2012

The Impact of General Chemistry Course Structure on Students' Exam Performance, Attitudes, and Problem Solving Strategies

Heather Anne Caruthers
Iowa State University

Recommended Citation

Caruthers, Heather Anne, "The Impact of General Chemistry Course Structure on Students' Exam Performance, Attitudes, and Problem Solving Strategies" (2012). *Graduate Theses and Dissertations*. 12637.
https://lib.dr.iastate.edu/etd/12637

The impact of general chemistry course structure on students' exam performance, attitudes and problem solving strategies

by

Heather Anne Caruthers

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY

Major: Chemistry (Chemical Education)

Program of Study Committee:
Thomas A. Holme, Major Professor
Thomas Andre
Thomas Greenbowe
Malika Jeffries-EL
Arthur Winter

Iowa State University
Ames, Iowa
2012

Table of Contents

CHAPTER 1: INTRODUCTION 1
National Issues in STEM Education 1
National Reforms in STEM Education 1
State of STEM Education at Iowa State University 3
Theoretical Framework 6

CHAPTER 2: IMMEX 12
Introduction 12
Artificial Neural Networks And the IMMEX system 14
Studying Problem Solving in CHEM 167 Using IMMEX 20
Discussion of Quad Scores Within and Across Problem Sets 23

CHAPTER 3: COURSE DESCRIPTIONS 28
Introduction 28
Course Descriptions 28
Lecture and Exam Content Coverage 31

CHAPTER 4: COURSE STRUCTURE EFFECTS ON EXAM PERFORMANCE AND ATTITUDE 47
Introduction 47
Establishing Equivalence 47
Overarching Theme 55
Research Question 55
Hypothesis 56
Method 56
Research Question 59
Hypothesis 59
Sub Hypothesis 59
Method 59
Exam Averages: Quartiles 61
Sub Hypothesis 65
Method 65
Sub Hypothesis 68
Method 68
Results 69
Findings/Conclusions about Exam Averages 69
Delta Z-Scores as Measures of Exam Performance 72
Sub Hypothesis 74
Delta Z-Scores by Quartiles 74
Delta Z-Scores by No Laboratory/Laboratory in Engineering Class 80
Delta Z-Scores by TA Experience 81
Findings/Conclusions for Delta Z-Scores 82
Attitude Data 83
Emotional Satisfaction 86
Attitudes about Chemistry Lab versus No Lab Engineering Course 96
TA Experience Correlations with Students' Attitudes 98

Attitude Findings 100

CHAPTER 5: ANALYSIS OF INTERVIEW DATA 101
Introduction 101
Population Description 102
Interview Description 103
Description of Participants in Analysis 104
Interview Questions used for Analysis 106
Research Questions 107
Analysis 108
Description of Problem Solving using Gram-to-Gram Conversion 110
Analysis of Familiar Stoichiometry Task 113
Problem Solving Behaviors 113
Time on Behaviors 115
Common Paths to Solving Familiar Task 117
Common Errors When Solving Familiar Task 120
Findings for Familiar Task 122
Analysis of the Unfamiliar Stoichiometry Task 123
Problem Solving Behaviors 123
Time on Task 126
Common Paths to Solving Unfamiliar Task 128
Bill – Wrong Answer with One Equation 129
Jeff – Wrong Answer with Two Equations 132
Cullen – Gave Up with One Equation 134
Seth – Gave Up Early with Two Equations 136
Joseph – Gave Up used Equations as Relationship 137
Relationships using Balanced Equations 141
Common Errors and Overall Findings for Unfamiliar Task 146
OVERALL FINDINGS FOR THE CHAPTER 147
Implications for Teaching 149
Limitations of Study 150
Future Work 151

CHAPTER 6: CONCLUSIONS 152
Summary of Findings 152
Implications for Teaching 154
Future Work 155

ACKNOWLEDGEMENTS 157

BIBLIOGRAPHY 159

APPENDIX 163
QUANTITATIVE CHAPTER 163
EXAM AVERAGES – COMPARING QUARTILES WITHIN COURSES 163
QUALITATIVE CHAPTER 175
IMMEX CHAPTER 181

