A Comparison Study – Is the Chemistry Modeling Approach to Teaching Energy Superior to the Traditional Approach?

Principal Investigator: Dr. Robert Culbertson
Co-Investigators: Donna Blakeney, Philip Root, and Andrea Strock

Action Research required for the Master of Natural Science degree with an emphasis in Physics.

Arizona State University

July 2009
# Table of Contents

Abstract 3
Description of Project 4
Rationale 6
Literature Review 9
Method 11
Timeline 15
Results
  Control Group Report 16
  Investigator 1 Report 21
  Investigator 2 Report 34
  Investigator 3 Report 42
  Qualitative Report 49
Study Conclusion 66
Implications for Further Research 67
Acknowledgements 68
References 69
Appendices
  A - Teacher Notes from Chemistry Modeling Units Used 72
  B - BECI 79
  C - CCI 86
  D - Questions Analyzed from the Inventories 94
  E - Sample Questions Used in the Interviews 98
Abstract:

Traditional teaching methods address the concept of energy as its own unit. Chemistry curriculum normally does not fully incorporate or explain chemical energy throughout the rest of the lessons. Despite the many resources available to Chemistry teachers and Chemistry students, a strong preference is shown towards avoiding the energy concept altogether. (Swackhamer, 2005) Students have great misconceptions about the concept of energy and how energy relates to Chemistry. (Swackhamer, 2005) A comparison study was performed analyzing the performance of treatment and non-treatment (control) classrooms. Treatment involved the use of reformed practice teaching, referred to in this paper as the modeling approach. Results of this study indicate that students who learned about the concept of Energy through a modeling approach improved their knowledge of energy more than a traditional Chemistry classroom control group.
Description of Project:

Traditional education in America is preventing students from addressing and correcting their preconceptions and misconceptions in chemistry and succeeding academically. (Horton, 2004) There are tremendous obstacles that students face in traditional education, especially when the majority of time in class involves sitting and listening to a teacher-focused curricula. Therefore, to accommodate the nature of students and how they interact with their learning environment teachers must implement new strategies. Thus, the classroom must be student-focused.

The traditional approach of chemical education that emphasizes the microscopic to macroscopic concept is not effective in addressing the students’ misconceptions. Students come with many preconceptions of the chemistry concepts to be taught to them. It seems as if the students do not fully understand the concepts and they fill in the missing pieces with their preconceptions or misconceptions. In order to overcome the students’ misconceptions, the students need to experience the concept first and then be able to explain it. (Horton, 2004) Models, which allow students to explain their experience, must be constructed, developed, tested and consistently applied if the students are to replace their misconceptions with an appropriate framework of knowledge. It is necessary to implement an alternate chemistry teaching method, utilizing the macroscopic to microscopic concept. In this study, a teaching method referred to as “modeling”, will be used to allow students to experience concepts through activities and labs and then develop and use models to explain these concepts.

Four different classes conducted this comparison study, one using the traditional approach and three using the modeling approach. Two teachers in this study taught in the Deer Valley Unified School District. Mr. Bryan Simon (pseudonym) taught at Deer Valley High School for over 20 years and has a degree in chemistry education. Mr. Simon taught his students utilizing the traditional approach. He had a combination of honors chemistry and regular chemistry classes. Mrs. Donna Blakeney taught at Deer Valley High School for 11 years and has a degree in secondary education with an emphasis in chemistry. Mrs. Blakeney taught her students utilizing the modeling approach. She had a combination of honors chemistry and regular chemistry classes. Two teachers in this study taught in the Chandler Unified School District. Mr. Philip Root taught at Chandler High School for 5 years and has a degree in secondary education with an emphasis in chemistry. Mr. Root taught his students utilizing the modeling approach. He had a combination of regular and honors chemistry classes. Mrs. Andrea Strock taught at Perry High School for 3 years and has a degree in secondary education with an emphasis in chemistry. Mrs. Strock taught her students utilizing the modeling approach. She had regular chemistry classes only. Throughout the study, the chemistry modeling approach (materials included in Appendix A) consistently addressed students’ misconceptions. There was a direct correlation to improving students’ understanding of energy storage and energy transfers with regard to endothermic and exothermic processes.

This action research examines both quantitative and qualitative data collected from the instructors. The quantitative data came from administering the Basic Energy Concept Inventory (BECI) and the Chemical Concept Inventory (CCI). Other forms of assessment included quizzes, unit tests, white board presentations, and classroom discussions. Interviews given
before and after the unit on energy provided the qualitative data. Videotape of these interviews allowed for future review.

The bulk of the quantitative data came from the Basic Energy Concept Inventory (BECI), that was given as a pre-test and post-test. Analyzed scores of select questions on the test pertained to energy in Chemistry only. The software programs “Microsoft Excel” (MS Excel) and “Statistical Package of Social Sciences” (SPSS) were used to analyze the BECI and CCI test questions and scores. Scores on these two tests were compared between the two methods of teaching - traditional versus modeling. One investigator also compares first year results to second year results. The results of the study demonstrate an increase in the scores of the pre-test to post-test for both methods. However, the students who were exposed to the modeling method experienced significantly greater increases in scores. In particular, the results of data analysis will show a direct correlation between modeling instruction and the improvement of students’ understanding of energy storage and energy transfer in endothermic and exothermic processes.
Rationale:

It has become apparent that students have difficulty understanding energy storage and energy transfers with regard to endothermic and exothermic processes in their chemistry class. Students arrive in our classes having many preconceptions and misconceptions concerning energy (Arons). It is difficult to change those preconceptions, as students are not receptive to new knowledge and ideas. It is taught that if a reaction feels hot then an exothermic reaction has occurred and if the reaction feels cold then an endothermic reaction has occurred. This concept has not been fully explained to them with regard to what is happening at the particle level or that a change in thermal energy has occurred. As a result, “energy is not well understood by our students. Students graduating from secondary schools generally cannot use energy to describe or explain even basic, everyday phenomena.” (Swackhamer 2005, 1) Standard chemistry textbooks, explain that energy is the capacity to do work or produce heat. The textbooks then explain energy in terms of physics concepts, such as lifting an object or kicking a ball. The textbooks do not explain energy in terms of chemistry concepts other than to say, “or produce heat.” What does this mean to the student who is just discovering this concept? During this study, investigators presented exothermic and endothermic processes in a manner designed to improve students’ knowledge of energy storage and energy transfers.

The purpose of this study was to address the students’ energy misconceptions in the Chemistry science class with the implementation of the modeling method. A passing grade will be a reflection of the student’s ability to demonstrate proficiency in the science standards of the following are the Arizona State Standards (Arizona Department of Education Standards, 2005, p. 10):

Strand 1: Inquiry Process
Concept 4: Communication
   (PO1, PO2, PO3, PO4)
Strand 5: Physical Science
Concept 3: Conservation of Energy and Increase in Disorder
   (PO2, PO3, PO6)
Concept 4: Chemical Reactions
   (PO10)
Concept 5: Interactions of Energy and Matter
   (PO4)

Research Context:

The pilot study conducted during the 2007-2008 academic school year at Deer Valley High School included the control group and investigator 1. The study continued a second year, during the 2008 – 2009 school year at Deer Valley High School. During the 2008 – 2009 school year, Chandler High School and Perry High School contributed their results.

Deer Valley High School is located in northwestern portion of the Phoenix Metropolitan area in Glendale, Arizona in the Deer Valley Unified School District. Deer Valley High School
offers a comprehensive education and has a media center, two computer labs, two gymnasiums, a large auditorium, and a variety of classrooms. The school offers various clubs and organizations for the students. These organizations include student government, band, yearbook, journalism, Physics club, and many more. Deer Valley also has an award winning athletic program where students can participate in a myriad of sports activities. In addition to the regular curriculum, the school offers special education, honors, advanced placement, and dual enrollment courses with Rio Salado Community College.

The Arizona school report card for the academic school year of 2004-2005 for Deer Valley High School was excellent. The achievement profile designates Deer Valley High School to be “Highly Performing”. (Arizona Department of Education, 2005) During the 2005-2006 school year there were 2222 students attending Deer Valley. Last year, 67% of the tenth grade class met or exceeded on the mathematics portion of the Arizona Instrument to Measure Standards (AIMS), 72% of the tenth grade class met or exceeded in reading, and 72% of the tenth grade class met or exceeded in writing (Arizona Department of Education, 2005).

Chandler High School is located in the eastern valley of the Phoenix metropolitan area in Chandler, Arizona in the Chandler Unified School District. Chandler High School offers a standard accredited curriculum that is supplemented by special education, honors, Advanced Placement, International Baccalaureate, and dual enrollment courses. Facilities at Chandler High include computer labs, media center, the Chandler Center for the Arts, two gymnasiums, large science classrooms with desk and lab areas, and various standard classrooms. The diversity of the student population lends itself to a variety of extracurricular programs and organizations including journalism, yearbook, student government, band, chorus, science club, dance clubs, cultural clubs, and more. Chandler High also boasts a varied and accomplished athletics program. The mission of Chandler High School is to educate its youth and prepare them for exciting opportunities of the future. Recognizing the strengths of its diverse community, the school builds upon its strong traditions to increase knowledge, pride, spirit, and character.

The Arizona school report card for the academic school year of 2007-2008 for Chandler High school validates its mission statement. The achievement profile designates Chandler High School as an “Excelling” school (Arizona Department of Education, 2009). During the 2007-2008 school year there were 3,071 students attending Chandler High. Last year, 82% of the tenth grade class met or exceeded on the mathematics portion of the Arizona Instrument to Measure Standards (AIMS), 82% of the tenth grade class met or exceeded in reading, and 76% of the tenth grade class met or exceeded in writing (Arizona Department of Education, 2009).

Perry High School is located Perry High School is located in the southeastern portion of the Phoenix Metropolitan area in Gilbert, Arizona in the Chandler Unified School District. Perry is a new school that has been open for only two years, and offers a variety of classes, clubs and organizations to freshmen, sophomores, and juniors. Organizations at PHS include student government, band, yearbook, journalism, Ecology club, and many more. Perry High School already has an award winning athletic program where students can participate in a myriad of sports activities. In addition to the regular curriculum, the school offers special education, honors, advanced placement, and dual enrollment courses with Rio Salado Community College.
The Arizona school report card for the academic school year of 2007 - 2008 for Perry High School was excelling. (Arizona Department of Education, 2008) During the 2007 - 2008 school year there were 797 students attending Perry High. Last year, 86% of the tenth grade class met or exceeded on the mathematics portion of the Arizona Instrument to Measure Standards (AIMS), 86% of the tenth grade class met or exceeded in reading, and 83% of the tenth grade class met or exceeded in writing (Arizona Department of Education, 2008).
Literature Review:

It is important to note that science education research is focusing on both the best way to teach students as well as how to improve science learning (Odell, 2005 and McDermott, 2001). If students are to become productive members of society, research must continue so science education can improve. This action research project focused on two aims. The first aim was to investigate the effectiveness of the chemistry modeling process when compared to the alternative traditional form of teaching. The instructor who used the traditional method, was not exposed to the modeling process, therefore did not use modeling techniques. The second focus of the study was to investigate the effectiveness of a first year modeler versus a second year modeler. Before completing this study, a review of current research efforts was conducted.

Hestenes characterizes classroom discourse as follows; “By discourse we mean all the interchanges in the classroom, including whiteboard presentations, which are crucial devices for focusing and directing discourse.” (Hestenes, 2007) Through white boarding sessions, the teacher is actually getting the student to discuss their activity, lab, or homework assignments with each other and the instructor. The students construct the discourse so they are “able to justify beliefs.” (Hestenes, 2007) Students use whiteboards to express themselves and to teach each other. After students put their thoughts on the whiteboard and discuss their findings, they share their knowledge or ideas with the entire class. Students work through their misconceptions to arrive at the correct answer themselves during the creation and presentation of the whiteboard.

Students come to chemistry class with many misconceptions in regard to energy and energy transfers. Even “A” students can have a high level of misconceptions when they come into your class. It is very important for teachers to address these misconceptions and for students to learn through their own experience (Horton, 2004). These misconceptions mislead the students in understanding the key concepts. (Levy Narcum, 2004) Students need models to help them see the flaws in their preconceptions and to overcome their misconceptions. Because misconceptions play a large role in the students’ learning, it is very important the teacher addresses those misconceptions quickly (Ozmen, 2003).

A specific difficulty that students run into once they are in the classroom is the different terminology used for energy-related concepts. Westphal studied student conceptions of thermal energy and the difference between temperature and heat. Westphal explains the difficulties students face in solving common chemistry problems related to “heat.” The issue is that students get confused when questions are phrased with “calculate the heat required to….” Westphal states that questions like this should read, “calculate the energy required to…” as energy is what is actually being referred to. (Westphal, 2003) With so many terms for similar concepts and so many similar terms for different concepts, students are easily confused. In a separate study, Barker points out that students need to come to consensus on which terms to use. Barker further states that it is best for students to experience phenomena first and then use their own words to help explain what they saw. In addition, continued use of particle diagrams allows for further clarification of concepts and identification of student misconceptions (Barker).
Furthermore, high school students are not able to relate the fundamental concepts involving energy to real life situations. Energy is presented but in such an incoherent way that the students are not able to develop a true understanding of energy. Students are either not able to correct any misconceptions, or may start to overcome their misconceptions but then revert back to their original beliefs (Swackhamer, 2005). According to Beall, students also “lock” on to notions presented in class that are not applicable to the thermodynamic problems they are trying to solve. Students will even ignore recently developed ideas and revert to prior conceptions. Beall found that in-class writings provided “a powerful means for identifying student problems and misconceptions so they can be remedied at the time” (Beall, 1994).

Current research supports that if students are to learn about energy and be able to relate those concepts to other areas of life, the instructor should teach energy in a coherent way so more students understand energy. Multiple representations, including visuals, are better at addressing misconceptions than just verbalization. In other words, the misconceptions are easier to identify when the student has visually shown what they mean and then can explain the visual. The teacher can then question the student using their own visuals as a means to get to the correct answers (Horton, 2004).

In summary, current research points to needs in science education that the modeling method of instruction is able to meet. Modeling instruction focuses on a phenomena first approach. The term “model” refers to the various mental constructs and representations that are developed after experiencing the phenomena. White boarding sessions, a cornerstone practice of modeling instruction, allow students to describe the phenomena in their own words, to develop and discuss their models, and to use their models to solve related problems. White boarding as a class discourse tool also serves the purpose of allowing students to come to consensus on terminology. At the same time, preparation of the whiteboards is a form of in-class writing, allowing teachers to easily diagnose student misconceptions.

