0.377
P-Value 0.380

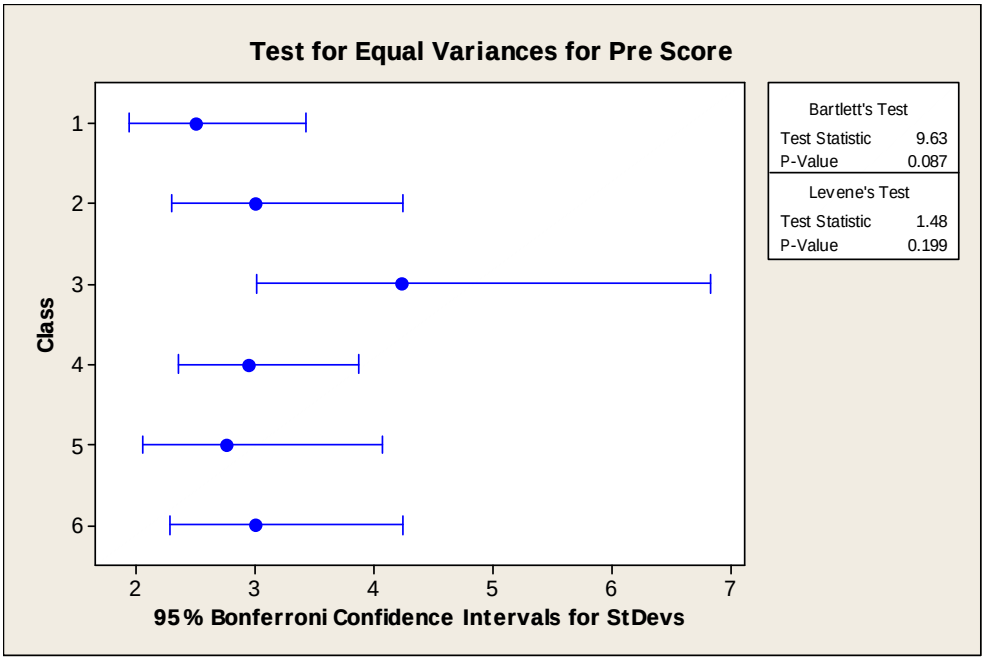

Probability Plot of Class 4 Pre Score - Post Score
Normal
Percent
99
95
90
80
70
60
50
40
30
20
10
5
1
-16
-14
-12
-10
-8
-6
-4
-2
0
Class 4 Pre Score - Post Score
Mean -8.224
StDev 3.021
N 58
AD 0.563
P-Value 0.139