CHAPTER 1: INTRODUCTION

National Issues in STEM Education

Science, technology, engineering and mathematics (STEM) education has been a focus of research as a way to maintain and improve the United States' position in the global economy (Chen & Weko, 2009). There have been concerns not only about getting students to enter into STEM fields, but how to keep students in these fields throughout an undergraduate degree program (Chen & Weko, 2009). There are also concerns about students in middle-school and high-school not being proficient in math and science, possibly due to improperly trained teachers (Kuenzi, 2008). There is of course also the concern that the students learn material that is relevant and that the coursework they take prepare them not only with the content they need, but also the problem solving and other higher-order thinking skills needed to adapt to ever-changing technology and job conditions. With these concerns in mind, it is not surprising that the government has provided a great many grants to try and improve STEM education in the United States (Kuenzi, 2008; Labov, Singer, George, Schweingruber & Hilton, 2009). With this funding comes research into teaching practices that improve student learning and retention of information as well as retaining students in STEM fields. Some of these teaching practices are presented below.

National Reforms in STEM Education

A National Academies white paper on positive practices in STEM Education at the undergraduate level has several suggestions for STEM instructors who are developing curricula (Froyd, 2008). These suggestions include preparing sets of learning outcomes, organizing students into small groups within a course, as well as grouping students into learning communities across courses within a particular STEM discipline, organizing content around scenarios instead of a list of topics, using formative assessment in a systematic way to give students feedback about their learning, designing or using in-class activities that engage students in the material, undergraduate research and professor initiated student-faculty interactions (Froyd, 2008). Many of these practices have been demonstrated to have positive impacts on students learning. Two standards, one of implementation and the other of student performance were used to determine the effectiveness of these practices in the classroom. The implementation standard relates to how easily and likely a faculty member is to implement the practice. The student performance standard relates to students retention in the STEM field and how much their performance on either exams or nationally normed exams improves from the use of the practice. Eight practices were evaluated; preparing a set of learning outcomes, organizing students into small groups, using learning communities, scenario-based content organization, feedback using systematic formative assessment, designing in-class activities to actively engage students, undergraduate research and faculty-initiated approaches to student-faculty interactions. Of these eight practices, five were rated as strong to good on the two standards used to assess the effectiveness. The practices that were the easiest to implement and lead to the most improvement in student performance and/or retention in the STEM field were the preparation of learning outcomes, organizing students into small groups, basing content organization on scenarios, using formative assessments to give students feedback on their performance and designing in-class activities to engage students in the material (Froyd, 2008). The use of small group learning, demonstrations as in-class activities, active learning strategies, and the use of clickers as a method for producing formative assessments have all been shown to be effective elsewhere (Lyon & Lagowski, 2008; Majerich & Schmuckler, 2007; Sharma, et. al, 2010, Walker, Cotner, Baepler & Decker, 2008, King 2011). Many of these suggestions are already being used to improve student learning and retention at Iowa State University in the chemistry department.

State of STEM Education at Iowa State University

At Iowa State University, there are four main general chemistry courses for science majors, agricultural majors, and engineering students, as well as remedial course for students who did not take chemistry in high school. Of particular interest are the engineering and science majors general chemistry courses. These two courses were chosen for comparison because there is large number of similarities between them but also some key differences, which allow for comparisons of course structure and there influence on students' exam performance, attitude, and problem solving skills. The two courses are taught in large lecture formats, with about 1000 students in each course. There are also opportunities for students to work in smaller groups, in recitations and depending on the course and major, laboratories run by graduate student teaching assistants (TA's). These TA's come in with different levels of experience and expertise in teaching general chemistry. That being said, the majority of TA's in the general chemistry courses are new graduate students. To prepare these new TA's for teaching, several days of TA training occur before the beginning of the semester, with the new graduates getting safety training, and practice teaching recitations and experiencing laboratory courses.