The comparison study presented in this paper addresses the needs expressed in the review of current literature. The study will illustrate that the modeling approach to teaching chemistry provides the opportunities and means necessary for students to learn about energy in a more coherent and meaningful way than through a traditional classroom approach.
Method:

Control:

The control population for this research was Mr. Bryan Simon’s traditional method of teaching energy in Chemistry. Mr. Simon is employed at Deer Valley High School. This school is located in northwestern portion of the Phoenix Metropolitan area in Glendale, Arizona in the Deer Valley Unified School District. This is a diverse school with approximately 2,200 students enrolled. The enrollment consists of approximately 78% Caucasian, 5% African American, 13% Hispanic, and 4% Asian/Pacific Islander. The demographics within the chemistry classes are very similar to those of the entire school. Mr. Bryan Simon taught his students utilizing the traditional approach to teaching chemistry. He has been teaching chemistry for over 20 years. Mr. Simon had two honors chemistry classes and three regular chemistry classes in the 2007 – 2008 school year.

Investigator 1:

Investigator 1 in this research was Mrs. Donna Blakeney. Mrs. Donna Blakeney is employed at Deer Valley High School. This school is located in northwestern portion of the Phoenix Metropolitan area in Glendale, Arizona in the Deer Valley Unified School District. This is a diverse school with approximately 2,200 students enrolled. The enrollment consists of approximately 78% Caucasian, 5% African American, 13% Hispanic, and 4% Asian/Pacific Islander. The demographics within the chemistry classes are very similar to those of the entire school. Mrs. Donna Blakeney taught her students utilizing the modeling approach. She has been teaching chemistry for eleven years. Mrs. Blakeney had two honors chemistry classes and three regular chemistry classes in the 2007 – 2008 school year, and two honors chemistry and two regular chemistry classes in the 2008 – 2009 school year.

Investigator 2:

Investigator 2 in this research was Mr. Philip Root. Mr. Philip Root is employed at Chandler High School. The school is located in the eastern valley of the Phoenix metropolitan area in Chandler, Arizona in the Chandler Unified School District. The school has a very diverse population with approximately 3,071 students enrolled. The enrollment consists of approximately 43% white students, 40% Hispanic students, 9% black students, 7% Asian students, and 1% Indian students. Mr. Philip Root taught his students utilizing the modeling approach. He has been teaching chemistry for five years. Mr. Root had two honors chemistry classes and one regular chemistry class.
Investigator 3:  

Investigator 3 in this research was Mrs. Andrea Strock. Mrs. Andrea Strock is employed at Perry High School. This school is located in the southeastern portion of the Phoenix Metropolitan area in Gilbert, Arizona in the Chandler Unified School District. Perry High School is a diverse school with approximately 1,595 students enrolled. The enrollment consists of approximately 65% White students, 18% Hispanic students, 8% Asian students, 8% African American students and 1% Indian students. Mrs. Andrea Strock taught her students utilizing the modeling approach. She has been teaching for three years. Mrs. Strock had four regular chemistry classes.

Procedure:

Permission:

The students in this study were all minors; therefore, every student that was included in this study gave their authorization provided by their signature. No students that was photographed or videotaped in this study was 18 years of age or older, therefore all these students consented to be videotaped, and gave their authorization with a signature. In addition to students providing their consent, authorization was also collected from their parents or legal guardians. If permission was not granted to include a student in the study, the results from that student were omitted, and that student was not photographed or videotaped.

Objectives:

The control teacher used a traditional approach to teaching chemistry. The energy unit was taught at the beginning of the school year between the unit on dimensional analysis but prior the unit on properties of matter. The objectives included:

- to describe the forms of energy
- to explain the law of conservation of energy
- to compare the temperature scales
- to explain the difference between temperature and heat

The students completed worksheets on energy terms and temperature scales. There was also a lecture and discussion on the law of conservation of energy. After the information was explained and reviewed, there was a unit test. Later during the school year, there were lectures and calculations on ‘thermal chemistry’. However, the concept of ‘thermal chemistry’ was never connected to the energy concept discussed at the beginning of the school year.

All classes in the treatment group, both regular and honors, were taught using the modeling approach to teaching chemistry. This unit came directly from the Arizona State University Chemistry Modeling Instruction Program. The specific units involved in this study were units 2, 3, 6 and 8. These units involve energy storage, energy transfers, and calculations involving energy. Unit 2 (Simple Particle Energy and States) introduces students to the basic
Investigators taught the energy units during the fall semester. The objectives included students being able to:

- relate observations regarding the addition of energy by warming to increased particle motion
- describe the characteristics of solids, liquids, and gases in terms of particles and their arrangement, attraction, and behavior
- recognize energy as a conserved, substance-like quantity that is always involved when a system undergoes change
- recognize that energy is stored in an object or system in several different ways (in this unit, this was restricted to kinetic and interaction energies)
- describe the ways that energy is transferred between the system and the surroundings through heating, working, and radiating
- draw energy bar graphs to account for energy storage and transfer in all sorts of changes
- identify what phase(s) are present in the various portions of a heating/cooling curve
- identify which energy storage mode is changing for the various portions of the heating/cooling curve
- sketch a heating/cooling curve that represents a situation in which a substance undergoes a change in temperature, phase, or both, to state the physical meaning of the heat of fusion and heat of vaporization for a given substance
- state the physical meaning of the heat capacity of a substance and use this factor to relate the mass and temperature changes to the energy absorbed or released during a change in temperature (with no phase change)
- distinguish between heat and temperature, and to predict the effect of the addition of energy to a system of particles in the solid or liquid state.

The sequence followed the Arizona State University Modeling Instruction Program utilizing their materials, activities, and labs. First, students distinguished between heat and temperature by watching a 5-minute video clip from the “Eureka” series or a similar video. After a discussion on this, students performed the “Icy Hot Lab” and plotted a graph of temperature versus time to observe the effect of particles in the solid and liquid state due to the addition of energy to the system. Investigator 1 then introduced the new energy storage mode similar to that used in the Physics Modeling Program. There was a whiteboarding session after the lab to assimilate the knowledge. New energy storage diagrams illustrate the students’ observations. Students completed four worksheets with whiteboarding sessions after each worksheet. There was a unit quiz and a unit test when all the information had been understood. Each investigator’s individual results address any slight modifications.
Time-Line:

Treatment:

This comparison study started during the Unit 2: Energy and States of Matter, of the Chemistry Modeling materials. The treatment began after the students had completed the unit on properties of matter but before the unit on bonding and ions. This sequence is slightly different from the traditional sequence, which has the energy unit after students learn dimensional analysis but prior to learning the states of matter.

The teacher materials explain the objectives and sequencing of the modeling method.

“In most high school science curricula, energy is addressed in a piecemeal manner with many different terms (heat, enthalpy, potential energy, kinetic energy, heat capacity) used to describe the role of energy. This seems to promote a non-coherent view of energy as something that is different in different circumstances rather than a singular view of energy as a universal element of change in both macroscopic and microscopic systems. The goal in this unit is to provide systematic development of the nature of matter and the role of energy in observable changes, beginning with phase changes. We have already established that matter is particulate; students have been guided to describe the phases of matter macroscopically and then infer particle arrangement for each phase. An attractive interaction between the particles must be inferred to account for the rigidity of solids and the fluid cohesiveness of liquids. In the previous unit, we have connected the motion of particles to kinetic energy and defined temperature as a measure of kinetic energy. In this unit, we examine what happens when sufficient energy is added to a solid to overcome the interactions between the particles. This is the point at which it is appropriate to define energy as a conserved, substance-like quantity that is always involved when a system undergoes change. The concepts of energy storage and transfer are fleshed out in class discussion.” (Arizona State University, 2007).

Assessment

Investigators administered the pre-test at the beginning of the school year to each student in all classes taught by the instructors. Investigators administered the post-test at the end of the school year to each student in all classes taught by the instructors. The Basic Energy Concept Inventory (BECI) and the Chemical Concept Inventory (CCI) was given as the pre-test, and post-test.

After the instructors completed their respective units and administered the pre-test, investigators conducted student interviews to collect qualitative data of what the students learned during the unit. Additional questions allowed students to demonstrate what they learned. Students explained their thinking as they answered these questions.
Results:

Control Group Report:

Introduction:

The instructor for the Control Group is Mr. Bryan Simon (pseudonym) from Phoenix, Arizona. He has been teaching chemistry for over 20 years. He teaches at Deer Valley High school in Glendale, Arizona. This is a “Highly Performing” school and approximately 72% of its students pass the math, reading, and writing standardized tests given by the State of Arizona each year. The sampled population for the Control Group consisted of 45 honors chemistry students from two class periods, and 73 regular chemistry students from three class periods. Although almost every student and their parents agreed to inclusion in the study, Deer Valley High School does not lock the students into the same teacher for the entire school year. Some students transferred from the treatment to the control class exposing students to both the treatment and the control. Therefore, a small segment of the initial students who transferred between classes were omitted from the data that underwent statistical analysis.

Procedure:

The Control Group students were taught utilizing the traditional method of teaching energy in chemistry using mostly lecture. After the lectures, students completed worksheets and finally took a unit test. There were demonstrations dispersed throughout the lectures and approximately three lab investigations used to enhance the lecture. At times, the few students who felt comfortable with the knowledge tutored other students. Since students’ misconceptions were not addressed, these same misconceptions were passed onto the student being tutored.

The BECI and the CCI were administered as a pretest in August, a mid-test in December, and a posttest in April. Investigators evaluated the results of the pretest and posttest at the end of the school year to assess longer term learning of the concepts.
Results:

The following figures report the distribution of scores on the selected questions from the BECI and CCI pre-test and post-test. The total number of possible questions was 15. All five classes were combined for the Control Group. On the right side of each histogram is the average score (Mean), the standard deviation of the data set (Std. Dev.) and the number of cases on the population (N).

![Control Group Pretest Scores](image1)
![Control Group Posttest Scores](image2)

Figure C1.1: Comparison of control group pretest scores to control group posttest scores

The pretest results of the control group show a fairly normal distribution. The variation among the scores, called the variance, in the pretest is smaller than that of the posttest scores. This is supported by the increase in standard deviation (a measure of variance) shown to the right of each graph above. The mean score of the group also increased from pre- to posttest, indicating that scores did improve for the control population. This increase in mean score is the basis of comparison for the performance of the three treatment groups.

As mentioned earlier, all five classes, including three regular and two core chemistry classes, were treated as one group or sample population. The reason for this was the similarity among all of the pre-test scores of the five control group classes. To verify that the classes can be treated as statistically the same, the Analysis of Variance Test (ANOVA) was used. The null hypothesis that there was no difference in the means of the pretests was tested at the 0.05 level of significance ($\alpha=0.05$). The results of the test are summarized in the following table.
Table C1.1: ANOVA: Single Factor Test for Control Group Pretests

Summary Statistics of Control Pretest Results:

<table>
<thead>
<tr>
<th>Groups</th>
<th>Count</th>
<th>Sum</th>
<th>Average</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control: 1st Hour</td>
<td>22</td>
<td>65</td>
<td>2.954545</td>
<td>1.664502165</td>
</tr>
<tr>
<td>Control: 2nd Hour</td>
<td>23</td>
<td>55</td>
<td>2.391304</td>
<td>1.339920949</td>
</tr>
<tr>
<td>Control: 3rd Hour</td>
<td>30</td>
<td>75</td>
<td>2.5</td>
<td>2.120689655</td>
</tr>
<tr>
<td>Control: 5th Hour</td>
<td>24</td>
<td>60</td>
<td>2.5</td>
<td>1.652173913</td>
</tr>
<tr>
<td>Control: 7th Hour</td>
<td>19</td>
<td>47</td>
<td>2.473684</td>
<td>1.040935673</td>
</tr>
</tbody>
</table>

ANOVA Statistical Test Results:

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P-value</th>
<th>F critical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>4.415097</td>
<td>4</td>
<td>1.103774</td>
<td>0.682798159</td>
<td>0.6052938</td>
<td>2.451988</td>
</tr>
<tr>
<td>Within Groups</td>
<td>182.6696</td>
<td>113</td>
<td>1.616546</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>187.0847</td>
<td>117</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results of the ANOVA test give a statistical test value (F), which is smaller than the critical value (F_{critical}). Specifically, 0.068 < 2.45. This means that the null hypothesis cannot be rejected and should be accepted. Statistically speaking, the five different classes scored essentially the same on the pretest, therefore, are treated as a single sample population.

The histograms above also show a fairly uniform distribution of scores for the post-test. The ANOVA test was again used to test the null hypothesis that there is no difference in the means of the posttest scores of the control. Again, the test result of the F value was less than the critical value (0.83 < 2.45). Again, the null hypothesis should be accepted meaning that all 5 classes achieved similar results on the posttest and can be treated as a single population in further analysis. These results are summarized in the table below.
Table C1.2: ANOVA: Single Factor Test for Control Group Posttests

Summary Statistics of Control Posttest Results:

<table>
<thead>
<tr>
<th>Groups</th>
<th>Count</th>
<th>Sum</th>
<th>Average</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control: 1(^{st}) Hour</td>
<td>22</td>
<td>105</td>
<td>4.772727273</td>
<td>4.279221</td>
</tr>
<tr>
<td>Control: 2(^{nd}) Hour</td>
<td>23</td>
<td>106</td>
<td>4.608695652</td>
<td>3.339921</td>
</tr>
<tr>
<td>Control: 3(^{rd}) Hour</td>
<td>30</td>
<td>153</td>
<td>5.1</td>
<td>9.196552</td>
</tr>
<tr>
<td>Control: 5(^{th}) Hour</td>
<td>24</td>
<td>127</td>
<td>5.291666667</td>
<td>8.476449</td>
</tr>
<tr>
<td>Control: 7(^{th}) Hour</td>
<td>19</td>
<td>77</td>
<td>4.052631579</td>
<td>3.163743</td>
</tr>
</tbody>
</table>

ANOVA Statistical Test Results:

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P-value</th>
<th>F\text{ critical}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>19.95071</td>
<td>4</td>
<td>4.987676524</td>
<td>0.826467</td>
<td>0.510982714</td>
<td>2.451988</td>
</tr>
<tr>
<td>Within Groups</td>
<td>681.9476</td>
<td>113</td>
<td>6.034934504</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>701.8983</td>
<td>117</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Did the control group show significant improvement from pre-test to post-test? This can be ascertained through the use of a paired Student’s t-test, which compares the means of paired sets of data (such as pre-tests and post-tests). The results of this test are shown in the table below.

Table C1.3: t-Test: Paired Two Sample for Means

<table>
<thead>
<tr>
<th>Testing the improvement of the Control Group</th>
<th>Control Posttest</th>
<th>Control Pretest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>4.813559322</td>
<td>2.559322034</td>
</tr>
<tr>
<td>Variance</td>
<td>5.999130813</td>
<td>1.599014921</td>
</tr>
<tr>
<td>Observations</td>
<td>118</td>
<td>118</td>
</tr>
<tr>
<td>Hypothesized Mean Difference</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Degrees of Freedom</td>
<td>117</td>
<td></td>
</tr>
<tr>
<td>t Statistical</td>
<td>10.12481666</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) two-tail</td>
<td>1.10094E-17</td>
<td></td>
</tr>
<tr>
<td>t Critical two-tail</td>
<td>1.980447532</td>
<td></td>
</tr>
</tbody>
</table>

The results of the paired t-test give a statistical ‘t’ value that is greater than the critical value (10.12 > 1.98). This means that the mean score on the posttest was higher than the mean score on the pretest for the control group.
The above item analysis shows how students improved from pretest to posttest as a function of the percentage of correct responses. Moderate gains are seen in most questions, except for #3 from the BECI (which was answered poorly in general by the whole tested population). Specific comparisons to the control group are made in each section of the investigators’ reports.