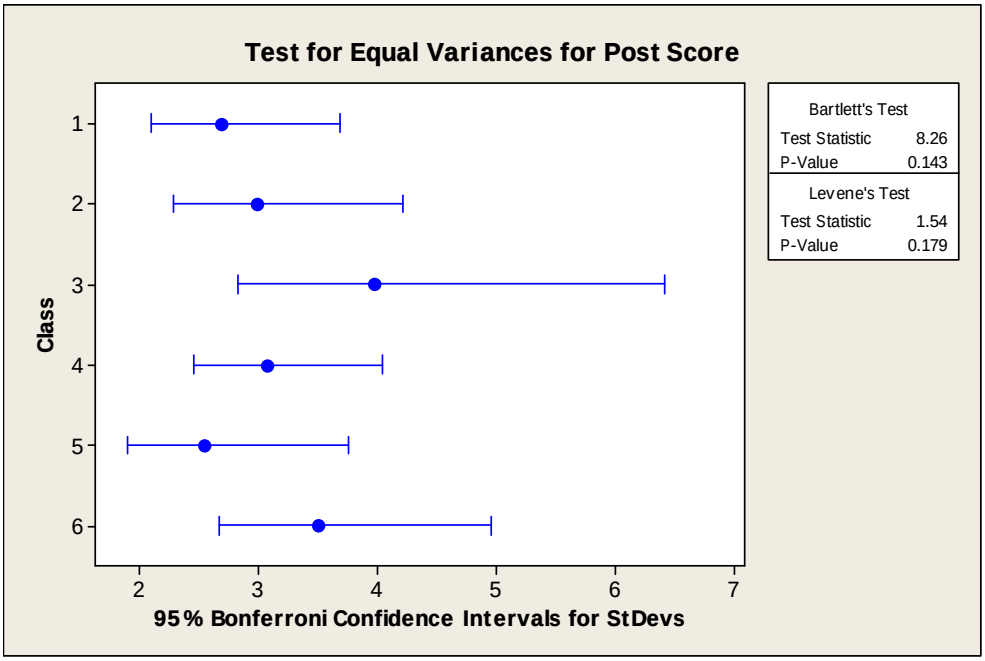

Probability Plot of Class 5 Pre Score - Post Score
Normal
Percent
99
95
90
80
70
60
50
40
30
20
10
5
1
-10.0
-7.5
-5.0
-2.5
0.0
2.5
5.0
Class 5 Pre Score - Post Score
Mean -3.469
StDev 3.005
N 32
AD 0.315
P-Value 0.528


Probability Plot of Class 6 Pre Score - Post Score
Normal
Percent
99
95
90
80
70
60
50
40
30
20
10
5
1
-7.5
-5.0
-2.5
0.0
2.5
5.0
Class 6 Pre Score - Post Score
Mean -0.6667
StDev 2.669
N 39
AD 0.613
P-Value 0.103

Each of these Anderson-Darling Tests for Normality gives us a p-value greater than our 0.05 level of significance. This means we fail to reject the null hypothesis that the data is normally distributed. We do not favor the alternative hypothesis that the data is not normally distributed. This means that there is some chance that the post-course scores which were subtracted from the pre-course scores, and whose result was used for the data of each of the classes, were normally distributed.

Next we test for homogeneity of variance for our ANOVA populations: the pre- and post-course scores. The results for these tests for equal variances are below.





Since both the Bartlett's and Levene's Tests for both the pre- and post-course scores give us p-values greater than our 0.05 level of significance, we fail to reject the null hypothesis that the variances are all equal. There is not sufficient evidence to conclude that at least one of the variances is different, therefore this assumption has been met. Since this is a class by class analysis of the variance of scores, this applies to both our ANOVA and Paired T-test. It encompasses all the data we have used. The ANOVA testing also requires that the ratio of the largest variance to the smallest variance not exceed three. This assumption is also met, as our largest variance is 17.885 and our smallest is 6.249, giving us a ratio that equates to 2.862.

Summary

The greenhouse effect is a natural phenomenon of the Earth's atmosphere that is vital to the survival of life. It is crucial that we as inhabitants understand the processes of the greenhouse effect, especially with the heightened awareness of its increases. It has been the goal of this project to: 1) Verify that the Greenhouse Effect Concept Inventory is indeed an effective tool for measuring a difference in knowledge after instruction in the greenhouse effect; 2) Measure and identify any difference in post-course mean score based upon the method of instruction used; and 3) Compare the pre-course greenhouse effect knowledge of students at Cal Poly against that of the students of the University of Arizona. Our hope is that this research will facilitate a higher quality of science education by identifying highly effective methods of instruction as well as areas of common misconception that should be better addressed in instruction.

In order to investigate whether or not the GECI is an effective tool for measuring a difference in student knowledge after instruction in the greenhouse effect at Cal Poly, we used an ANOVA test in order to compare pre- and post-instruction GECI scores. We obtained a p-value of 0.387 in our ANOVA for the pre-instruction scores which means that there was not sufficient evidence to show a difference in the pre-instruction scores of the six Cal Poly classes studied. From this result, we have shown that we can consider all of our Cal Poly classes to have a

similar knowledge of the greenhouse effect prior to instruction according to the GECI. In our second ANOVA, we compared the post-instruction scores of the six Cal Poly classes. We obtained a p-value of 0.000 for this post-instruction score ANOVA. This means we can say with 95% confidence that at least one of the six classes had a mean GECI score that differed in a significant way from that of the other Cal Poly classes. This lends evidence that the GECI did measure some change in knowledge from pre- to post-instruction.