In addition to the large lecture format and use of graduate student TA's, other similarities between the Engineering and Science Majors courses include the use of experienced professors, both with general chemistry teaching in general, but with these courses specifically, and the use of active learning strategies. These active learning strategies include using formative assessments that will be discussed later, asking students to work in small groups to discuss content that was just presented or to work on tasks together, and asking students questions along the way, to move the lecture forward, to name a few (Holme T. , 1992). Active learning strategies like the ones discussed above are shown to improve students learning (Walker, Cotner, Baepler, & Decker, 2008). The major differences between the two courses are the content coverage, whether the laboratory is required or optional, and the motivational tool used in each course to encourage students to learn material they missed earlier in the semester, either resurrection points (Herschbach, "Resurrection" Points, 1997) on the final exam, or a replacement exam between the last hour exam and the final.

There are particular curricular challenges in the Engineering course that leads the general chemistry course for that group of students to be a one-semester survey course. There are three Engineering majors that are required to take the Science Majors course, Materials Engineering, Civil Engineering and Chemical Engineering. The other eight kinds of Engineering major offered at Iowa State, ranging from Aerospace, Computer, Electrical, Agricultural to Biological Systems Engineering, are required to take the Engineering general chemistry course. The students take between 121.5 and 129.5 credits over four years. Many of these students are taking the Engineering general chemistry course, along with three introductory engineering courses and calculus. Most of the students take general chemistry their first semester at Iowa State because it prepares them for their future coursework, and it's the only time it would fit into their schedules. There are so many other Engineering courses they are required to take, that a one-semester survey course is the only option they have to learn chemistry.

But within these course constraints, there are opportunities to help students in this large lecture atmosphere. The lectures in both the Science Majors course and the Engineering course were audio and video recorded so that students could get the information from the class if they were absent, or could review the information outside of class if there was material that didn't make sense to them. Another resource available to the students in both courses, is the Help Center, a room that is staffed by graduate student TA's where students in any general chemistry course can go doing the week, to get assistance with course material that they don't understand, or help with their laboratory reports. Online homework is also used in both courses as a way to challenge students to apply concepts from class, in the Engineering course, and to work with new material as well as applying known concepts in the Science Majors course. Finally, formative assessment, that is determining if students understand the material recently presented, is important in any course. In the Science Majors and Engineering courses the formative assessments usually took the form of clicker quizzes. These quizzes were usually related to recently presented material and the student was expected to work by him or herself to come up with an answer first, then talk to a neighbor about their results. The students entered their responses using a personal response device or clicker that was identifiable as theirs, so that they could get credit for giving an answer. These clicker quizzes gave the professors almost immediate feedback as to what the students did or did not understand about recent material (King, 2011). The similarities and differences between the two courses will be discussed in more detail in chapter three.

Theoretical Framework

The theoretical framework for the analysis of this work is the Unified Learning Model, which combines the underlying ideas of several current models into one model. The Unified Learning model has three components, working memory, knowledge, and motivation (Shell, Brooks, Trainin, Wilson, Kauffman, & Herr, 2010). The three basic principles of learning in this model are that learning is product of working memory allocation, working memory's capacity for allocation is affected by prior knowledge, and that working memory allocation is directed by motivation (Shell, Brooks, Trainin, Wilson, Kauffman, & Herr, 2010). Since working memory is the focus of this model, it is important to not overwhelm the students' working memory in order for them to learn new material. The students' prior experiences, especially with high school chemistry can effect what they pay attention to in class. Finally, how students are motivated also effects what they're paying attention to when they come to class, or do the homework. In the Science Majors and Engineering classes, the professors present information in relatively small chunks so as not to overload their working memories. The homework, especially in the Science Majors class is broken down into parts to help students to work towards to final concept, to keep their working memory from being overwhelmed. The professors in both classes also focus on making connections between the chemistry content and the students' prior knowledge/experiences and topics of interest to their majors. In this way they're making connections to help students store the information into long-term memory. Also, by making connections to prior content or experiences, they are helping students chunk information, to reduce the strain on their working memory while they are learning new material. In addition, making connections to prior knowledge changes how and what the working memory focuses on when students' are learning new material. The professors took stock of their students' prior chemistry knowledge by giving them the chemistry questions of the Toledo Test at the beginning of the semester. Finally, motivation affects what students' pay attention to in class. This means if the student is motivated to learn for learning's sake, they may pay more attention in class than someone who is working to earn enough points to pass the class. Motivating students to learn missed material on previous exams takes place in both classes, and the effects of the tools for doing this will be assessed in terms of students' exam performance in chapter four. A student's level of motivation may fluctuate across a lecture or a semester, but overall students' seem to be motivated to do the amount of work, and earn the amount of points they need to earn the grade they desire in a course. This is the backbone of the analysis on exam performance in chapter four. Motivating students to pay attention, or allocate enough of their working memory to learn new material, is a main way professors can get students to pay attention in their class. The influence of the motivation tool used in two courses, resurrection points in Engineering and a replacement exam in the Science Majors course, on the students' final exam performance will be part of the analysis in chapter four.