Conclusion:

The results presented here are very similar to what would be expected by students taught using traditional methods. The data above also suggested that the test populations improved their scores from the pretest to the posttest. The mean values of the posttest’s values are larger than the pretest’s value. However, variance in each data set also increased as evidence by the wider curve in the posttest. Even though the Control Group students improved, they did not seem to increase their score as much as the students in the treatment did. Students did not seem to have a grasp of the basic concepts that are lectured to them throughout the school year. This will be shown through further analysis of the investigators and their treatment groups. This analysis will compare the improvements of the treatment group to the improvements of the control group.
Investigator 1 Report:

Introduction:

Investigator 1 is Mrs. Donna Blakeney from Phoenix, Arizona. She has been teaching chemistry for 11 years. She teaches at Deer Valley High school in Glendale, Arizona. This is a “Highly Performing” school and approximately 72% of its students pass the math, reading, and writing standardized tests given by the State of Arizona each year. The sampled population for Investigator 1, during the pilot study conducted in the 2007 – 2008 school year, consisted of 49 honors chemistry students from two class periods, and 63 regular chemistry students from three class periods. The sampled population for Investigator 1, during the 2nd year the study conducted in the 2008 – 2009 school year, consisted of 47 honors chemistry students from two class periods, and 27 regular chemistry students from two class periods. Although almost every student and their parents agreed to inclusion in the study, during both years, Deer Valley High School does not lock the students into the same teacher for the entire school year. Some students transferred from the treatment to the control class exposing students to both the treatment and the control. Therefore, Investigator 1 removed a small portion of the initial students who transferred between classes from the data before statistical analysis.

Procedure:

There was little variation from the Chemistry modeling as developed at Arizona State University. Investigator 1 enrolled in the 1st semester Chemistry Modeling Workshop in the summer of 2007. Investigator 1 implemented the knowledge learned in that workshop as presented using the worksheets, labs, and activities in the same order and manner as demonstrated during the workshop. Throughout the school year, students were involved in lab investigations that were not “cook book” labs. They developed procedures after class discussions on different phenomenon. Students analyzed their data and demonstrated that the numbers they were collecting actually mean something. Students verbalize the concepts discovered during labs to their group and the class through white boarding sessions. Students discussed any misconceptions and improved their ways of thinking. It appeared as the school year continued, students better discussion skills and improved their knowledge of the basic concepts in chemistry; specifically energy.

Investigator 1 administered the BECI and the CCI as a pretest in August, a midtest in December, and a posttest in April. Investigator 1 evaluated the results of the pretest and posttest at the end of the school year to assess longer term learning of the concepts. After administering all tests of the BECI and the CCI, Investigators decided to evaluate only 15 items of these tests.

During the 2nd year of this study, the same materials and testing were completed. However, the 2nd year allowed for additional time spent developing the modeling materials to include additional activities that complement those materials presented in the original workshop. As a 2nd year modeler, Investigator 1 was more comfortable with the materials. The presented
information was more coherent throughout the school year. Students benefited by this investigators increased knowledge and skills, and therefore increased their knowledge and skills.

Results:

Pilot Study:

The first year of this study was conducted as a pilot study. Data from the control group and the treatment group were compared to assess if this study warranted further investigation. The pilot study showed enough improvement in students’ knowledge that a 2nd year was necessary to confirm this. The following figures report the distribution of scores on the selected questions from the BECI and the CCI pretests and posttests. The total number of possible questions was 15. On the right side of each histogram is the average score (Mean), the standard deviation of the data set (Std. Dev.) and the number of cases on the population (N).

![Figure 1.1: Comparison of Investigator 1 pretest scores to Investigator 1 posttest scores in 2007](image1.png)

The pretest results of the pilot study show a normal distribution. All classes, regular and honors, were treated as one group or sample population. The reason for this was the similarity among all of the pretest scores. To verify that these classes can be treated as statistically the same, the Analysis of Variance Test (ANOVA) was used. The null hypothesis, that there was no difference in the means of pretest scores, was tested at the 0.05 level of significance (α=0.05). However, the fact that the posttest distribution appears to be bimodal (two peaks) might indicate that the honors chemistry classes and the regular chemistry class may be different statistical group. Even with the bimodal distribution in the posttest scores, the variation among the scores,
called the variance, in the pretest is smaller than that of the posttest scores. This is supported by the increase in the standard deviation shown to the right of each graph above. The pretest mean score was 2.47 and the posttest mean score was 5.88. This increase in the mean score indicates that scores did improve for the pilot study group. The results of the test are summarized in the following table.

Table 1.1: ANOVA: Single Factor Test for Treatment (2007) and Control Group Pretest Scores

<table>
<thead>
<tr>
<th>Groups</th>
<th>Count</th>
<th>Sum</th>
<th>Average</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investigator 1: 1st Hour</td>
<td>29</td>
<td>80</td>
<td>2.75621</td>
<td>1.903941</td>
</tr>
<tr>
<td>Investigator 1: 2nd Hour</td>
<td>20</td>
<td>47</td>
<td>2.35</td>
<td>1.607895</td>
</tr>
<tr>
<td>Investigator 1: 3rd Hour</td>
<td>21</td>
<td>54</td>
<td>2.57143</td>
<td>3.357143</td>
</tr>
<tr>
<td>Investigator 1: 6th Hour</td>
<td>15</td>
<td>34</td>
<td>2.26667</td>
<td>2.638095</td>
</tr>
<tr>
<td>Investigator 1: 7th Hour</td>
<td>27</td>
<td>62</td>
<td>2.29629</td>
<td>2.139601</td>
</tr>
<tr>
<td>Control: 1st Hour</td>
<td>22</td>
<td>65</td>
<td>2.95455</td>
<td>1.664502</td>
</tr>
<tr>
<td>Control: 2nd Hour</td>
<td>23</td>
<td>55</td>
<td>2.39130</td>
<td>1.339921</td>
</tr>
<tr>
<td>Control: 3rd Hour</td>
<td>30</td>
<td>75</td>
<td>2.5</td>
<td>2.12069</td>
</tr>
<tr>
<td>Control: 5th Hour</td>
<td>24</td>
<td>60</td>
<td>2.5</td>
<td>1.652174</td>
</tr>
<tr>
<td>Control: 7th Hour</td>
<td>19</td>
<td>47</td>
<td>2.47368</td>
<td>1.040936</td>
</tr>
</tbody>
</table>

ANOVA Statistical Test Results:

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>$F$</th>
<th>P-value</th>
<th>$F_{critical}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>9.19462142</td>
<td>9</td>
<td>1.021625</td>
<td>0.527308</td>
<td>0.853915</td>
<td>1.922614</td>
</tr>
<tr>
<td>Within Groups</td>
<td>426.2358134</td>
<td>220</td>
<td>1.937436</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>435.4304348</td>
<td>229</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The ANOVA test was used at the 0.05 level of significance to test the null hypothesis that the mean pretest scores of the treatment and control groups are the same. The results of the ANOVA test give a statistical test value (F), which is smaller than the critical value ($F_{critical}$). Specifically, $0.527 < 1.92$, this means that the null hypothesis cannot be rejected and should be accepted. Statistically speaking, the different classes in the treatment and control groups scored essentially the same on the pretest. This means that both groups are starting out at the same baseline score for the pretest. Thus, posttest scores for both groups can be compared directly using statistical tools such as a t-test.

To show that the Treatment Group posttest scores improved over the pretest scores, a paired two-sample t-test was conducted. This t-test compares data sets that are paired, such as the pretest and posttest scores. The null hypothesis, tested at the 0.05 level of significance, was that the posttest scores are the same as the pretest scores. The results are summarized below.
A Comparison Study for Teaching Energy 23

Table 1.2: t-Test: Paired Two Sample for Means
Tested group: 2007 Treatment Group

<table>
<thead>
<tr>
<th></th>
<th>Posttest</th>
<th>Pretest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>5.875</td>
<td>2.473214286</td>
</tr>
<tr>
<td>Variance</td>
<td>4.885135135</td>
<td>2.233510296</td>
</tr>
<tr>
<td>Observations</td>
<td>112</td>
<td>112</td>
</tr>
<tr>
<td>Hypothesized Mean Difference</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Degrees of Freedom</td>
<td>111</td>
<td></td>
</tr>
<tr>
<td>t Stat</td>
<td>13.89705386</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) two-tail</td>
<td>4.79193E-26</td>
<td></td>
</tr>
<tr>
<td>t Critical two-tail</td>
<td>1.981566695</td>
<td></td>
</tr>
</tbody>
</table>

The results of the paired two-sample t-test give a statistical t value that is greater than the critical t value (13.90 > 1.98). The null hypothesis that the mean posttest and pretest scores of the treatment groups are the same is rejected. The treatment group scored significantly better on the posttest.

The next step in the analysis then was to determine if the Pilot Study showed significant improvement in posttest scores of the treatment group as compared to the control group. This can be achieved using a student’s two sample t-test assuming unequal variances. This test compares the means of two sets of data (such as the posttests of the treatment and control groups). The unequal variances test was used since the variance in the scores of the treatment and control group were, in fact, unequal. The results of this test are shown in table 1.3 below.

Table 1.3: Results of t-Test: Two-Sample Assuming Unequal Variances
Tested Groups: Investigator 1 2007 vs. Control Group

<table>
<thead>
<tr>
<th></th>
<th>Investigator 1 2007 Posttest</th>
<th>Control Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>5.875</td>
<td>4.813559322</td>
</tr>
<tr>
<td>Variance</td>
<td>4.885135135</td>
<td>5.999130813</td>
</tr>
<tr>
<td>Observations</td>
<td>112</td>
<td>118</td>
</tr>
<tr>
<td>Hypothesized Mean Difference</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Degrees of freedom</td>
<td>227</td>
<td></td>
</tr>
<tr>
<td>t Statistical</td>
<td>3.453645867</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) two-tail</td>
<td>0.000659733</td>
<td></td>
</tr>
<tr>
<td>t Critical two-tail</td>
<td>1.970469461</td>
<td></td>
</tr>
</tbody>
</table>

The results of the two sample t-test give a statistical t value that is greater than the critical t value (3.45 > 1.97). The null hypothesis that the mean posttest scores of the 2007 treatment and control groups are the same is rejected. This supports that the mean score on the posttest for the pilot study was significantly higher than the mean score on the pretest for the control group.
Since both groups started out at the same level, the t-test demonstrates that the students of the treatment group did better than the control group to a statistically significant degree. The significance can be quantified through the effect size. “Effect size is a measure of the strength of the relationship between two variables in a statistical population.” (Effect) The effect size basically shows how effective the treatment was over the control or how significant the increase in the treatment compared to the control. Using the means and standard deviations of the treatment and control group, the Cohen’s d value can be calculated which states how much of an effect the treatment had on the test population over the control group. For Cohen's $d$, an effect size of 0.2 to 0.3 is a "small" effect, 0.4 to 0.5 is a "medium" effect, and above 0.8 is a "large" effect. (Effect) For the above data, the Cohen’s d value is equal to 0.455. This is as a medium effect size over the control group, meaning that the mean score of the treatment group is in about the 68th percentile of the control group scores. (Coldarci et. al.)

Table 1.4: Effect size statistics

<table>
<thead>
<tr>
<th>Metric</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cohen's d value</td>
<td>0.454999744</td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>0.221831706</td>
</tr>
</tbody>
</table>

Figure 1.2: Item Analysis of selected BECI and CCI questions for Investigator 1 in 2007

Questions pertaining to the energy standards were selected for analysis in the study. The percentage of correct responses on the pretest and posttest are summarized in the graph above. All questions show improvement as a result of the treatment. Certain questions show large improvement, such as BECI numbers 3, 8, 11, and 12, and CCI number 9, 16, 17. Others, such as BECI numbers 7, 10, and 13, showed improvement but were still answered poorly by the students. Reasons for this will be discussed in the conclusion.
2\textsuperscript{nd} Year of Study Compared to Control Group:

![Histograms showing pretest and posttest scores](image)

Figure 1.3: Comparison of Investigator 1 pretest scores to Investigator 1 posttest scores in 2008

The pretest results of the 2\textsuperscript{nd} year in this study for Investigator 1 show a normal distribution. All classes, regular and honors, were treated as one group or sample population. The reason for this was the similarity among all of the pretest scores. To verify that these classes can be treated as statistically the same, the Analysis of Variance Test (ANOVA) was used. The null hypothesis, that there was no difference in the means of pretest scores, was tested at the 0.05 level of significance ($\alpha=0.05$). The variation among the scores, called the variance, in the pretest is smaller than that of the posttest scores. The increased variation in the data is quantitatively shown in the standard deviation values that increase from pretest to posttest. The pretest mean score was 2.7 and the posttest mean score was 8.66. This increase in the mean score indicates that scores did improve for the 2\textsuperscript{nd} year treatment group for Investigator 1. The results of the test are summarized in the following table.
Table 1.5: ANOVA: 
Single Factor Test for Treatment (2008) and Control Group Pretest Scores

<table>
<thead>
<tr>
<th>Groups</th>
<th>Count</th>
<th>Sum</th>
<th>Average</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investigator 1: 2nd Hour</td>
<td>22</td>
<td>61</td>
<td>2.772727273</td>
<td>1.898268</td>
</tr>
<tr>
<td>Investigator 1: 3rd Hour</td>
<td>25</td>
<td>68</td>
<td>2.72</td>
<td>2.376667</td>
</tr>
<tr>
<td>Investigator 1: 5th Hour</td>
<td>9</td>
<td>19</td>
<td>2.111111111</td>
<td>1.611111</td>
</tr>
<tr>
<td>Investigator 1: 7th Hour</td>
<td>18</td>
<td>58</td>
<td>3.222222222</td>
<td>1.477124</td>
</tr>
<tr>
<td>Control: 1st Hour</td>
<td>22</td>
<td>65</td>
<td>2.954545455</td>
<td>1.664502</td>
</tr>
<tr>
<td>Control: 2nd Hour</td>
<td>23</td>
<td>55</td>
<td>2.391304348</td>
<td>1.339921</td>
</tr>
<tr>
<td>Control: 3rd Hour</td>
<td>30</td>
<td>75</td>
<td>2.5</td>
<td>2.12069</td>
</tr>
<tr>
<td>Control: 5th Hour</td>
<td>24</td>
<td>60</td>
<td>2.5</td>
<td>1.652174</td>
</tr>
<tr>
<td>Control: 7th Hour</td>
<td>19</td>
<td>47</td>
<td>2.473684211</td>
<td>1.040936</td>
</tr>
</tbody>
</table>

ANOVA Statistical Test Results:

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P-value</th>
<th>F critical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>14.34338187</td>
<td>8</td>
<td>1.792922734</td>
<td>1.033163</td>
<td>0.412705</td>
<td>1.98929</td>
</tr>
<tr>
<td>Within Groups</td>
<td>317.5732848</td>
<td>183</td>
<td>1.735373141</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>331.9166667</td>
<td>191</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The ANOVA test was again used at the 0.05 level of significance to test the null hypothesis that the mean pretest scores of the treatment and control groups are the same. The results of this ANOVA test give a statistical test value (F), which is smaller than the critical value (F critical). Specifically, 1.033 < 1.989, which means that the null hypothesis cannot be rejected and should be accepted. Statistically speaking, the different classes in the 2008 treatment and the control group scored essentially the same on the pretest. This means that both groups are starting out at essentially the same baseline pretest scores for the pretest. Thus, posttest scores for both groups can, again be, compared directly using statistical tools such as a t-test.
A paired two-sample t-test was conducted to show that the treatment group posttest scores improved over the pretest scores. The null hypothesis, tested at the 0.05 level of significance, was that the posttest scores are the same as the pretest scores. The results are summarized below.