Tukey-Kramer Tests performed at the 95% confidence level identified trends in the post-instruction mean scores based upon the method of instruction. These Tukey Tests showed that Classes 1 and 4, which were instructed using the Lecture Tutorial Activity, achieved mean GECI post-instruction scores that did not show any statistically significant difference. When compared to the other classes in the study, Classes 1 and 4 were found to have confidence intervals which were significantly different, and more specifically at a higher score range, than those of the other classes. Classes 3 and 5 were found to have no statistically significant difference in post-instruction score. These classes were instructed in the greenhouse effect but did not engage in the Lecture Tutorial Activity. The confidence intervals for the mean post-instruction scores of Classes 3 and 5 were higher than lower than those of Classes 1 and 4, yet higher than those of Classes 2 and 6. Classes 2 and 6 had the lowest post-instruction mean score confidence intervals. These classes showed no significant difference between their scores and both received no instruction in the

greenhouse effect. Each instructional method produced results unique to the two classes upon which it was used. The Lecture Tutorial Activity showed the greatest gains in mean score, while no instruction produced the least gain in score from pre- to post-instruction. The other two classes instructed with only lecture fell between the scores of the Lecture Tutorial Activity and no-instruction classes. Finally, using Paired T-Tests we investigated which classes had significant differences in their scores from pre- to post-instruction. All classes, with the exception of Class 6, showed a significant difference in mean score from pre- to post-instruction and a confidence interval which told us that these differences were gains. It is important to note, however, that Class 2 had a p-value which was much greater, at 0.046, than those of all classes other than Class 6. Class 2 showed the smallest gain by far from pre- to post-instruction though it was still found to be statistically significant. This is because they did receive very limited instruction in topics related to the greenhouse effect. In summary of the instructional methods, the Lecture Tutorial Activity produced the highest learning gains, followed by lecture instruction only, and the classes with no instruction showed the least gain in mean score.

In our investigation of the Cal Poly and University of Arizona populations, a Two-Sample T-Test showed that the populations did not have significantly different pre-instruction mean scores. The student populations are similar in pre-instruction knowledge. It seems that time, location, and any other factors have not significantly changed the knowledge of the student population before college instruction. For

further comparison, we used another Two-Sample T-Test to examine whether Cal Poly Class 4 responded to the Lecture Tutorial Activity in a similar manner to that of the University of Arizona ASTR-Lecture+LTA class. Both of these classes were instructed through the use of lecture and the Lecture Tutorial Activity. This T-Test gave us a p-value of 0.626 which means tells us that the gains shown in each class did not differ from each other in a significant way. Both the Cal Poly Class 4 and University of Arizona ASTR-Lecture+LTA class responded in a similar fashion to the Lecture Tutorial Activity. Though their gains and mean scores were highly similar, many of the questions which most Cal Poly students answered correctly were dissimilar to those that University of Arizona students answered correctly. This is an interesting result and the reasons for this are an area for future study. In summary, the Cal Poly and University of Arizona populations are highly similar, both showing significant gains in response to the Lecture Tutorial Activity.

We have verified the GECI as measuring differences in knowledge of the greenhouse effect after instruction. We have also found the Lecture Tutorial Activity to be a more effective method of instruction than purely lecture, which in turn is more effective than no instruction. In the implementation of both the GECI and Lecture Tutorial Activity, students at both Cal Poly and the University of Arizona showed similar mean scores and gains in response to similar instruction. These results are significant as their conclusions suggest ways in which we can both

identify and more effectively address misconceptions in students' understanding of the greenhouse effect.

Appendix

Greenhouse Effect Concept Inventory

(GECI)

On the scantron sheet, bubble in the best answer for each question. Please do not write on this survey.

1) Which of the following is the most abundant greenhouse gas?

   a) carbon dioxide (CO_2)
   b) water vapor (H_2O)
   c) methane (CH_4)
   d) oxygen (O_2)
   e) ozone (O_3)

2) Earth's atmosphere is warmer than it would be without a greenhouse effect. Which form of energy is absorbed by the atmosphere and mainly causes this increased temperature?

   a) radio
   b) infrared
   c) visible
   d) ultraviolet
   e) x-ray

3) On average, the total amount of energy leaving the Earth system to space

   a) is greater than the amount of energy arriving from space.
   b) is less than the amount of energy arriving from space.
   c) is roughly equal to the amount of energy arriving from space.
   d) depends upon the concentration of greenhouse gases in the atmosphere.
   e) depends upon the status of ozone in the atmosphere.