By comparing these two courses, the influence of differences in the course structures on their exam performance, attitude towards chemistry, and problem solving skills can be determined. Questions about movement relative to the course median, in terms of exam averages, due to course structure differences, i.e. survey versus two-semester sequence, differences in student final exam performance in the two courses, the impact of laboratory on student exam performance in the Engineering course, and the influence TA experience on student exam performance in the two courses will all be addressed in chapter four. The impact of the course structure, laboratory in the Engineering course and TA experience on students' attitudes about chemistry will also be discussed in chapter four. In the interview analysis chapter, the impact of how problem solving was presented in the context of stoichiometry, on students' problem solving strategies when presented with a familiar and a novel stoichiometry task are addressed.

In order to answer these questions about exam performance, attitudes about chemistry and problem solving skills, course performance data, including exam scores, homework and quiz scores, including clicker scores were collected. Information about laboratory and recitation enrollment were collected for both courses, to establish co-enrollment in the Engineering course, and to assign TA experience levels to all students. Determining if the students in the two courses were equivalent to start with is an important aspect of the analysis that follows, so two tests of prior knowledge, a departmental placement exam given over the summer and the Toledo Test, given during the first week of the semester, were used to try and establish an equivalent group across the two courses. This process will be described in detail in chapter four.

To assess changes in students' attitude toward chemistry, the Attitude towards the Subject of Chemistry Inventory version 2 (ASCI v2) was given during the first week of the course, and after the weeklong semester break. The ASCI v2 is a semantic differential with eight items that uses a seven-point scale to determine students' attitudes about chemistry (Xu & Lewis, 2011). There are two factors within the ASCI v2, emotional satisfaction and intellectual accessibility. The ASCI v2 is based on the work of Bauer (2008), who developed a 20-item survey with four factors and one pair of items, used to assess college students' changes in attitude about chemistry.

The final piece of data collected was interviews with 40 students from the Engineering and Science Majors courses. The interviews involved asking students about their chemistry background, then asking them to talk out loud while they worked on three stoichiometry and three thermochemistry questions. For each topic, one of the questions was a simulation that students worked through while answering questions. A copy of the interview guide is in the Appendix. The interviews were video and audio recorded and transcribed prior to analysis. The analysis of the non-simulation stoichiometry questions and why they were chosen will be discussed in chapter five.

"One of the essential unanswered questions about effective STEM practice is what approaches to teaching, learning, and assessment transcend the disciplines (and are thus appropriate for use in almost any setting) and what approaches are more discipline-specific. Also unclear is which practices that seem to work well within a discipline can be used in multidisciplinary or interdisciplinary approaches to teaching and learning"(Labov, Singer, George, Schweingruber, & Hilton, 2009). There is a call for research into practices that can be useful in multiple STEM disciplines that will improve student learning. Because the course structures are the main area of comparison for the Engineering and Science Majors courses, questions can be answered about what aspects of the courses are leading to improved exam performance and attitudes about chemistry. Course structure differences between the two courses include whether or not the laboratory is required, replacement exams or resurrection points to motivate students to learn on their own, and the depth and breadth of content coverage. If any aspects of the two courses lead to improvements in these areas, then they can be transferred to other courses in chemistry, or other disciplines, because the differences will be due to the course structure, not the content itself.

As far as the national importance of the qualitative research presented in the study, there are no comparisons in the problem solving literature between students solving algorithmic tasks that can be solved with a common procedure and algorithmic tasks that appear to require the same common procedure at first glance, but that can be solved using a simpler method, in this case algebra, upon closer inspection. The analysis of the problem solving behaviors exhibited by students while working on these types of tasks will add to the literature base in this area of problem solving research.