### Table 1.6: t-Test: Paired Two Sample for Means

<table>
<thead>
<tr>
<th>Tested group: 2008 Treatment group</th>
<th>Posttest</th>
<th>Pretest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>8.662162162</td>
<td>2.783783784</td>
</tr>
<tr>
<td>Variance</td>
<td>14.1993706</td>
<td>1.952610144</td>
</tr>
<tr>
<td>Observations</td>
<td>74</td>
<td>74</td>
</tr>
<tr>
<td>Hypothesized Mean Difference</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Degrees of Freedom</td>
<td>73</td>
<td></td>
</tr>
<tr>
<td>t Statistical</td>
<td>13.18414021</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) two-tail</td>
<td>5.39304E-21</td>
<td></td>
</tr>
<tr>
<td>t Critical two-tail</td>
<td>1.992997097</td>
<td></td>
</tr>
</tbody>
</table>

The results of the paired two-sample t-test give a statistical t value that is greater than the critical t value (13.18 > 1.99). The null hypothesis that the mean posttest and pretest scores of the treatment groups are the same is rejected. The treatment group scored significantly better on the posttest.

To illustrate that the posttest scores of the treatment group are significantly better than the posttest scores of the control, a two-sample t-test is again used. The t-test (level of significance 0.05) was used to test the null hypothesis that the mean posttest scores of the 2008 treatment and control groups are the same.

### Table 1.7: Results of t-Test: Two-Sample Assuming Unequal Variances

<table>
<thead>
<tr>
<th>Tested Groups: Investigator 1 2008 vs. Control Group</th>
<th>Investigator 1 2008 Posttest</th>
<th>Control Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>8.662162162</td>
<td>4.813559322</td>
</tr>
<tr>
<td>Variance</td>
<td>14.1993706</td>
<td>5.999130813</td>
</tr>
<tr>
<td>Observations</td>
<td>74</td>
<td>118</td>
</tr>
<tr>
<td>Hypothesized Mean Difference</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Degrees of freedom</td>
<td>112</td>
<td></td>
</tr>
<tr>
<td>t Statistical</td>
<td>7.811729612</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) two-tail</td>
<td>3.29844E-12</td>
<td></td>
</tr>
<tr>
<td>t Critical two-tail</td>
<td>1.981371752</td>
<td></td>
</tr>
</tbody>
</table>

The results of this two sample t-test also give a statistical t value that is greater than the critical t value (7.81 > 1.98). The null hypothesis that the mean posttest scores of the 2008 treatment and control groups are the same is rejected. Since both groups started out at the same baseline level, the t-test demonstrates that the students of the treatment group did better than the control group to a statistically significant degree. As previously stated, “Effect size is a measure of the strength
of the relationship between two variables in a statistical population.” (Effect) The effect size basically shows how effective the treatment was over the control or how significant the increase in the treatment compared to the control. Using the means and standard deviations of the treatment and control group, the Cohen’s d value can be calculated which states how much of an effect the treatment had on the test population over the control group. For Cohen’s $d$, an effect size of 0.2 to 0.3 is a "small" effect, 0.4 to 0.5 is a "medium" effect, and above 0.8 is a "large" effect. (Effect) The effect size expressed through the Cohen’s d value for the 2008 treatment group compared to the control is equal to 1.21. This is a large effect size over the control group, meaning that the mean score of the treatment group is in about the 88$^{th}$ percentile of the control group scores. (Coldarci et. al.)

<table>
<thead>
<tr>
<th>Table 1.8: Effect size statistics:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cohen’s d value</td>
</tr>
<tr>
<td>Pearson Correlation</td>
</tr>
</tbody>
</table>

![Figure 1.4: Item Analysis of selected BECI and CCI questions for Investigator 1 in 2008](image)

Questions pertaining to the energy standards were selected for analysis in the study. The percentage of correct responses on the pretest and posttest are summarized in the graph above for the 2$^{nd}$ year for Investigator 1. All questions show improvement as a result of the treatment. In addition, the percentage correct on many of the items increases to over 50%, which was seldom seen in the same chart for the control group. Certain questions show large improvement, such as BECI numbers 3, 4, 7, 10, 11, 12, and 22 and CCI number 16. Reasons for this will be discussed in the conclusion. This graph appears to have similar increases as those seen during the Pilot.
Study in 2007. In addition, the increases of the 2008 Treatment Group were greater than that of the 2007 Pilot Group.

2\textsuperscript{nd} Year of Study Compared to Pilot Study:

![Histograms showing comparison of pretest scores](image)

Figure 1.5: Comparison of Investigator 1 2007 scores to Investigator 1 2008 scores

The pretest results of the Pilot Study of 2007 and the 2\textsuperscript{nd} year Treatment Group of 2008 show a normal distribution. All classes, regular and honors, were treated as one group or sample population. The reason for this was the similarity among all of the pretest scores. To verify that these classes can be treated as statistically the same, the Analysis of Variance Test (ANOVA) was used. The null hypothesis, that there was no difference in the means of pretest scores, was
tested at the 0.05 level of significance (\(\alpha=0.05\)). The variation among the scores, called the variance, in the pretest is smaller than that of the posttest scores in both tests. This is supported by the increase in the standard deviation shown to the right of each graph above. The pretest mean score for the Pilot Study was 2.47 and for the 2\(^{nd}\) year Treatment Group was 2.7. The posttest mean score for the Pilot Study was 5.88 and for the 2\(^{nd}\) year Treatment Group was 8.66. This increase in the mean score indicates that scores did improve for both the Pilot Group, as well as the 2\(^{nd}\) year Treatment Group. To continue, the 2\(^{nd}\) year Treatment Group had a larger increase (69\%) over the Pilot Group (58\%). This demonstrates that a 2\(^{nd}\) year modeler has improved modeling skills over a 1\(^{st}\) year modeler as evidenced by the larger increase. The results of the test are summarized in the following table.

### Table 1.9: ANOVA:
Single Factor Test for Treatment (2007) and Treatment (2008) Posttest Scores

<table>
<thead>
<tr>
<th>Groups</th>
<th>Count</th>
<th>Sum</th>
<th>Average</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investigator 1: 1(^{st}) Hour (2007)</td>
<td>29</td>
<td>80</td>
<td>2.758621</td>
<td>1.903941</td>
</tr>
<tr>
<td>Investigator 1: 2(^{nd}) Hour (2007)</td>
<td>20</td>
<td>47</td>
<td>2.35</td>
<td>1.607895</td>
</tr>
<tr>
<td>Investigator 1: 3(^{rd}) Hour (2007)</td>
<td>21</td>
<td>54</td>
<td>2.571429</td>
<td>3.357143</td>
</tr>
<tr>
<td>Investigator 1: 6(^{th}) Hour (2007)</td>
<td>15</td>
<td>34</td>
<td>2.26667</td>
<td>2.638095</td>
</tr>
<tr>
<td>Investigator 1: 7(^{th}) Hour (2007)</td>
<td>27</td>
<td>62</td>
<td>2.296296</td>
<td>2.139601</td>
</tr>
<tr>
<td>Investigator 1: 2(^{nd}) Hour (2008)</td>
<td>22</td>
<td>61</td>
<td>2.772727</td>
<td>1.898268</td>
</tr>
<tr>
<td>Investigator 1: 3(^{rd}) Hour (2008)</td>
<td>25</td>
<td>68</td>
<td>2.72</td>
<td>2.376667</td>
</tr>
<tr>
<td>Investigator 1: 5(^{th}) Hour (2008)</td>
<td>9</td>
<td>19</td>
<td>2.111111</td>
<td>1.611111</td>
</tr>
<tr>
<td>Investigator 1: 7(^{th}) Hour (2008)</td>
<td>18</td>
<td>58</td>
<td>3.222222</td>
<td>1.477124</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P-value</th>
<th>(F_{\text{critical}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>16.28826</td>
<td>8</td>
<td>2.036033</td>
<td>0.952197</td>
<td>0.475164</td>
<td>1.991033</td>
</tr>
<tr>
<td>Within Groups</td>
<td>378.4698</td>
<td>177</td>
<td>2.138247</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>394.7581</td>
<td>185</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The ANOVA test was used at the 0.05 level of significance to test the null hypothesis that the mean pretest scores of the 2007 and 2008 treatment groups are the same. The results of this ANOVA test give a statistical test value (F), which is smaller than the critical value (\(F_{\text{critical}}\)). Specifically, 0.952 < 1.991, this means that the null hypothesis cannot be rejected and should be accepted. Statistically speaking, even though the student populations were from different classes, spanning two different school years, they scored essentially the same on the pretest. This means that both groups are essentially starting out at the same baseline score for the pretest. Thus, posttest scores for both groups can again be compared directly using statistical tools such as a t-test.
Table 1.10: t-Test: Two-Sample Assuming Unequal Variances
Tested Groups: Investigator 1 2008 vs. Investigator 1 2007

<table>
<thead>
<tr>
<th></th>
<th>Investigator 1 2008 Posttest</th>
<th>Investigator 1 2007 Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>8.662162</td>
<td>5.875</td>
</tr>
<tr>
<td>Variance</td>
<td>14.19937</td>
<td>4.885135</td>
</tr>
<tr>
<td>Observations</td>
<td>74</td>
<td>112</td>
</tr>
<tr>
<td>Hypothesized Mean Difference</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Degrees of freedom</td>
<td>106</td>
<td></td>
</tr>
<tr>
<td>t Statistical</td>
<td>5.743362</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) two-tail</td>
<td>8.96E-08</td>
<td></td>
</tr>
<tr>
<td>t Critical two-tail</td>
<td>1.982597</td>
<td></td>
</tr>
</tbody>
</table>

The results of the two sample t-test give a statistical t value that is greater than the critical t value (5.74 > 1.98). The null hypothesis that the mean posttest scores of the 2007 and 2008 treatment groups are the same is rejected. Since both groups started out at the same baseline level, the t-test demonstrates that the students of the 2008 treatment group did better than the 2007 treatment group to a statistically significant degree. The effect size expressed through the Cohen’s d value for the 2008 treatment group compared to the control is equal to 0.902. Statistical tables classify this as a large effect size over the control group.

Table 1.11: Effect Size statistics:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cohen's d value</td>
<td>0.902271</td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td>0.411225</td>
</tr>
</tbody>
</table>

Conclusion:

After analyzing the data, it would appear that the modeling approach to teaching energy is more effective to teaching using the traditional approach. Investigator 1’s Pilot Group of 2007 scored higher on the BECI and CCI tests than that of the Control Group. The test scores for the 2nd year Treatment Group of 2008 are significantly higher than that of the Control Group. This not only supports the hypothesis that the modeling approach is superior to teaching energy than the traditional approach but it also shows support that as modelers improve their own skills, their students’ knowledge also increases.

Students in the 2nd year of this study showed a much more insightful understanding of energy than either the Control Group or the Pilot Group. Through the use of Socratic questioning skills learned before the Pilot Study and practiced during the Pilot Study, students of
the 2nd year of this study were able to process their own thinking much better. This is shown in a larger increase on the mean of the posttest scores from the 2nd year Treatment Group than the Pilot Group. During white boarding sessions and group discussions, students seemed to explain their knowledge at a higher level. Students in the 2nd year also, exhibited better questioning skills of others when they did not grasp concepts. Throughout the 2nd year of this study, Investigator 1 modeled better Socratic questioning and therefore students eventually learned these techniques and were able to question each other more efficiently. During the debriefing of investigations and worksheets, students did most of the questioning of each other in the 2nd year of this study.

During item analysis of the selected test questions, it appears that students in the both the Pilot Study and the 2nd year Treatment Group had larger gain on certain questions. This demonstrates that the modeling approach helps in student understanding of these items in particular. Question 3 on the BECI test asks how you would keep your food hot or cold; wrap it in an insulator or a conductor. Throughout the modeling materials students utilize particle diagrams to show what is happening at the molecular level. Students’ knowledge of what is happening in the molecular level of an insulator versus a conductor shows in the number of correct answers on that item. Question 11 on the BECI test asks what happens to the energy of batteries after the light of a flashlight goes out. Through many discussions that energy is transferred and showing this concept in multiple representations, students realize that energy is only transferred and is not “used up”. Unfortunately, not all test items had such large increases between the pretest scores and the posttest scores. For example, number 10 on the BECI, asks students to explain their answer about how cold wood and steel would feel after taking them out of the freezer. Many students answered this wrong in the Pilot Study, saying that steel absorbs cold better than wood does. On the other hand, students in the 2nd year Treatment Group did a much better answering this question. This could be attributed by improving my questioning skills and by realizing some of the misconceptions students come to the chemistry class with. Finally, there were items on both exams that all populations; the Control Group, the Pilot Group, and the 2nd year Treatment Group had difficulty. Additional study should be done to see if there are problems with these test items.

The purpose of this comparison study was to test our hypothesis that teaching energy utilizing the modeling approach is superior to the traditional approach. Based on the data, there is statistic evidence to support this. However, since some of the gains were not as substantial as initially expected, there is further investigation that must be completed. A larger population is necessary and further testing should be done by means of future studies.
Investigator 2 Report:

Introduction:

Investigator 2 taught chemistry in a suburb of Phoenix, Chandler, Arizona. The sampled population for investigator 2 consisted of 52 honors chemistry students from two class periods, and 13 regular chemistry students from one class period. The reason not all honors and regular chemistry students were included is that students are allowed to change classes at the semester and do not keep the same teacher all year. Therefore, a small segment of the initial student population left at semester and some new students entered. In consideration of population issues, students who dropped or added at semester were omitted from the data that underwent statistical analysis. In addition, investigator two taught two Advanced Placement/International Baccalaureate Chemistry courses. These students were not studied since they had already completed a first year chemistry course and would not be receiving the same treatment as the regular and honors chemistry classes.