4) Which of the following is a primary characteristic of greenhouse gases?

   a) They can destroy certain molecules in the atmosphere.
   b) They bend and magnify sunlight entering the atmosphere.
   c) They can trap certain molecules in the atmosphere.
   d) They can bounce around more in the atmosphere.
   e) They are transparent to some forms of energy but not all.

5) The greenhouse effect is a very _____ process probably caused by _____.
   a) recent ; burning of fossil fuels, industry, agriculture, and other human activities.
   b) old ; plants that increase humidity and create conditions similar to those in a greenhouse found at a plant nursery.
   c) recent ; depletion of the ozone layer which allows more ultraviolet sunlight to reach the Earth's surface.
   d) old ; interactions between naturally occurring gases and various forms of energy in the atmosphere.
   e) recent ; natural processes including volcanic emission and changes in solar activity.

6) During the nighttime, Earth's surface mainly gives off (radiates) which form of energy?

   a) radio
   b) infrared
   c) visible
   d) ultraviolet
   e) Earth's surface does not give off energy during the nighttime

7) Which of the following most strongly affects Earth's overall surface temperature?

   a) heat released by factories and other industrial activities
   b) the destruction of the ozone layer allowing more sunlight into the atmosphere
   c) the flow of different forms of energy through the atmosphere
   d) air pollution trapped in the atmosphere by gases
   e) sunlight being magnified and focused by gases in the atmosphere

8) Which one of the following is <u>not</u> a greenhouse gas?

   a) carbon dioxide (CO_2)
   b) water vapor (H_2O)
   c) methane (CH_4)
   d) oxygen (O_2)
   e) ozone (O_3)

9) Which of the following best describes the relationship between the greenhouse effect and global warming?

a) The greenhouse effect and global warming are the same thing.
b) An increase in the greenhouse effect may be causing global warming.
c) Global warming may be causing an increase in the greenhouse effect.
d) The greenhouse effect and global warming are likely unrelated.
e) There is no definite proof that either the greenhouse effect or global warming exist.

10) The Sun mainly gives off (radiates) which two forms of energy?

a) ultraviolet and x-ray
b) ultraviolet and infrared
c) visible and ultraviolet
d) infrared and visible
e) radio and infrared

11) A planet that has a greenhouse effect

a) receives more UV sunlight because it lacks ozone in its atmosphere.
b) has an atmosphere that absorbs and then gives off certain forms of energy but not all.
c) receives more energy because it is closer to its central star.
d) has an altered atmosphere due to living organisms.
e) does not radiate any energy away into outer space.

12) Earth's surface is heated mainly by which two forms of energy?

a) ultraviolet and x-ray
b) ultraviolet and infrared
c) visible and ultraviolet
d) infrared and visible
e) radio and infrared

13) If human civilization had never developed on Earth, would there be a greenhouse effect?

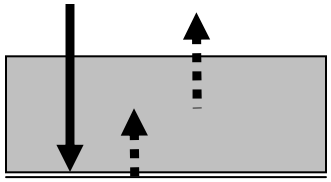
a) Yes, the greenhouse effect is caused by naturally occurring gases in the atmosphere.
b) Yes, the greenhouse effect is caused by plants giving off gases during photosynthesis.
c) No, the greenhouse effect is caused by humans burning fossil fuels and releasing pollutants.
d) No, the greenhouse effect is caused by humans depleting ozone in the atmosphere.
e) No, there is no conclusive evidence that a greenhouse effect exists.

14) At which of the following does the Sun give off (radiate) energy at its maximum intensity?

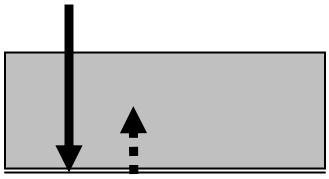
a) radio
b) infrared
c) visible
d) ultraviolet
e) x-ray

15) Each diagram below shows Earth's surface, the atmosphere, and outer space. The solid arrow represents incoming energy from the Sun that is absorbed by the surface. The dashed arrow represents energy that is radiated or given off by the surface and atmosphere. The thickness of the arrow roughly represents the amount of energy. Select the diagram that best represents the transport of energy among Earth's surface, the atmosphere, and outer space.