Outline of Five Chapters to Come

Chapter two will discuss a project studying students' problem solving skills when faced with complex problems. The IMMEX project used complex real-world problems to assess changes in students' problem-solving skills over the course of five problem sets in the Fall 2009 semester in the Engineering course. Students were classified along efficient and effective lines into four quadrants, or Quad Scores, for each assignment. Changes in these Quad Scores over the course of semester will be discussed, in particular how they relate to students comprehension of the context of the questions.

Chapter three moves back to the Fall 2010 data from the Engineering and Science Majors courses. The chapter reviews the course structure and content coverage in both courses in the lecture and on the four one-hour exams throughout the semester. The course content in both the Engineering and Science Majors courses is also aligned with the general chemistry content map for the undergraduate chemistry curriculum(Murphy, Holme, Zenisky, Caruthers, & Knaus, 2012). This description of the similarities and differences between the two courses, particularly the content coverage in the two courses, sets the stage for the analysis of the quantitative and qualitative data used to compare students exam performance, attitudes towards chemistry and problem solving skills in the two courses.

Chapter four compares the exam performance of students in the two courses as well as changes in their attitude toward the subject of chemistry. Exam averages, as well as changes in exam performance as measured by delta z-scores between exam pairs, are discussed. The final exam is one of the focuses of analysis as the motivational tools used in the each course are either on the final exam in the Engineering course or right before it in the Science Majors course. Another aspect of the analysis in the quantitative chapter is pre/post changes in the students' attitudes towards chemistry. The Engineering and Science Majors courses were compared to determine if one course structure led to major changes in attitude over another.

The fifth chapter compared students' problem solving skills when faced with familiar and unfamiliar tasks. Stoichiometry tasks were chosen as the topic for comparison between courses as it underpins much of the rest of the content in a general chemistry course. A discussion of how content related to a subset of stoichiometry, gram-to-gram conversion, was presented in each course occurs. The influence of how stoichiometric problem solving is presented in each course on students' problem solving when presented with familiar and unfamiliar tasks is analyzed. The conclusion chapter presents a summary of the findings of the study and implications for teaching. Future work is also presented.

CHAPTER 2: IMMEX

Introduction

In addition to instilling in students the factual knowledge about chemistry, the chemistry curriculum is expected to teach students about problem solving. In order to measure this problem solving, professors may put short answer or open-ended questions on their exams and say that they are measuring student problem solving. But are they actually measuring student problem solving, particularly in a general way? Measuring student problem solving requires taking into account several difficulties, including the fact that problem solving is a complex task that is influenced by the task difficulty, the student's prior knowledge and metacognitive abilities, and the fact that problem solving is dynamic, so the strategies used by students change with experience (Stevens & Thadani, 2007). Also the questions used to study problem-solving need to be problems for students not exercises (Bodner G. , 2003). A problem is defined as a gap between where you are now, the information you have available, and where you want to be, or the information you have at the goal state (Bodner G. , 2003). Bodner also adds that there must be a level of uncertainty about how to get from the initial state to the goal state, and that the main difference between an exercise and a problem is not complexity or difficulty, but familiarity (Bodner G. , 2003). Finally, to be useful from an instructional standpoint, measurements should be made quickly and be readily understood by the student and the instructor, so that interventions can be offered when necessary (Stevens & Thadani, 2007).

Student problem solving has been studied using a variety of methods including talk-alouds, where the student verbalizes their thought process when working on a problem, and by using a combination of a measure of self-efficacy, and types of problems to categorize and solving quantitative questions (Bowen, 1994; Taasoobshirazi & Glynn, 2009; Stevens, Ikeda, Casillas, Palacio-Cayento, & Cylman, 1999). The present study will utilize IMMEX (Interactive Multimedia Exercises), an online system that tracks what links a student clicks on within an assigned problem set, and uses that data to determine the student's problem solving abilities and strategies. The use of this system, will allow for many of the challenges of measuring student problem solving to be overcome. The IMMEX system requires students to solve ill-defined problems that relate chemical concepts learned in class to real-life situations. The ability to study problem solving as it relates to these different concepts could be used as part of a programmatic assessment in a chemistry department. By using ACS exams, one could measure students' conceptual and factual knowledge about topics covered in a particular course, and using these IMMEX problems, you could study how the students apply these concepts to real-world situations.