Implementation:

Both regular and honors chemistry students were taught using the modeling approach to teaching chemistry. Due to site-based issues within the department, the order of instruction was changed slightly. Modeling Chemistry unit 2 introduces students to Kinetic Molecular Theory and the common gas laws. This unit helps further establish a particle model for matter that helps set up Unit 3, which is the main introduction to Matter & Energy concepts. The treatment for investigator 2 used the introductory materials from unit 2 to discuss particle behavior and the differences between solids, liquids, and gases. Then, instead of completing the unit with gas behavior and gas law relationships, the investigator proceeded to Unit 3, which focuses on matter & energy through study of phase change. Gas laws and relationships were taught in the second semester. Units 6 and 8, which emphasize energy involved in chemical change, were also taught in the second semester. Aside from these changes in order, the same major labs, videos, and activities were conducted, similar to the other investigators. White boarding sessions were consistently used throughout all units, including units 2 and 3, which focused specifically on energy concepts.

Procedure:

In order to assess the results of the modeling treatment, investigator 2 administered the same tests that were given to the control group: the chemical concept inventory (CCI) and the basic energy concept inventory (BECI). Both inventories were given as a pre-test during the first week of school. The inventories were given as a mid-test during the first semester after students completed the Unit 3 energy materials. Finally, the inventories were given a third time as a post-test towards the end of the second-semester. The pre- and post- tests were used in data analysis to assess longer term learning of the concepts.
Results Part 1: Descriptive Statistics

The following figures report the distribution of scores on the selected questions from the BECI and CCI pre- and post-tests. The total number of possible questions was 15. On the right side of each histogram is the mean score (Mean), standard deviation of the data set (Std. Dev.) and number of cases in the population (N).

The results of the pretests for the regular chemistry sample show a fairly simple distribution, slightly skewed to the right. There is little variance in the data as students scored in the 1-3 range. The posttest showed much more variation in student score. This increased variance is shown through the increase in the standard deviation between the data sets. The posttest also appears that it might be bimodal, but the spike at 1 correct response is only a frequency of 2. The small number of this population makes the distribution more difficult to normalize.

The results of the pretests for the honors chemistry sample show a fairly simple distribution, slightly skewed to the right. There is little variance in the data as students scored in the 1-3 range. The posttest showed much more variation in student score. This increased variance is shown through the increase in the standard deviation between the data sets. The posttest also appears that it might be bimodal, but the spike at 1 correct response is only a frequency of 2. The small number of this population makes the distribution more difficult to normalize.
The results of the honors pretests also show a distribution of scores in the low number correct range. The skew in this data is more to the left towards fewer correct answers. The posttest on the right shows how the mean scores improved as the normalized curve shifted to the right. There is a slight increase in the variance of student scores since some students still answered few questions correctly, while many more answered a larger number correctly. The increase in variance is shown in the increase of standard deviation between the two data sets. The scores in figures 2.3 and 4 fit 2.4 a normalized curve better, likely due to the larger sample size of this population.

Results Part 2: Inferential Statistics

The first task in interpreting the data is analyzing the pretest scores of the treatment population. The null hypothesis that the pretest scores for the one regular and two honors class are the same was tested at a 0.05 level of significance (α=0.05) using the Analysis of Variance test (ANOVA). While the pretest means are different, the ANOVA test results in a statistical test value (F) that is smaller than the critical value (F critical), specifically 1.01<4.03. The result is that the null hypothesis is accepted: the means of the 3 classes cannot be considered statistically different, even though one class was regular and the other two were honors. Both classes started at the same statistical baseline score. The statistical test data and summary are included in table 2.1 below.

Table 2.1: ANOVA: Single Factor Test for Investigator 2 Pretests

<table>
<thead>
<tr>
<th>Groups</th>
<th>Count</th>
<th>Sum</th>
<th>Average</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investigator 2: 0 Hour</td>
<td>13</td>
<td>30</td>
<td>2.307692308</td>
<td>0.564102564</td>
</tr>
<tr>
<td>Investigator 3: 3rd Hour</td>
<td>23</td>
<td>70</td>
<td>3.043478261</td>
<td>2.407114625</td>
</tr>
<tr>
<td>Investigator 3: 4th Hour</td>
<td>29</td>
<td>103</td>
<td>3.551724138</td>
<td>3.97044335</td>
</tr>
</tbody>
</table>

ANOVA Statistical Test Summary:

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P-value</th>
<th>F critical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>3.31337216</td>
<td>1</td>
<td>3.31337216</td>
<td>1.009380872</td>
<td>0.319888751</td>
<td>4.034309546</td>
</tr>
<tr>
<td>Within Groups</td>
<td>164.1289355</td>
<td>50</td>
<td>3.282578711</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>167.4423077</td>
<td>51</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The data in the descriptive statistics, figures 2.1-2.4, suggests that the tested populations improved their scores from the pretest to the posttest. The means of the posttests are higher than the pretests. However, the variance in each data set also increased as evidenced by the wider curve in each of the posttest graphs. This is confirmed by the increase in the standard deviation of each group from pretest to posttest. In order to show that the increase in mean scores was statistically significant, a paired two sample t-test (two sample dependent t-test) was used at α=0.05. The results showed a t-test value of 12.22, which was greater than the critical t value of 2.00. The null hypothesis that the mean score on the posttest is less than or equal to the mean on the pretest was rejected. The posttest scores of the students in all three classes were statistically
different and significantly higher than their pretest scores. The statistical test data and summary are included in table 2.2 below.

Table 2.2: t-Test: Paired Two Sample for Means
Tested population: Investigator 2 Posttest and Pretest Scores

<table>
<thead>
<tr>
<th></th>
<th>Investigator 2 Posttest</th>
<th>Investigator 2 Pretest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>7.707692308</td>
<td>3.123076923</td>
</tr>
<tr>
<td>Variance</td>
<td>9.210096154</td>
<td>2.890865385</td>
</tr>
<tr>
<td>Observations</td>
<td>65</td>
<td>65</td>
</tr>
<tr>
<td>Hypothesized Mean Difference</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Degrees of freedom</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>t Stat</td>
<td>12.21745061</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) two-tail</td>
<td>2.2086E-18</td>
<td></td>
</tr>
<tr>
<td>t Critical two-tail</td>
<td>1.997729633</td>
<td></td>
</tr>
</tbody>
</table>

To compare the posttests of the treatment group with the control group, an ANOVA test was completed ($\alpha=0.05$) to test the following null hypothesis: there is no difference between the pretest scores of investigator 2 and the control group. The ANOVA test resulted in an $F$-value of 2.25, which was larger than the critical value of 2.06. This means that the null hypothesis must be rejected. The sample population of Investigator 2 is not statistically similar to the control group. The conclusion is that the treatment and control groups have significantly different posttest scores. The statistical test data and summary are included in table 2.3 below.

Table 2.3: ANOVA: Single Factor Test for Investigator 2 vs. Control Pretest Scores

<table>
<thead>
<tr>
<th>Groups</th>
<th>Count</th>
<th>Sum</th>
<th>Average</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investigator 2: 0 Hour</td>
<td>13</td>
<td>30</td>
<td>2.307692308</td>
<td>0.564102564</td>
</tr>
<tr>
<td>Investigator 3: 3rd Hour</td>
<td>23</td>
<td>70</td>
<td>3.043478261</td>
<td>2.407114625</td>
</tr>
<tr>
<td>Investigator 3: 4th Hour</td>
<td>29</td>
<td>103</td>
<td>3.551724138</td>
<td>3.97044335</td>
</tr>
<tr>
<td>Control: 1st Hour</td>
<td>22</td>
<td>65</td>
<td>2.954545455</td>
<td>1.664502165</td>
</tr>
<tr>
<td>Control: 2nd Hour</td>
<td>23</td>
<td>55</td>
<td>2.391304348</td>
<td>1.339920949</td>
</tr>
<tr>
<td>Control: 3rd Hour</td>
<td>30</td>
<td>75</td>
<td>2.5</td>
<td>2.120689655</td>
</tr>
<tr>
<td>Control: 5th Hour</td>
<td>24</td>
<td>60</td>
<td>2.5</td>
<td>1.652173913</td>
</tr>
<tr>
<td>Control: 7th Hour</td>
<td>19</td>
<td>47</td>
<td>2.473684211</td>
<td>1.040935673</td>
</tr>
</tbody>
</table>

ANOVA Statistical Test Summary:

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>$F$</th>
<th>P-value</th>
<th>$F$ critical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>31.8529503</td>
<td>7</td>
<td>4.550421471</td>
<td>2.252251829</td>
<td>0.032239204</td>
<td>2.062240348</td>
</tr>
<tr>
<td>Within Groups</td>
<td>353.5678147</td>
<td>175</td>
<td>2.020387513</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>385.420765</td>
<td>182</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Normally, a $t$-test would be used to compare the posttest scores of the treatment and control groups to show a statistical difference between the two. Since these populations are starting out at different base levels (different mean pretest scores), the $t$-test comparison would not result in a fair test. Instead, another form of inferential statistic must be employed,
calculation of a Hake gain. A Hake gain, symbolized \(<g>\), is a method of comparing how much a population improved, the average gain, to how much room the population had to improve, the maximum possible average gain (Hake). The following items summarize the mean pretest and posttest scores and the calculated normalized gains of Investigator 2’s treatment group and the control group. (As stated in the Results for the Control group, the posttest scores of the control may be treated as one group). The gains of the control group were compared to both honors and regular chemistry treatment groups.

![Comparison of Pretest and Posttest Scores](image)

Figure 2.5: Mean total gains for Investigator 2 and the Control

### Table 2.4: Summary of Normalized Gains for Investigator 2 and the Control

<table>
<thead>
<tr>
<th>Testing Sample:</th>
<th>Pretest:</th>
<th>Posttest:</th>
<th>Normalized Gain (&lt;g&gt;):</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investigator 2: 0 Hr</td>
<td>2.31</td>
<td>4.46</td>
<td>0.17</td>
</tr>
<tr>
<td>Investigator 2: 3rd Hr</td>
<td>3.04</td>
<td>7.87</td>
<td>0.40</td>
</tr>
<tr>
<td>Investigator 2: 4th Hr</td>
<td>3.56</td>
<td>9.03</td>
<td>0.48</td>
</tr>
<tr>
<td>Control</td>
<td>2.56</td>
<td>4.81</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Instead of merely comparing the posttest to the pretest, the Hake calculations normalized the gain so that conclusions can be drawn about the significance of the improvement from pretest to posttest for a population. Hake classifies a high-gain course as one with \((<g> > 0.7)\), a medium-gain course as one with \((0.7 > (<g>) > 0.3)\), and a low-gain course as one with \((<g> < 0.3)\) (Hake). The above results show that the 0 hour section of the treatment group and the control group were considered to be low-gain, while the 3<sup>rd</sup> and 4<sup>th</sup> hour sections were considered medium-gain. Two of the three treatment populations showed normalized gains that are significantly higher than those of the control group.
Questions pertaining to the energy standards were selected for analysis in the study. The percentage of correct responses on the pretest and posttest are summarized in the graph above. All questions show improvement as a result of the treatment. In addition, the percentage correct on many of the items increases to over 50%, which was seldom seen in the same chart for the control group. Certain questions show large improvement, such as BECI numbers 7, 8, 10, 11, and 15. Others, such as BECI numbers 3 and 13, showed improvement but were still answered poorly by the students. Reasons for this will be discussed in the conclusion.

Conclusion:

Overall, the treatment group exhibited higher mean posttest scores on energy related test items than the control group. This illustrates the effectiveness of the modeling curriculum and supports the hypothesis that a modeling approach to teaching energy concepts results in greater learning than a traditional approach.

While most of the treatment group experienced significant improvements, conclusions can also be drawn from the one group that did not. The regular chemistry class of the treatment group exhibited gains similar to that of the control group in the low-gain range. There are possible circumstances that can detract from successful implementation of modeling curriculum. The regular chemistry class population took the course as a “zero” hour. This class met from 6:30 to 7:20 a.m. daily. The issue here is that zero hour never had the 100-minute block classes each week that the other treatment classes had. The result: lesson plans without the block would often result in discussions and whiteboard sessions being cut short. At 6:30 a.m. the next day, the students were not always motivated or in tune to pick up where white boarding and discussion left off. Another reason for lower gains is that this population of students did not exhibit the same sort of buy-in to the modeling process as the other treatment groups or as previous years. The class was composed of mainly juniors, while the honors classes were made up of a majority of sophomores. These students are likely more engrained with traditional
teaching approaches and resistant to modeling treatment. Overall, while the modeling approach is a significant improvement in teaching energy concepts, teachers should still be aware of what factors attempt to work against them.

The two honors class treatment groups experienced medium-gains in their scores. These classes took the course during normal school day periods that included block days twice a week. They were also more receptive to the modeling approach over the course of the year. Although lesson plans used with all treatment groups were similar, these classes were more actively engaged than their zero hour counterparts. Other factors such as prerequisite math skills and cognitive abilities entering the course may also play a role.

None of the treatment group fell into the high-gain category as defined by Hake. In Hakes original study in which the Hake gain was defined, Hake studied high school, college, and university classes. Classes were divided based on their use of either interactive engagement methods, similar to modeling techniques, or traditional methods in teaching physics concepts. In Hake’s study, all 14 of the traditional courses and 7 of the interactive engagement courses fell into the low-gain region. The other 41 of 48 interactive engagement courses fell into the medium-gain region. None of the courses studied fell into the high-gain region (Hake). The results of this investigation for modeling instruction find similar results. Most of the treatment groups fell into the medium-gain region showing an improvement over traditional methods. These results correlate with Hake’s, illustrating that interactive engagement techniques, such as the modeling treatment, are not successful in every course, but are successful in most courses. In this case modeling does represent a statistically significant improvement over traditional methods.

These conclusions can be further supported by the item analysis. Students from the treatment group showed large improvements in answering some of the selected items. Questions 8 and 15 from the BECI both deal with the distinction between heat and “coldness” involved in energy storage and transfers. Students from the treatment group experienced much larger gains as compared to the control group. Most students answered question 9 about energy transfers from different substances correctly in both pre- and posttests in treatment and control population. However, the follow-up explanation question number 10 saw a much greater correct answer percentage gain for the treatment group. Questions 3, 4, and 13 were answered poorly with smaller gains for both this treatment and the control groups. These questions were related to energy storage and heat capacity. Questions 3 & 4 are both heat capacity related and were answered much better by the other treatment groups. This may illustrate a point weakness in the application of the modeling treatment by Investigator 2. While overall percentage increases in test scores support the modeling method, finer analysis reveals that there are minor differences among modeling instructors themselves.