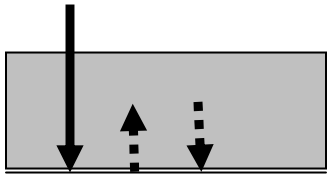
a)



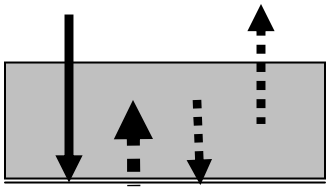
b)



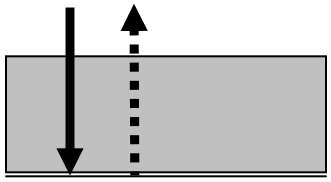
c)



d)



e)



16) Which of the following are the two most abundant greenhouse gases in Earth's atmosphere?

a) carbon dioxide (CO_2) and methane (CH_4)
b) ozone (O_3) and carbon dioxide (CO_2)
c) nitrogen (N_2) and oxygen (O_2)
d) oxygen (O_2) and carbon dioxide (CO_2)
e) water vapor (H_2O) and carbon dioxide (CO_2)

17) Earth's atmosphere mainly gives off (radiates) which form of energy?

a) radio
b) infrared
c) visible
d) ultraviolet
e) Earth's atmosphere does not give off energy

18) Due to the greenhouse effect, Earth's overall surface temperature is affected primarily by

a) an increase in energy entering from space.
b) a decrease in energy leaving to space.
c) both an increase in energy entering from and a decrease in energy leaving to space.
d) energy being permanently trapped in the atmosphere.
e) an increase in the amount of energy exchanged between the surface and atmosphere.

19) You walk from a region of shade into a region of direct sunlight and notice you start to feel warmer. Which of the following most correctly describes the cause of the temperature increase?

a) You absorb more ultraviolet energy than you give off (radiate) as visible energy.
b) You absorb more visible energy than you give off (radiate) as infrared energy.
c) You reflect more ultraviolet energy than you give off (radiate) as infrared energy.
d) You reflect more visible energy than you give off (radiate) as infrared energy.
e) You reflect more infrared energy than you give off (radiate) as visible energy.

20) During the daytime, Earth's surface mainly gives off (radiates) which form of energy?

a) radio
b) infrared
c) visible
d) ultraviolet
e) Earth's surface does not give off energy during the daytime

21) What is your gender?

a) Male
b) Female

22) What year are you in university?

a) Freshman
b) Sophomore
c) Junior
d) Senior
e) Other

23) Which of the following best characterizes your academic major(s)?

a) Science
b) Non-science
c) Double major (science and non-science)
d) Other

24) Prior to this semester, which of the following best describes the highest level of coursework you have had on the greenhouse effect?

a) I have never taken a class that covers the greenhouse effect.
b) Before college, I took a class that briefly touched upon the greenhouse effect.
c) Before college, I took a class that dealt extensively with the greenhouse effect.
d) In college, I took a class that briefly touched upon the greenhouse effect.
e) In college, I took a class that dealt extensively with the greenhouse effect.

25) During this semester, have any of your other courses besides this course covered the greenhouse effect?

a) Yes
b) No

Thank you for your participation.

Please wait for instructions on returning survey.

Lecture Tutorial Activity

Energy Flow through the Atmosphere and the Greenhouse Effect

Section 1: The solar spectrum

Objects give off different amounts of light depending upon their temperature. Figure 1 below shows the energy spectrum for our Sun along with the percent of energy given off by the Sun in the ultraviolet (UV), visible (VIS), and infrared (IR) portions of the electromagnetic spectrum.