Each problem set in IMMEX is designed to have a general description of the situation, and have links to all the data and background information a student might need to solve the problem. Each problem set has different examples, or clones, with different values for variables, or different compounds to identify, and the exact example given to each student is randomized. A computer system tracks data behind-the-scenes about what links the students click on within a problem set and collects that information into a database. The information in that database is then analyzed using artificial neural networks, to produce thirty-six histograms of the different menu items, and the likelihood of the item being clicked on (Stevens, Ikeda, Casillas, Palacio-Cayento, & Cylman, 1999; Vendlinski & Stevens, 2002; Cooper, Stevens, & Holme, 2006).

Artificial Neural Networks And the IMMEX system

Artificial neural networks can be used for pattern recognition in a variety of circumstances. Artificial neural network models can be understood as analogous to biological neurons. A neuron in the brain is called a node in an artificial neural network (ANN). The synapse is like the connection between nodes. The synaptic efficiency or size of the signal received by a neuron is similar to the connection weight in an ANN. And finally, the firing efficiency of a neuron in the brain, or sending a signal to another neuron when the neuron has been excited over a threshold point, is called node output in an artificial neural network (Mehrota, Mohan, & Ranka, 1997). Nodes within an artificial neural network, can do one of three jobs, pass information to another node, do a computation with the information, or produce an output or result. The node that passes along information, including output of nodes, to other nodes are input nodes. The nodes that produce results are called output nodes. The simplest artificial neural network is one with one input node, one node to do calculations and one output node. The output node is usually designed in such a fashion, that it outputs one result if the calculation result is above some predetermined threshold and another if it's below the threshold, usually one and zero. From there, it is possible to increase to two input nodes, one calculation node and an output node. At this point, the connection weights between the input and calculation nodes become important. Once a calculation has been assigned to a node, and a threshold has been set for the output node, the most important part of using an artificial neural network becomes adjusting the connection weights to optimize the system. For example, the calculation is summing the information from the input node times the connection weight, $f(x) = \Sigma (w_i * l_i)$ over all the input nodes, where w is the connection weight and l is the input from the input node. In the case of the simplest ANN, the connection weight would be one, as there is only one input node passing information to the calculation node. Once there is more than one input node, the system needs to adjust the connection weights between input nodes and the calculation node so that the calculations are being carried out in an efficient manner. The goal of any pattern matching artificial neural network is to have a particular set of inputs produce the same outputs consistently.

There are many ways to adjust the connection weights to insure that the inputs are correctly grouped together. The main differences depend on the type of learning environment being used, either supervised (confirmatory) or unsupervised (exploratory). In each case, the connection weights start out at a randomly assigned value, usually between 0 and 1, and are adjusted from there. In a supervised learning environment, a set of test cases are presented to the ANN, and the weights are adjusted, usually by adding or subtracting a small value, until the cases have all be separated into the appropriate group. This may take many presentations of the test cases, and adjustments may be made after each case is presented or after the whole batch of cases has been presented. The key is that the operators of the system know which group the cases belong to, and what the desired outcome is for the calculation. The operators just allow the system to make adjustments until the desired outcome is reached. At that point, training is complete and new cases can be presented to the trained network, and assigned to groups (Mehrota, Mohan, & Ranka, 1997).

In an unsupervised learning environment, the operator does not know ahead of time which groups the inputs belong to. The goal of using exploratory ANNs is to group information together based on the similarity of given characteristics within the data. In this environment, again the connection weights are set to a random value, usually between zero and one. The network tries to optimize the connection weights between the input nodes and the calculation nodes, either by calculating distances, or by having several layers of nodes, some of which have inhibitory connections to other nodes, that can be used to determine the overall “winning” node, the one with the highest activation or a combination of both. In the case of the distance calculation, the goal is to adjust the connection weights so that the distance between similar patterns of information is minimized and maximized between different patterns of information. With the “winning” node method, there are connections between the input nodes and the goal is to decide which node has the highest initial input value. To do this, the function, $f(total) = max\ (0,total)$, where $total = \Sigma\ (w_i * x_i)$., and w is the connection weight and x is the input from the input node. Within this layer of nodes, there are inhibitory connections, that decrease the value of f(total) for a given node, so that over many iterations, one node will have a value of one and the rest of the nodes will have a value of zero. The node with the value of one will be declared the winner.