This data supports that students who have undergone the modeling treatment have better conceptual understanding of how energy is stored and transferred between systems and surroundings. The modeling treatment results in a better model of energy, attacks misconceptions such as a concept of “coldness”, and supports distinctions between terms like temperature and energy. These outcomes were achieved through the consistent use of modeling techniques such as interactive lab experiences and student discourse about concepts and
problems facilitated through white boarding sessions. Students who did not receive this treatment still improved their scores, but had more difficulty with questions distinguishing heat, temperature, and “coldness.” They did not perform as well on items relating to energy storage and transfer. Finally, there are still items that both populations struggled with. This suggests that these are concepts students have the most difficulty with, or there could be an issue with the writing of these test items.

In summary, analysis of the question items and the overall results from these treatment groups suggests that the modeling treatment is a statistically significant improvement over the traditional approach in the teaching of energy concepts and standards. Students who have undergone the treatment experienced significant gains in the test items that required the coherent application of energy concepts in chemistry contexts. While this does not prove the claims of the research group, the data definitely supports that modeling techniques are an improvement to traditional approaches to teaching energy.
Investigator 3 Report:

Introduction:

Investigator 3 is Andrea Strock from Gilbert, Arizona. Andrea Strock teaches Chemistry and Physics in the southeastern portion of the Phoenix Metropolitan area in the Chandler Unified School District. Strock has been teaching for three years. The 2008 – 2009 school year was Strock’s second time teaching Chemistry, and the first time utilizing the Chemistry Modeling approach. Last year 119 of Strock’s’ regular chemistry students were part of this action research project. The reason not all of the regular chemistry students were included is that students were allowed to change classes throughout the school year. Various students dropped out of accelerated chemistry programs into regular chemistry, and some of the regular students dropped Chemistry altogether. It was for this reason, a small segment of the initial student population changed. In consideration of population issues, students who were dropped or were added during the school year were omitted from the statistical analysis when the data was analyzed.

Procedure:

Last year Strock administered the BECI and the CCI to her students. For this action research, Strock proctored both of the tests during the first week of school for the pre-test, and both post-tests during the last week of school. The CCI was conducted as part of the final exam. At the time the students took the post-test, they had completed the treatment unit approximately 16 - 20 weeks prior. Despite this larger time gap, the pre- and post- tests were analyzed to assess long-term learning of the energy concepts. In between the pre-test and the post-test various modeling techniques were employed, including, but not limited to: student-centered activities, conducting labs before defining terms, white board meetings, and student discussions. In Units 2 and 3, taught in the fall semester, students performed labs, such as the ‘Icy Hot Lab’, in which they warmed ice and liquid water to produce water vapor. During this lab, students had to take careful temperature measurements and watch a clock in order to create a temperature versus time graph. From this graph students had to infer what was happening on a molecular level to the particles as temperature increased, how particle motion was affected, and what phase(s) are present in heating/cooling curves. Students watched small simple video clips from the ‘Eureka’ series of videos to summarize their findings. Students deduced that particles move more in the gas form than in the solid form; and that this movement is referred to a ‘kinetic energy’; and as particles bump into each other, ‘interaction energy’ is transferred. Energy bar charts were introduced to students as a method to use to account for energy storage and energy transfer. Students ultimately were able to distinguish between heat and temperature and predict how energy would change within a system and its surroundings.
Results:

After administering both the BECI and the CCI, it was decided that only 15 select questions from these tests would be assessed. Strock was interested in looking at how these students fared compared to the control group students. The most obvious place to look for comparisons was the assessments. To start, pre-test scores were compared.

A single factor ANOVA test was run on the groups to determine whether or not they were from the same population of students. The null hypothesis for the control group was accepted. This means that although the control group consisted of a combination of honors and regular chemistry classes, all the scores were so similar, they could not be considered statistically different.

Based on descriptive statistics, the pre-test mean on the selected energy questions for Strock’s classes was 3.07 and the pre-test mean for the control group was 2.56. Statistically this could imply that the populations are similar, unfortunately, Strock had a few students in one of her classes who were taking all AP classes, along with one regular chemistry class as a ‘decompress class’; this caused the pre-test score mean for this one class to be abnormally high. As depicted above in Figure 3.1, Strock’s pre-test scores show greater variance than those of the control group. When the ANOVA test was run on Strock’s four regular chemistry classes to compare them, it was determined that the groups were statistically different. (See Table 1 below) The null hypothesis, which states that the population means are the same, was rejected. Instead of simply comparing posttest scores of the treatment and control groups, a Hake test was used to compare normalized gains.
Table 3.1: ANOVA: Single Factor Test for Treatment and Control Group Pretest Scores

<table>
<thead>
<tr>
<th>Groups</th>
<th>Count</th>
<th>Sum</th>
<th>Average</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investigator 3: 1&lt;sup&gt;st&lt;/sup&gt; Hour</td>
<td>29</td>
<td>96</td>
<td>3.31034</td>
<td>2.952380914</td>
</tr>
<tr>
<td>Investigator 3: 2&lt;sup&gt;nd&lt;/sup&gt; Hour</td>
<td>29</td>
<td>53</td>
<td>1.82759</td>
<td>2.504926108</td>
</tr>
<tr>
<td>Investigator 3: 3&lt;sup&gt;rd&lt;/sup&gt; Hour</td>
<td>30</td>
<td>65</td>
<td>2.16667</td>
<td>0.902298851</td>
</tr>
<tr>
<td>Investigator 3: 5&lt;sup&gt;th&lt;/sup&gt; Hour</td>
<td>30</td>
<td>76</td>
<td>2.53333</td>
<td>2.188505747</td>
</tr>
<tr>
<td>Control: 1&lt;sup&gt;st&lt;/sup&gt; Hour</td>
<td>22</td>
<td>65</td>
<td>2.95455</td>
<td>1.664502165</td>
</tr>
<tr>
<td>Control: 2&lt;sup&gt;nd&lt;/sup&gt; Hour</td>
<td>23</td>
<td>55</td>
<td>2.3913</td>
<td>1.339920949</td>
</tr>
<tr>
<td>Control: 3&lt;sup&gt;rd&lt;/sup&gt; Hour</td>
<td>30</td>
<td>75</td>
<td>2.5</td>
<td>2.120689655</td>
</tr>
<tr>
<td>Control: 5&lt;sup&gt;th&lt;/sup&gt; Hour</td>
<td>24</td>
<td>60</td>
<td>2.5</td>
<td>1.652173913</td>
</tr>
<tr>
<td>Control: 7&lt;sup&gt;th&lt;/sup&gt; Hour</td>
<td>19</td>
<td>47</td>
<td>2.47368</td>
<td>1.040935673</td>
</tr>
</tbody>
</table>

ANOVA Statistical Test Results:

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P-value</th>
<th>F critical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>46.0252</td>
<td>8</td>
<td>5.75315</td>
<td>3.307555555</td>
<td>0.00138</td>
<td>1.98165</td>
</tr>
<tr>
<td>Within Groups</td>
<td>373.97</td>
<td>215</td>
<td>1.7394</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>419.996</td>
<td>223</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 3.2: Comparison of control group pretest scores to Strock’s pre-test scores

As shown in Figure 1, above, the highest score in the control group’s pretest population was a 6 out of 15, while the highest score in Strock’s population was a 9 out of 15. This difference was enough to offset the populations slightly, but for the most part each population’s pre-test scores started out relatively similar.

Due to the fact that the ANOVA test could not be employed, it was concluded that a traditional statistical comparison between these populations using the t-test was inappropriate.

The two populations were compared using Hake gains. The intent of using Hake gains was to show that the improvements from pre-test to post-test of the treatment group were greater
A Comparison Study for Teaching Energy

than the improvements made by the control group. The following bar graph shows the Hake gains for Strock’s treatment group as they compare to the control’s treatment group:

![Comparison of Pretest and Posttest Scores](image)

Figure 3.3: Histogram of the average Hake gains for Strock’s group versus the control’s group

<table>
<thead>
<tr>
<th>Testing Sample</th>
<th>Pretest:</th>
<th>Posttest:</th>
<th>Normalized Hake Gain:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investigator 3: 1st Hr</td>
<td>3.31</td>
<td>8.72</td>
<td>0.46</td>
</tr>
<tr>
<td>Investigator 3: 2nd Hr</td>
<td>1.83</td>
<td>5.97</td>
<td>0.31</td>
</tr>
<tr>
<td>Investigator 3: 3rd Hr</td>
<td>2.60</td>
<td>9.37</td>
<td>0.55</td>
</tr>
<tr>
<td>Investigator 3: 5th Hr</td>
<td>2.57</td>
<td>6.93</td>
<td>0.35</td>
</tr>
<tr>
<td>Control Group</td>
<td>2.56</td>
<td>4.81</td>
<td>0.18</td>
</tr>
</tbody>
</table>

As shown in the table above, while the control group did have an increase in average mean from pre- to post-test scores, the control group’s mean was still well below the lowest performing class in Strock’s population. The treatment group, by far, out performed the control group. Additionally, all four of Strock’s classes had a mean Hake gain score of 0.42, which falls into the moderate range for Hake gains, while the control group’s Hake gain of 0.18 falls into the low-gain range for Hake scores, thus showing that the treatment group is statistically different from the control group. The results obtained by Strock’s population are consistent with the results found in the treatment group overall and support the belief that the modeling approach to teaching energy is superior to the traditional approach.

The histograms below clearly show that the pretest sample for Strock’s treatment groups was skewed towards very low scores, and the posttest sample was skewed higher. The results of the pretests show a distribution of scores in the low number correct range. The skew in this data
is more to the left towards fewer correct answers. The posttest on the right shows how the mean scores improved as the normalized curve shifted to the right. There is a slight increase in the variance of student scores since some students still answered few questions correctly, while many more answered a larger number correctly. The increase in variance is shown in the increase of standard deviation between the two data sets. The scores for the post-test fit a normalized curve better, likely due to the greater distribution of scores within the population. The mean score increased from 2.56 to 7.75. Because the posttest sample is shown to be from a statistically more advanced population of students, Strock concluded that the students who received the treatment showed significant gains in their understanding of the concept of energy in chemistry. Furthermore, Strock concluded that her sample population performed similarly to the overall treatment group, based on the reports from Investigators 1 and 2 previously discussed.

In addition to looking at how the pre- and post-test scores compared with the control population, Strock was interested in looking at the scores on the selective energy questions from the BECI and the CCI. (See Figure 5, below) Interestingly enough, the students had gains on every single energy question, however not as high of a gain as might be hoped.

Question 17 on the CCI asked the students to explain their reasoning for their answer to question 16, which asked which received more heat, equal amounts of water, or alcohol being heated to 50°C. Student responses were inconclusive due to the fact that they varied greatly. Obviously, this was not fully understood by the students, because the post-test percentage was only slightly above 30%.

Question 22 on the BECI required that students understand the properties of water and its phase changes. A lot of time was spent discussing this topic in the fall semester, with energy bar charts, heating curves and cooling curves used to help emphasize this concept. Obviously, this was a concept grasped more clearly by the students.
Conclusion:

As a whole, Strock’s treatment group had significantly higher scores on their CCI and BECI assessments than the control group. This is evidenced by the population’s greater Hake gains than the control group’s population. The success of the treatment group is attributed to the process they went through of learning about chemistry with an emphasis on the energy concept, using the modeling curriculum.

This being Strock’s first year teaching chemistry using the modeling approach means no previous averages for the CCI or BECI to be compared to, but it was Strock’s knowledge of modeling from the past couple of years that was a benefit to her teaching ability.

Extrapolating from the data, it is particularly interesting to note that the treatment group in Strock’s regular chemistry class exceeded the performance by the control’s honors chemistry class. There are several possible reasons why this occurred. Much of Strock’s class time was spent on whiteboard discussions in the classroom. Strock asked her students to support their writings on the whiteboards verbally with evidence for their board. Often Strock would attempt to tie in a real world application so that her students could apply their conceptions immediately. While this sounds ideal, some of the students in Strock’s population did not ‘buy-in’ to the modeling approach to teaching. These students wanted to do as little as possible, and move around the classroom as little as possible during the school day. Looking at the original student response data, it is quite evident who these students are based on their low post-test scores.

The hypothesis that teaching energy using the modeling approach is superior to the traditional approach was supported based on the data, however it cannot be said beyond a doubt that this is the ultimate answer to teaching energy in chemistry. The treatment did not always give the largest gains, although the gains seen were greater than those from the control group.

Figure 3.6: General chemistry treatment group performance on the 15 selected energy questions taken from the BECI and the CCI
This is evidenced in the item analysis. Question 3 had a large gain for Strock’s classes. Question three asks about how you would keep your food hot or cold; wrap it in an insulator or a conductor? The modeling material does a nice job of describing what is happening at the molecular level in an insulator and a conductor. Strock’s classes led extensive discussions on why you pack your lunch in a lunch bag, and not an aluminum box; and why your house is made out of bricks and glass, and not silver and iron. These extensive discussions stuck with the students and ultimately this was one area of large improvement for them. On the other hand, some discussions did not happen and this is where students still had misconceptions, and could not apply their learning of concepts adequately. For example, number 13 on the BECI, asks students about the difference in temperature between a wood plate and an aluminum plate one minute after being taken out of the refrigerator. Many students answered this wrong, saying that both would be at the same temperature because the refrigerator was at a constant temperature. Ultimately, there were items on both exams that both the control group population and the treatment populations struggled with. This implies that the test itself may have some problems.

Strock believes that the treatment used on her students led them to earn higher scores on the CCI and BECI and have an increased understanding of energy than they would have otherwise. By analyzing the combined selective questions of the CCI and BECI questions individually you can see that overall there was an increase in the mean scores for her classes. This increase was very sizeable and quite impressive to Mrs. Strock.

Before conducting this research Strock admits that she was not a huge proponent of teaching energy in chemistry as a central concept, however after carrying out the treatment on her students she would be a huge advocate of emphasizing the concept of energy in as many chemistry units as possible. Strock’s students this past year exhibited a much deeper understanding of the information that they can extract from a graph than in years past. Strock’s students were able to apply energy bar graphs to understanding various endo- and exothermic processes. In addition, they could apply energy concepts to the stoichiometry of exothermic and endothermic reactions, a topic addressed in later units.

It is Strock’s belief that the most important part of the treatment was ensuring student understanding of energy and how it can be incorporated into multiple units in the subject of chemistry. Strock was particularly concerned with her student’s qualitative data and their reasoning abilities. In discussions with her students it was evident that they had learned about energy and were thinking about how to reason out their thoughts, although sometimes it was not as eloquent as she would have hoped. If nothing else, Strock can say that she emphasized a very important concept that can be beneficial to the students and easily used for cross-reference in other science subjects.
Qualitative Data:

The investigators also collected and examined qualitative data to see if it also supported the conclusion that the treatment improved students’ understanding of energy. The qualitative data primarily consisted of interviews, student white boards and student work. As shown below, many students drew graphs to help them solve energy problems after the treatment was administered. Despite the fact that the students are not speaking meaningfully, the investigators concluded that the students who received the treatment solved energy problems correctly, with sound reasoning, more frequently than students who did not receive the treatment did.