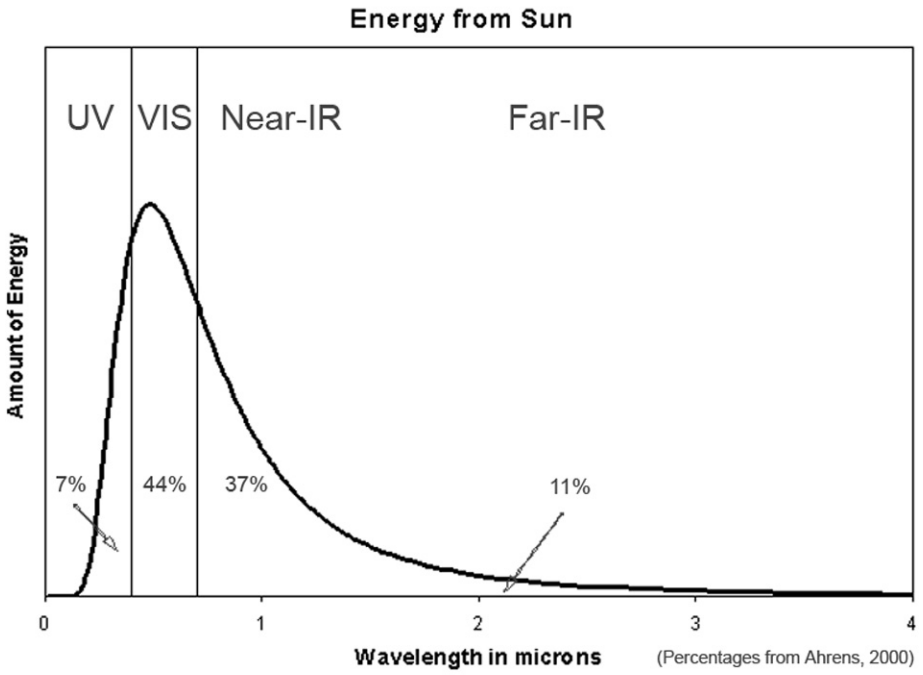


Figure 1

1) Which TWO forms of light account for the majority of energy coming from the sun: ultraviolet, visible, or infrared? Which of the three accounts for the least energy? Provide numbers to back up your answer.

2) Based upon Figure 1, why is ultraviolet light NOT an important energy source for heating the surface of the Earth?

3) Consider the following debate between two students regarding the energy given off by the Sun.

Student #1 – I think that the Sun gives off most of its energy at ultraviolet wavelengths because ultraviolet light is more intense than visible light and you always hear about ultraviolet light causing sunburns.

Student #2 – Even though UV photons are more energetic than visible photons, the Sun simply gives off less ultraviolet photons and gives off way more visible and infrared photons. So I think that these longer wavelength photons account for most of the energy coming from the Sun.

Do you agree and/or disagree with either of these students? Explain your reasoning.

Section #2: Atmospheric absorption of light

Earth's surface temperature is affected by energy that is absorbed at the surface. However, a photon's ability to travel through our atmosphere and reach the ground depends upon its wavelength. Figure 2 below shows that certain energies of light are absorbed in our atmosphere more than others. The figure also lists the primary gas molecules responsible for the absorption.

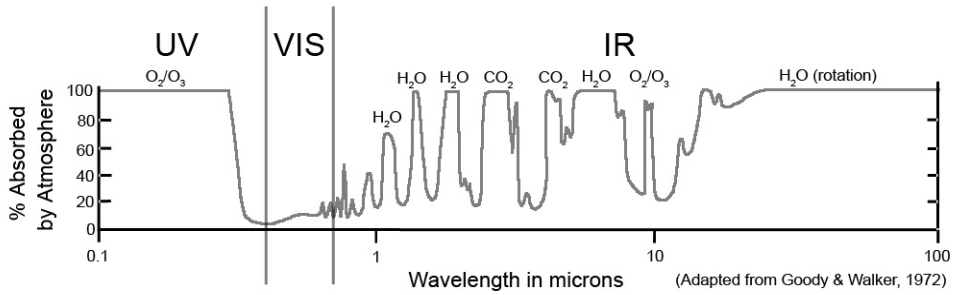


Figure 2

4) What gas molecules are primarily responsible for the absorption of each of the following forms of light in our atmosphere?

Type of Light	Molecule(s) Responsible for Absorption
Ultraviolet	
Visible	
Infrared	

5) Comparing the visible and the infrared parts of the spectrum, which would you say has an easier time getting through our atmosphere? Which experiences more absorption?

Section 3: Spectrum from Earth's surface

Once visible light from the Sun reaches the surface of Earth, it can either be reflected back towards space as visible light or be absorbed by the ground. Reflected light does not change the temperature of the surface, while absorbed light causes the temperature of the surface to increase. Ground that is heated then gives off infrared light based upon its increased temperature. As an example, black asphalt absorbs more visible light and gives off more infrared light than a white sidewalk on a hot day.