The distance calculation method and the winning node calculation can be combined into a simple competitive learning network. In an example of a simple competitive learning network, a layer of input nodes are connected to a layer of output nodes and the output layer has intra-layer inhibitory connections. Each input node is connected to all the output nodes and the connection weights are all set to a random value to start. The only thing that varies within the network is the connection weights that are updated based a calculation that allows one output node to be assigned to each unique input. Therefore, the connection weights should be representative of each input pattern. Each node in the output layer is described by the vector, or combination of the connection weights between all the input nodes and that output node. So for example, if there are 3 input nodes, and 4 output nodes, then the second output node’s weight vector would be ($w_{2,1}$, $w_{2,2}$, $w_{2,3}$). So the goal of a competitive learning network is to find the output node that best matches the weight vector of a set of inputs.

The network starts all the connection weights out at random values and picks an input pattern from the training set, and determines which output node best matches the input vector, using the distance calculation described above. The output nodes then output their information, usually a one for the winning node, and a zero for all the others. Each weight is then updated for the winning node using the pre-determined rule, which makes the winning node's vector more similar to the input vector (Stevens, 2008). This process is repeated until the weight vectors converge, or move closer to a predetermined value, at which time the training in complete (Mehrota, Mohan, & Ranka, 1997).

There are other methods for designing unsupervised (exploratory) artificial neural networks, but a competitive learning network is the one used by IMMEX, and so is the most important for the work reported here. The IMMEX system uses a competitive artificial neural network to categorize the menu items that a student clicks on while solving a problem. The IMMEX artificial neural network uses thirty-six nodes to classify these combinations of menu items and produces a pattern of nodes wherein nodes that are near each other in space, represent similar methods for solving the problem. Each node is represented using a histogram that shows the probability that each menu item would be selected by a student when solving a problem (Stevens, 2008). Given a node number and the histogram, a researcher will know what items a student looked at while solving a particular example within a problem set. If a series of nodes were given, the researcher could study how a student's performance changed over time. Using Hidden Markov Modeling, one can look at the probabilities of a student in one state, or combination of nodes, moving to other states

(Vendlinski & Stevens, 2002). Using other information, about the solve rate, or whether or not a student solved the example, along with the menu items viewed, a researcher can determine the efficiency and effectiveness of the strategy used by the student. This information can lead a professor to suggest an intervention for lower performing students. If the problem sets were given in a pre-test, post-test fashion, then the professor could use the information from the artificial neural network to determine if their intervention helped students become more efficient problem solvers.

Measures of the efficiency and individual item difficulty can also be gathered from the database. A measure of efficiency can be combined with a solve rate, and be normalized so it can be compared across problem sets. These values are called quad scores, and can be used to show where a student is on a spectrum of guessing to being an efficient problem solver. As part of previous research using IMMEX, an artificially limited resource must be given to the students, usually virtual money, to ensure that the students are being selective about what links within the problem set they utilize. A series of IMMEX questions could be used to determine whether students are using efficient strategies to solve ill-defined problems using chemistry concepts in real-world situations.

The data gathered from IMMEX assignments is analyzed using an approach called learning trajectories(Stevens, Soller, Cooper, & Sprang, 2004). Comparisons of how novices and experts solve problems are used to develop learning trajectories(Stevens, Soller, Cooper, & Sprang, 2004). These learning trajectories can be visualized a set of quadrants that students may move through while working on a problem set. As there is one correct answer for each example within a given task, effectiveness is determined based on whether a student selected the correct answer to a particular example with in the assigned task. Table 1 shows what these four quadrants would look like. Students in Quad 1 are mainly guessing and in Quad 2, the students are working hard on improving their strategy, but they are still getting the question wrong. In Quad 3, the students are getting the questions correct but they are looking at more items than is absolutely required to solve the problem and in Quad 4, the students are getting the question correct and they are looking at the minimum amount of information needed to solve the problem. Students in Quad 4 would be described as efficient and effective problem solvers.