Each of the questions used in the student interviews were taken from the CCI (see Appendix C), and the extended CCI (see Appendix E). The investigators consider these problems to be fairly challenging for high school chemistry students because the distracters on these questions are quite convincing. Unfortunately, for the comparison study, these interviews were only conducted on the treatment group, and not the control group. However, they are still good exemplars of student thinking before and after treatment.

Students’ scripted interviews:

This first interview is of two female students (S1 and S2) and the teacher (T) on the pretest interview. This interview is typical of a pretest interview; students had very little comprehension of the concept, but they are positive that they are correct. They have no idea why but they are certain they are correct.

T – Is it OK to videotape you?

S1 – Yes, it is OK for you to video me.

S2 – Yeah, its OK and then she read question #25.

S1 – OK, so temperature equals –5°C (writes this on the white board). OK and goes to 5°C (write this)

S2 – This is… (inaudible)

T – Could you girls talk louder?

S1 – I don’t know what to do.

T – Just talk aloud, whatever you say is OK.

S1 – OK, the graph should go up. So it should start down and go up.

S2 – It says the temperature is going up
S1 – So it goes up.

S2 – Yeah, but it doesn’t start at 0.

S1 – OK, but its going up at 5 and passing through so it’s A (correct answer is D)

S2 – AAAHHHH, Yeah – A

S1 – It’s A … we think.

T – OK, can you draw it please?

S1 – Yes Ma’am, here you draw it.

S2 – OK (copies the graph directly from the question)

T – OK, So what would be an example of a situation where you would see this graph or these temperature changes?

S1 – ice

S2 – snow melting

T – Ice or snow melting? Could you just label that graph for me really quickly? (students label graph)

T – So how would you verbalize temperature relating to time given that picture that you drew? (pause)

S1 – What do you mean?

T – When you look at that graph, how would you relate temperature to time?

S2 – The amount of… wait… (pause)

S2 – Would you say that the temperature increases at time… (pause)

T – Where does your temperature start increasing?

S1 – About like a little bit after (she points to the graph)

T – OK, so from the very beginning how does your temperature and time relate?

S1 – its constant

T – OK and what about the numbers tell you, you are going to have a constant temperature?
S1 – Cause your going from –5 to 5 and you have to pass 0 to get to 5.

T – OK and what happens at 0 to make you think we have a constant temperature?

S1 – mmmm – that’s a good question. (pause)

S2 – Well, there a phase change starts immerring… I don’t know.

T – Do you have any reason to doubt your answer?

S1 – (shakes head yes)

T – What’s your reason for doubting?

S1 – AAAH HH.. Cause none of them are right… none of the graphs.

T – What makes you think you are wrong?

S1 – I don’t know.

S1 – I don’t think we’re wrong we are right

T – You think you’re right?

S1 – yes!

S2 – With this graph – yeah.

T – Confident?

S1 & S2 – Yeah!

This second interview is of a male (M) student, a female (F) student and the teacher (T) on the pretest interview. They are not certain what to do. However, it is interesting to watch as the female student has some correct ideas but is not sure if she is correct. She is unable to convince the male student what she is saying is correct; so she backs down and agrees with him.

T – Are you OK if I videotape you.

F – Yeah.

M – No problem.
T – OK, Will you please read the question aloud?

M – OK (he then reads question #23)

M – What we need is a graph that goes from 0°C to 5°C

F – So the graphs gonna go up?

M – yes and its going to increase… Is going to start at 0 but it’s not going to go up this way (he point to graph C)

M – Is time…Is time the x or the y value?

F – This looks like a constant because it is the constant heat of the electric heater…

M – Well I don’t… We have to understand what our x and y is. Right?

M – We have to understand if its time over temperature or temperature over time.

F – Oh, OK. (pause)

F – I think it is temperature on the y axis.

M – OK, so we have a y axis (drawn a vertical line) and an x axis (drawn a horizontal line)

M – Temp…. waters temp is just capital T and the time is lower case.

M – It goes from 0 to 5° (pause)

M – It takes 5 minutes so time is 5 minutes (start marking the y axis)

F – This is temp.

M – I forgot it was temperature (writes on the x axis)

M – and then it goes to 5 minutes either way it is 5

M – AAHHH – label that from 0 to 5 (F labels the y axis)

M – Temp (points to the T)

F – Yes, the big T is temp

M – OK
M – So our options are … (reads all the answers out loud)… Well if is gonna go from 32°F to 41°F over the course of 5 minutes that is gradual so its not gonna be an instant increase

F – But its not gonna stay the same for the entire time

M – Well a water heater.. it says… (reads the question again)… that means like a hot plate… and it’s a mixture of ice… the ice is going to increase slowly and then increase rapidly as it goes.

F – So wouldn’t a gradual go like that (motions her marker like A) and says the A one.

M- I don’t want to say constant.  I’m thinking B (reads the answer out loud)

F – Well, that’s not like our Icy Hot Lab.  It’s going up and it has some little plateaus and stuff.

M – I’m saying B (write B on the board and draws the B graph)

M – As a final conclusion we came to the answer of B (reads the answer again)… (the correct answer is A)

This third interview is of a male (M) and the teacher (T) on the posttest interview. This is the same male student that was in the second pretest interview above. He verbalizes his work and demonstrates what he is doing very well. He starts out drawing the graph correct. Then, he changes it to an incorrect representation. Finally, he realizes he drew the graph incorrectly and gives the correct answer. It takes him a while to come up with the correct answer.

M – (reads question #23 and write down information and starts a graph)

T – So, Could you explain what you’re doing right now?

M – Right now I’m taking…uummm… I just brought down the 0°C and the 5°C so I remember what two numbers are so I don’t have to keep looking at the top of the paper.

M – and I’m going to draw a graph showing the time and the temperature

M – I’m gonna put the temperature on the y axis and time on the x axis (write it down – pause while he writes the numbers on the graph)

M – right now I’m just demonstrating going from 0 to 5 degrees which is the span we have but I’m going to drawn the while graph and pull it out later so I’m gonna take it up to 100°C to its boiling point. (student draws the graph line correctly but then erases the 0°C line and moves it to 5°C)

M – So, right there… on this graph… even before looking at the answers I realize… that let’s say that the temperature started at a –10°C.  It would increase when it goes up to 0, which would be starting to melt from its freezing point…. And going from 0 all the way to 100 and 5°C would
have the constant of the melting point first and then a slight increase going up to its gas point… oh… boiling point.

T – What temperature is that “constant” at?

M – AAHHH… the constant is right now…well it’s suppose to be at 0°… but I drew the map incorrectly… I apologize… UUUMMM, so, I think the best answer is (reads all answers)... so from 0°C to 5°C.. (pause)...mmm…I’m going to have to go with A… (the correct answer is A)

T – Why did you choose A?

M – The temperature stays constant for a while and then rise because if it’s coming from…as my graph shows… from a freezing of 0 it would just be melting from where ever it is below and then as a slight increase going to 5° would cause it to start to melt and so then it would start to rise so it would go from a constant to a rise because the water would need time to rise up...

T - OK
Student answers on Interview Questions:

Question 6: The circle on the left shows a magnified view of a very small portion of liquid water in a closed container. What would the magnified view show after the water evaporates? There is a particle diagram of liquid water on the left and students have to select the particle diagram for evaporated water. (See the details of this question in the Appendix E)

For the last problem I chose D because it's more free floating and less condensed.

Student ZW wrote as his pretest interview: “For the last problem I chose D because its more free floating and less condensed.”

Student ZW’s answer shows the particle diagram with oxygen and water separated and not bonded as water. This shows a misconception that oxygen and hydrogen separate from each other in the gas form, instead of water vapor staying as H2O in the gas form.

E because it’s still water in gas form since water can only be separated by electrolysis.

Student AF wrote in her posttest interview: “E because it’s still water in gas form since water can only be separated by electrolysis.”

Student AF’s answer was correct and he knew it was still water. This shows that the misconception about water breaking apart into Hydrogen and Oxygen in gas form was addressed during the school year and this misconception was addressed. Previously, students discover that the only way hydrogen and oxygen particles can separate out of water (with no other chemicals involved) is through electrolysis.
Student SL wrote in her posttest interview: “The molecules are still in the form of water so they do not separate from each other (Hydrogen & Oxygen) but the water molecules separate from each other.”

Student SL’s answer was correct and show the water molecule just further apart to represent the evaporated water. This shows that the concept of solids versus liquids versus gases was comprehended, from Unit 2’s materials.

Question 27: What would be the temperature of the water in the Styrofoam cup a short time after the reaction is complete (long enough for the heat through the copper, but short enough to neglect heat transfer through the Styrofoam)? (See the details of this question in the Appendix E)

Student CR wrote in his pretest interview: “Because the copper kept heat and the Styrofoam doesn’t the temperature has to go down.”

Student CR’s answer was originally A, which was correct. He changed his answer by the time the interview was complete.
Student KP wrote in his posttest interview: “A - Because the solution in the cooper decreases which then makes the solution in the cup decrease.”

Student KP’s answer was originally C but he adjusted his answer to the correct answer after thinking through the problem.

Question 28: Students had to explain their prediction from question 27. (See the details of this question in the Appendix E)

Student CR wrote on his pretest interview: “the temp is decreasing causing no chemical change”

Student CR’s answer was E which was incorrect. He attempted to answer this question by process of elimination but crossed off the correct answer first thing.
Student ZW wrote as his pretest interview: “For the first problem I thought about energy and its transfer, and I thought if the energy is lost when you do this then the energy will be lost in the surroundings which would be the water. The system gives off heat and loses energy so that is why I choose E.”

Student ZW’s answer shows much thinking and an OK explanation, with many big words. However, he answers the questions incorrectly as the correct answer is D.

Question 23: Choose the graph which best describes the change in temperature of the water (T) as a function of time (t), neglecting any heat loss to the environment. Students then have four graphs to select from. (See the details of this question in the Appendix E)

Two students, AB and CN, both wrote on theirs pretest interview no explanation but answered B, which is incorrect as the correct answer is A. These students did not have any idea where to start when answering this question. This is typical of pretest answers. Please note there are no diagrams/pictures as there were none on these papers.
Student AG wrote in his posttest interview: “It takes a while for solids to melt, but then it gets easier to heat, thus increasing (progressive curve), uses knowledge of temperature & thermal & interaction energy.”

Student AG’s answer was correct and diagram explained this very well. His answer shows that he comprehended the heating and cooling curve graph, which depicts the process, and he was able to apply the forms of energy change that coincide with heating curves. The students learned this in Unit 3.
Student JI wrote in her posttest interview: “It was not completely melted yet so it keeps a straight line till it is all liquid goes consistently it takes more energy to warm up a solid then a liquid because it has less space for energy to get in.”

Student JI had a wonderful explanation with her drawing; however, she marked the answer incorrectly. This shows that quantitative data cannot definitely be the final answer; often times, students understand the material, and accidentally mark the wrong answer choice.

Question 25: Choose the graph which best describes the change in average temperature of the aluminum (T) as a function of time (t), neglecting any heat loss to the environment. Students then have four graphs to select from. (See the details of this question in the Appendix E)

Student AS wrote on her pretest interview: “Once the aluminum hits zero, it stays constant then rapidly increase Option A displays that the best”

Student AS’s answer was A, which was incorrect (the correct answer is D). She attempted to answer this question by simply looking at the graphs and selecting one that made the most sense to her without thinking about what substance the question was asking about.
Student CB wrote in her posttest interview: “Metal take a long time to loose it’s shape & become a liquid. It also shows that if you are at a standard rate that you have to be at a certain temp of which has not yet reached it’s peak.”

Student CB’s answer was correct and she actually wrote a detailed explanation that was also correct.
Student White Boards:

Throughout the school year, students conducted many board meetings with discussions. The investigators were interested in seeing what students thought temperature looked liked as a function of time as water was heated. The following picture, as shown below, shows a sample graph that many students drew before conducting the lab.

Photo 1: White board predicting what the ‘Icy Hot Lab’ graph would look like as the temperature of ice increases to a boil

While the lab was conducted, students were asked to record the temperature of the ice and eventually the water every 30 seconds as the system was heated. The drawing below shows a sample of the results that 99% of the groups presented in their post-lab board meeting.

Photo 2: Graph after the Icy Hot Lab was complete.
Approximately one week after the ‘Icy –Hot’ Lab was completed, students were asked the following warm-up question: While the Apollo 11 space shuttle was in space, all of the power had to be shut down, due to an explosion. For a few days, the men on board were exhaling CO\(_2\), and due to the cold temperatures, the gas was condensing on the walls. Create a temperature-time graph, on a whiteboard, depicting this situation assuming the temperature onboard the L.E.M. got as cold as -32°F. The graphs below show that many groups correctly accomplished this warm-up, but some groups still showed misconceptions.

Photo 3: Student is confused on the concept of temperature decreasing as the energy is added.

Photo 4: Students showing a cooling curve based on their knowledge of heating curves.
Photo 5: Jim’s group showed the start of the fusion diagram.

Student Notebooks:

Furthermore, investigators in the treatment process used interactive science notebooks, in which their students took all their notes, placed all their worksheets, and corrected all of their assessments. These notebooks were valuable student-thinking tools for several reasons. They helped students organize the inquiry process, collect evidence, and work at their own level. Students found it easier to write their laboratory reports based on their hands-on learning and having everything in one place. Student notebooks are a great way to access and process student learning through different modes of differentiated instruction. Finally, the notebooks were a great formative assessment for teachers to identify evidence of student learning and identify student misconceptions. Unfortunately, due to the size of the notebooks, they cannot be included in this paper but will be referenced.

Overall, inspection and analysis of student work presented here shows that the students who received the treatment were likely to solve their temperature-time problems correctly, and were willing to incorporate dialogue about energy into their discussions. The investigators conclude from the qualitative data that students were successful in solving energy-related chemistry problems, and this did not decrease their overall level of problem solving for non-energy concepts. In fact, those students were able to recognize the role of energy and apply their model of energy to the problems had a greater success rate on their first attempt than those students who did not. The treatment helped students incorporate a coherent, workable model of energy into their knowledge schema.
Student Quotes from the school year:

The following are student quotes demonstrating that students felt like they learned more in the treatment classes than the control classes. Students also enjoyed the treatment classes more since it was more challenging and hands-on.

♦ “Mrs. Blakeney, this class teaches us so much as we are actually doing something.”

♦ “I learned more in your class in one month then I did in the entire semester of the other class.”

♦ “I wish I was in this class all year so I could really learn chemistry”

♦ “Mrs. Strock, can we do this the rest of the year?”