6) The Sun is approximately 6000K at the surface and gives off most of its energy as visible light; the Earth's surface is much cooler at about 288K. What type of light do you think the Earth's surface gives off: ultraviolet, visible, or infrared light? Explain your reasoning.

7) Does Earth's surface give off light at night? If so, what type? If not, why not?

8) Consider the following debate between two students regarding the energy given off by Earth's surface.

Student #1 – The Sun mainly gives off visible light and so does Earth's surface because I can see it during the daytime.

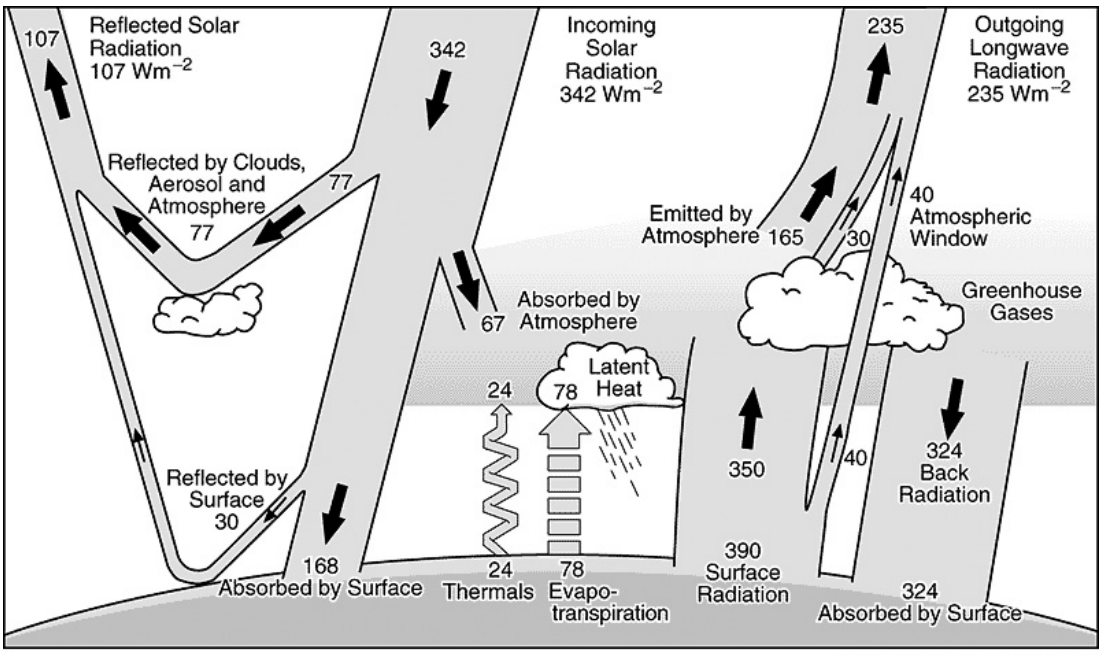
Student #2 – But that's just reflected sunlight. Earth's surface is much cooler than the Sun and mostly gives off energy closer to the kind that our bodies give off – infrared light. I'm not sure, but I think the surface probably radiates infrared light during both the daytime and the nighttime based upon its temperature.

Do you agree and/or disagree with either of these students? Explain your reasoning.

9) Based upon your answer to Question 5, will the light given off by Earth's surface easily travel back through the atmosphere to space or will it be absorbed by molecules in the atmosphere? Explain your reasoning.

Section 4: Energy flow and the greenhouse effect

Figure 3 below shows the flow of energy originally from the Sun through the Earth system (surface and atmosphere). The numbers listed describe the rate of energy flow through system (units of watts per square meter). A larger number indicates that more energy is flowing through that labeled pathway.



(Source: Kiehl and Trenberth, 1997)

Figure 3

10) In Figure 3, fill in the four empty boxes with the type of light involved: UV, visible, or IR.

11) What two types of light primarily heat Earth's surface. What one type of light primarily heats Earth's atmosphere?

12) Is more energy absorbed by the Earth's surface in the form of light from the Sun or light from the atmosphere? Provide values to justify your answer.

13) Due to the energy absorbed by Earth's surface from the atmosphere, is the surface temperature warmer or cooler than it would be without this energy input?