♦ “I love this class! It’s more challenging than my Honors classes!”
**Study Conclusion:**

Statistical evidence supports that the students who received the treatment had a more coherent conceptual model of energy than those students did in the control group. After viewing the qualitative data, it is evident that the treatment gives students a visual reference to communicate their knowledge that they would not otherwise have. Visually communicating energy through bar charts and graphs helped students to follow their fellow classmate’s logic easily as opposed to just listening to their verbal communication. The treatment made a difference in the students’ problem solving skills based on the qualitative data presented. As the semesters progressed, it seemed that the students involved in the treatment had a more profound understanding of energy in chemistry than the students involved in the control group.

In addition, both honors and regular chemistry students can benefit from the modeling approach to teaching energy concepts. In the control group, honors and regular chemistry students progressed, but to a limited degree. The regular and honors chemistry students in the treatment groups exhibited posttest scores and normalized gains that were statistically significantly larger than the control group. This supports that the modeling approach is beneficial in teaching energy concepts to both regular and honors chemistry populations.

Conclusions can also be drawn from the item analysis of the treatment and control groups. In general, students in the treatment group exhibited larger percentage gains than the control, which would be expected following the statistical analysis. The treatment group did better on specific test items involving understanding of endothermic and exothermic processes. The treatment group also performed better on questions addressing energy transfer and the misconception of “coldness.” Students from the treatment group also better answered follow up questions that addressed prior problems.

A closer look at the item analysis for the treatment groups of each investigator yields an additional overall conclusion. There are some finer differences even between modeling teachers. Even though the modeling treatment outperformed a traditional treatment, not all modeling treatments resulted in the same gains on specific questions. An excellent example of this is questions 3 and 4 from the BECI. Investigators 1 and 3 saw much larger percentage gains in correct responses on these items than investigator 2. Question 3 asks about what materials would be best for storage of hot and cold food. Question 4 addresses energy exchange between two metals of the same identity, but different mass and initial temperature. These concepts may have been discussed more or less depending on both the instructor of the classroom and the students in the classroom. Modeling discussions and whiteboard sessions are student led and do not always end up in the same place. The data supports that the application of the modeling treatment; though superior to traditional methods, does differ slightly amongst instructors. This may even be a result of the student-centered focus of the modeling process.

Overall, the advantages of modeling instruction are often evident to teachers who use modeling. They see the benefits throughout the school year as they teach various chemistry concepts. This study focused on statistically examining and supporting the advantages of using modeling methods to teach energy. The results of the study demonstrate that the modeling approach is a significant improvement over traditional education methods in improving student understanding of energy concepts related to chemistry.
Implication for Further Research:

As measured by the BECI test, CCI test, and qualitative data, this research was successful in demonstrating that modeling is a superior teaching method in helping students fashion a robust concept of energy. However, this conclusion is based on relatively small sample sizes. It is believed that the effect size and Hake gains found in this research are real, however more work must be done to determine how large an effect size or Hake gain there would be with larger, more diverse samples of the high school student population.

Furthermore, not all of the populations were consistent between the treatment groups, and between the treatment groups and the control. Further study should be done which improves upon the sampling techniques used. This would improve results by examining a more representative sample of the student population from among high schools and districts.

In addition, scheduling issues may affect the ability to implement a successful modeling curriculum. Further studies could examine the effect of different school schedules on the success of modeling instruction. Possible schedules that could be examined are block schedules that compress a course in one semester, block schedules that meet fewer days but for longer time periods throughout the school year, and traditional schedules of approximately 50-minute class periods.

Finally, a further study of the degree of student improvement with respect to the number of years the teacher has been using modeling instruction would be beneficial. The increase in student knowledge going from 1st year chemistry modeling teacher to 2nd year chemistry modeling teacher was considerable. Additional research should be conducted to examine this learning curve and also see if it extends to 3rd and 4th year teachers and teachers with extensive use of modeling, often referred to as “veteran modelers.”
Acknowledgements:

Rarely one can thank people that have assisted in an endeavor. We feel gratitude and great thanks to be able to tell these people how much their time and efforts have enriched our lives. We would like to thank, first and foremost, our committee consisting of Dr. Robert Culbertson, Dr. Carl Covatto and Dr. Richard Bauer for their support and flexibility throughout this whole process. We would like to thank Dr. Colleen Megowan - Romanowicz for her guidance and expertise during the research and writing of this paper. Without her assistance, this project would not have as stellar as it is. We would like to thank Jane Jackson for her assistance and the financial support that she provides to the MNS program. Without her aid to us individually, through her pursuit of grants, we would not have been able to complete our degrees. We would like to thank Larry Dukerich for his unwavering efforts in developing the Chemistry modeling materials so more teachers can improve their teaching methods to enhance the teaching of so many students. We would like to thank Guy Ashkenazi for his contributions to the modeling program, including the additional energy questions for the CCI, and for allowing us to use these questions in our student interviews. We would like to thank Michael Crofton, Jim Archambault, Angel McClure and Theresa Burch for their willingness to help us learn about this process through their experience. We would like to thank Aracelli Vizcarra and Ixtchell Ledesma for all of their help with the paperwork for the MNS program. We would like to thank our control teacher, Mr. Bryan Simon for allowing us to use him and his students in our study. We would like to thank David Flores for interviewing students to allow for an unbiased interpretation. We would like to thank our students for their willingness to be our test subjects. Finally, we would like to thank our families for the sacrifices that they made during the entire action research process, for us.
References:


Appendices:

- **A – Teacher Notes from Chemistry Modeling Units Used**

  The following are the instruction goals and sequence of the teacher notes as taken from the Arizona State University Chemistry Modeling Units. These are the units included in this study as it models how energy is woven throughout chemistry. Unit 2 introduces the energy concept, Unit 3 builds on the basic concepts, and Units 6 and 8 use the concept of energy to complete calculations.

**Instructional Goals and Sequence of Teacher Notes – Unit 2**

**Unit 2 – Energy & States of Matter – Part 1**

**Instructional goals**

1. Recognize that the model of matter that we use during this unit is essentially that proposed by Democritus.

2. Relate observations of diffusion to particle motion and collision in both liquid and gas phases.

3. Relate observations regarding the addition of energy by warming to increased particle motion.

4. Explain, at the particle level, how a thermometer measures the temperature of the system.

5. Explain the basis for the Celsius temperature scale.

6. State the basic tenets of Kinetic Molecular Theory (KMT) as they relate to gases:
   - Particles of a gas:
     a. are in constant motion, moving in straight lines until they collide with another particle or a wall of the container in which they are enclosed.
     b. experience elastic collisions; i.e., they do not eventually “run down”.
     c. do not stick to other particles.
   - The speed of the particles is related to their temperature.
   - The pressure of a gas is related to the frequency and impact of the collisions of the gas particles with the sides of the container in which they are enclosed.

7. The 3 variables P, V and T are interrelated. Any factor that affects the number of collisions has an effect on the pressure. You should be able to:
   - Predict the effect of changing P, V or T on any of the other variables.
8. Explain the basis for the Kelvin scale. Keep in mind that one must use the absolute temperature scale to solve gas problems.

9. Use factors to calculate the new $P$, $V$ or $T$. Make a decision as to how the change affects the variable you are looking for.

**Sequence**


2. Demonstration/discussion: diffusion of liquids in hot and cold liquids. Prepare a storyboard contrasting the arrangement and behavior of the particles in hot and cold water diffusion.

3. Observe the three states of matter and relate observations of rigidity vs. fluidity and density differences to the arrangement and behavior of matter particles. Eureka videos 1 – 3

4. Define temperature, Eureka videos 4 – 5, demo on thermal expansion of liquids, explanation of how a thermometer works, worksheet 1 – temperature and motion of molecules

5. Notes on pressure and how devices can measure pressure

6. Worksheet 2 – measuring pressure

7. PVT lab – part 1: $P$ vs. $V$, PVT lab – part 2: $P$ vs. $n$

8. PVT lab – part 3: $P$ vs. $T$, discussion on laws

9. KMT; how theories differ from laws

10. Solving PVT$n$ problems, ws 3

11. Whiteboard ws 3, unit 2 review

12. Unit 2 test
Instructional Goals and Sequence of Teacher Notes – Unit 3

Unit 3 – Energy & States of Matter – Part 2

Instructional goals

1. Relate observations regarding the addition of energy by warming to increased particle motion.

2. Describe the characteristics of solids, liquids and gases in terms of particles and their:
   • Arrangement: use particle diagrams to account for motion and density differences; describe the process of how the arrangement of matter particles changes during phase changes.
   • Attraction: infer the necessity of an attractive force between particles at close range from observations of differences in cohesiveness of the three phases;
   • Behavior: describe and contrast motion the three phases singly, and during phase changes.

3. Recognize energy as a conserved, substance-like quantity that is always involved when a system undergoes change.

4. Recognize that energy is stored in an object or system in several ways; for now we restrict our discussion to:
   • Thermal – due to the motion of the particles. The thermal energy depends on the mass and the velocity of the particles. The temperature of a system is a measure of its thermal energy.
   • Phase – due to the arrangement of the particles in solid, liquid and gaseous phases. Attractions lower the energy of a system; therefore, solids have the lowest phase energy because the particles are bound most tightly, liquids have greater energy because they have more freedom of motion, and gases have the greatest amount of energy because the particles have overcome the attractions that hold solids and liquids together.

5. Describe the ways that energy is transferred between the system and the surroundings. These are:
   • Heating – transfer of energy through the collisions of particles
   • Working – transfer of energy when macroscopic objects exert forces on each other
   • Radiating – transfer of energy by the emission or absorption of light

6. Draw energy bar graphs to account for energy storage and transfer in all sorts of changes. Make up a sample situation and sketch the bar graph.
7. Given a heating/cooling curve for a substance, identify what phase(s) is/are present in the various portions of the curve, and what the melting and freezing temperatures for the substance are.

8. Given a heating/cooling curve for a substance, identify which energy storage mode is changing for the various portions of the curve.

9. Given a situation in which a substance at a given temperature undergoes a change (in temperature, phase or both), sketch a heating/cooling curve that represents the situation.

10. State the physical meaning of the heat of fusion ($H_f$) and heat of vaporization ($H_v$) for a given substance and use these factors to relate the mass of a substance to the energy absorbed or released during a phase change (at the melting or freezing temperature).

11. State the physical meaning of the heat capacity ($c$) of a substance and use this factor to relate the mass and temperature changes to the energy absorbed or released during a change in temperature (with no phase change).

**Sequence**

1. Distinguish between heat and temperature, Eureka episode #6. Predict the effect of the addition of energy to a system of particles in the solid or liquid state.

2. Lab: Icy Hot, collect data, plot graph of temperature vs. time

3. Introduce new energy storage mode: phase energy, $E_{ph}$, post-lab discussion

4. Worksheet 1 in class, assign worksheet 2

5. Whiteboard ws 1 and 2, notes on heat capacity

6. Notes on quantitative problems, worksheet 3

7. Multi-step problems, worksheet 4
8. Whiteboard ws 4
9. Quiz 2, unit 3 review
10. Unit 3 test

Instructional Goals and Sequence of Teacher Notes – Unit 6

Unit 6 – Chemical Reactions: Particles and Energy

Instructional goals

1. Describe chemical changes in terms of rearranging atoms to form new substances.
2. Recognize that the total number of particles (sum of the coefficients) can change during a reaction because of differences in the bonding ratios of each substance.
3. Recognize that the total number of atoms does not change during a reaction because every reactant atom must be included in a product molecule.
4. Learn to describe reactions in terms of macroscopic observations.
5. Learn to describe reactions in terms of microscopic behavior of atoms.
6. Learn to write balanced equations to represent these changes symbolically.
7. Explain that the coefficients in a chemical equation describe the quantities of:
   a. the individual atoms or molecules involved
   b. the moles of the substances involved.
8. Observe basic patterns in the way substances react and learn to generalize them to other reactions students encounter.
   • Synthesis reactions
   • Decomposition reactions
   • Combustion reactions
   • Single replacement reaction
   • Double replacement (ionic) reactions
9. Describe endo- and exothermic reactions in terms of storage or release of chemical potential energy.
Sequence

1. Nail Lab
2. Rearranging Atoms Activity
3. Worksheet 1: Balancing Equations
4. Worksheet 2: More Balanced Equations
5. Quiz: Balancing Equations
6. Worksheet 3: Writing and Balancing word equations
7. Reaction Types Lab
8. Summary of reaction Types
9. The role of energy in reactions
10. Unit Test

Instructional Goals and Sequence of Teacher Notes – Unit 8

Unit 8 – Further Applications of Stoichiometry

Instructional goals

1. Use the concept that $P \propto n$ to determine the partial pressure of a particular gas in a mixture.
2. Determine by experiment the volume of a mole of gas at STP.
3. Develop and use the ideal gas law equation to determine the number of moles in a sample of gas not at standard conditions.
4. Relate the molar concentration (molarity) of a solution to the number of moles and volume of the solution.
5. Use information regarding molarity and volume to determine the number of moles of a reactant or product in a given volume of solution.
6. Sketch energy bar graph diagrams to represent energy storage and eventual transfer in exothermic and endothermic reactions.

7. Extend the use of the BCA table, first used in situations where the mass of reactant or product was known, to cases involving volume of a gas or solution or energy of reaction.

**Sequence**

1. Revisit gases and KMT. Relate pressure to the number of particles. Develop concept of partial pressure. Worksheet 1

2. Review ws 1, pre-lab discussion involving the problem: how could we determine the volume of a mole of gas?

3. Lab: molar volume of a gas

4. Wrap up discussion of lab; introduce ideal gas law.

5. Worksheet 2 – stoichiometry involving volumes of gases.

6. Molar concentration, worksheet 3


8. Heat of combustion lab

9. Post-lab discussion, worksheet 4

10. Review worksheet 4, review for test

11. Unit test

• **B - BECI**

The Basic Energy Concept Inventory presented here was taken directly from the Chemistry Modeling Materials located on the modeling website. [Omitted for security.]

• **C - CCI**

The Chemical Concepts Inventory presented here was taken directly from the Chemistry Modeling Materials located on the modeling website. [Omitted for security.]
D - Questions Analyzed from the Inventories

Students of this study took the BECI and CCI as previously stated. Certain questions of each test pertained to energy in chemistry. Investigators analyzed questions 3, 4, 7, 8, 9, 10, 11, 12, 13, 15, 16, and 22 on the BECI. They analyzed questions 9, 16, and 17 of the CCI. … [Omitted for security.]

E - Sample Questions Used in the Interviews

The developer of these questions, Guy Ashkenazi, provided verbal permission during the 2007 chemistry modeling workshop to use his 6 questions in our study.

These questions are now part of the Chemistry Modeling Materials …

During the interviews, students answered the following questions. The updated CCI has these six questions attached to the end. Since these questions pertained to energy and temperature change, investigators selected these questions as basic energy knowledge that all students should have a grasp on by the end the school year. [Omitted for security.]