14) How does the total rate of energy coming in from space (incoming solar energy) compare to the total rate of energy leaving out to space (all energy reflected or given off by earth and atmosphere)? Provide values to justify your answer.

15) Based upon your answer to Question 14, does the total energy in the Earth system increase, decrease, or stay the same over short time scales? Explain your reasoning.

The flow of energy shown in Figure 3 is the source of the natural "atmospheric greenhouse effect." Visible light penetrates the atmosphere and is absorbed by the surface. The heated surface gives off infrared light that is then absorbed by the atmosphere. The heated atmosphere gives off infrared light out to space and also back down to Earth's surface, making the surface temperature warmer than it would be without a greenhouse effect. The amount of energy entering and leaving the Earth system is balanced, but the Earth's surface temperature is warmer because the surface is heated both by visible light from the Sun and infrared light from the atmosphere.

16) In Question 4, you listed several gases. Which of these are primarily responsible for absorbing and emitting infrared light? What characteristic makes them greenhouse gases?

17) Consider the following debate between two students regarding the greenhouse effect.

Student #1 - So the greenhouse effect is caused by infrared light being trapped in Earth's atmosphere by greenhouse gases. Visible light from the sun heats the ground, but the infrared light given off by the ground gets permanently trapped in the atmosphere and can never escape.

Student #2 – I think that's close. But based on Figure 3, all of the arrows balance and just as much energy leaves the planet as comes in. I think the greenhouse effect makes the surface hotter than it would be without greenhouse gases because the ground gets visible light from the Sun AND infrared light from the atmosphere given off back to the surface.

Do you agree and/or disagree with each of these students? Explain your reasoning.

Institutional Review Board Approval Materials

This appendix provides copies of documentation from the Cal Poly Institutional Review Board (IRB) in the Human Subjects Protection Program. Included is a copy of the Subject Informed Consent form used with the GECI survey.

INFORMED CONSENT TO PARTICIPATE IN A VALIDATION OF THE GREENHOUSE EFFECT CONCEPT INVENTORY (GECI) SURVEY INSTRUMENT

A research project on the use a concept inventory survey instrument to determine student understanding of the greenhouse effect is being conducted by John M. Keller, Ph.D, in the Department of Physics at Cal Poly, San Luis Obispo. The purposes of the study are 1) to validate the use of a concept inventory designed to test pre- and post-instruction understanding of the greenhouse effect, and 2) to investigate changes in student understanding of the greenhouse effect over time.

You are being asked to take part in this study by responding to the questions on the attached survey. You may choose not to answer some or all of the questions. A post-instructional survey will also be administered at the end of the quarter. Your participation will take approximately 15 minutes of class time today and an additional 15 minutes of class time at the end of the quarter. Please be aware that you are not required to participate in this research and you may discontinue your participation at any time without penalty. Your responses to this survey will in no way impact your grade.

There are no known risks from your participation and no direct benefit from your participation is expected. Potential benefits associated with the study include the ability to better address student misconceptions about the greenhouse effect in educational settings. There is no cost to you except for your time and you will not be compensated for your participation.

Your confidentiality will be protected by Dr. Keller, who will be the only individual with access to survey information prior to the removal of identifying information. You are asked to provide your name on the survey so that pre- and post-instruction surveys can be matched. However, survey information will remain

confidential and all identifying information will be removed after the survey match has been completed.

If you have questions regarding this study or would like to be informed of the results when the study is completed, please feel free to contact John M. Keller, Ph.D., at 805-756-2095, jmkeller@calpoly.edu. If you have questions or concerns regarding the manner in which the study is conducted, you may contact Steve Davis, Chair of the Cal Poly Human Subjects Committee, at 756-2754, sdavis@calpoly.edu, or Susan Opava, Dean of Research and Graduate Programs, at 756-1508, sopava@calpoly.edu.

If you agree to voluntarily participate in this research project as described, please indicate your agreement by completing the attached survey. Please keep a copy of this form for your reference, and thank you for your participation in this research.

Sincerely,

John M. Keller, PhD.

Bibliography

Keller, John. *Development of a Concept Inventory Addressing Students' Beliefs and Reasoning Difficulties Regarding the Greenhouse Effect*. 2006.
PhD Dissertation submitted to faculty of Department of Planetary Sciences at the University of Arizona