Table 1 Quad Scores and their definitions

Quad 2 ***Efficient*** ***Not Effective*** Working to improve strategy but still getting question wrong	**Quad 4** ***Efficient*** ***Effective*** Question correct and minimum amount of information used
Quad 1 ***Not Efficient*** ***Not Effective*** Mainly Guessing	**Quad 3** ***Not Efficient*** ***Effective*** Question correct but looking at more than minimum information

When a novice student first works on an IMMEX assignment, their understanding of the topic may be fragmented and incomplete, therefore they tend to click on a lot of items to decide what pieces are relevant to the process of answering the question, most likely ending up in Quad 1. As the student's problem solving ability improves, they tend to become more selective in the information they viewing when trying to solve the problem, eventually moving into Quads 2, 3, and 4. The goal for professors is for their students to move into Quad 4, where the student is using efficient and effective strategies for solving the problem. Ill-structured problems set in real-world contexts, such as those presented using the IMMEX system, have assumptions about prior knowledge, connections between different areas of knowledge and how knowledge develops that are described by constructivism and situated cognition (Jonassen, 2000). As students' problem solving skills improve, they are able to connect information from coursework and prior experience with the level of success they find when solving an IMMEX task, and move toward a more efficient and effective state, where they are only using the information necessary to successfully solve the task.

Studying Problem Solving in CHEM 167 Using IMMEX

IMMEX problem sets have been designed to cover a number of topics in general chemistry and organic chemistry, in a variety of real-world situations, see Table 2 in Appendix. So far, these activities have been used or are being used to study student problem solving in two general chemistry courses and three organic courses. A series of IMMEX questions, covering the topics of identifying elements or compounds, and states of matter (Model Madness), stoichiometry (How Much to Order), gas laws (Gas Laws on Planet Ardanda), thermochemistry (RXN) and the qualitative identification of an unknown (Hazmat), were assigned as homework in the Engineering class in the Fall 2009 semester. The Model Madness assignment was given fairly early on in the semester, to allow students to review content from high school chemistry as well as familiarize themselves with the IMMEX system. The other assignments were given as homework after the corresponding content had been covered in lecture and were open for five days.

Using data from the 650 students who gave their consent, quad scores, the normalized value that is a combination of a measure of efficiency combined with solve rate, were compared for each problem set to see if there is any improvement over the course of the semester in student strategies for problem solving, see graphs that follow. The graphs show the percentage of students in each quadrant after they have completed at least the minimum number of examples in each problem set, usually five, see Figure 1 through Figure 5. The graphs only show students' work on the final example within each task, because based on prior work with IMMEX, students' stabilize into a Quad Score after about 5 examples in a task (Stevens, Johnson, & Soller, 2005). Since at least five examples were completed in each task, only the stabilized Quad Scores are shown in Figures 1 through 5.

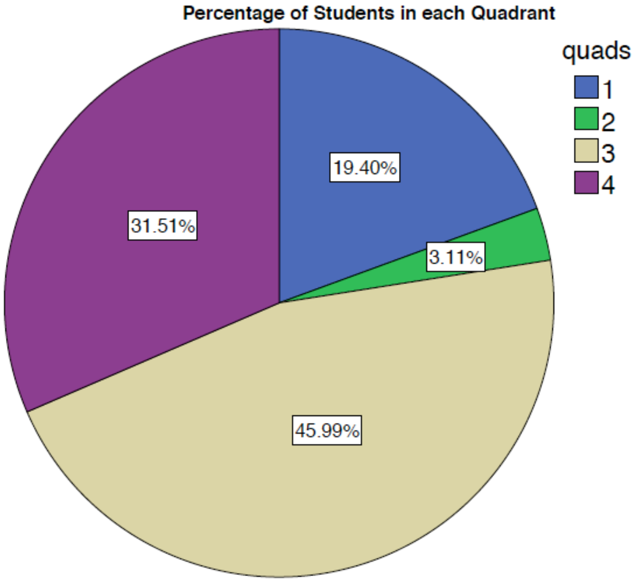


Figure 1 Percentage of Students with each Quad Score for Model Madness

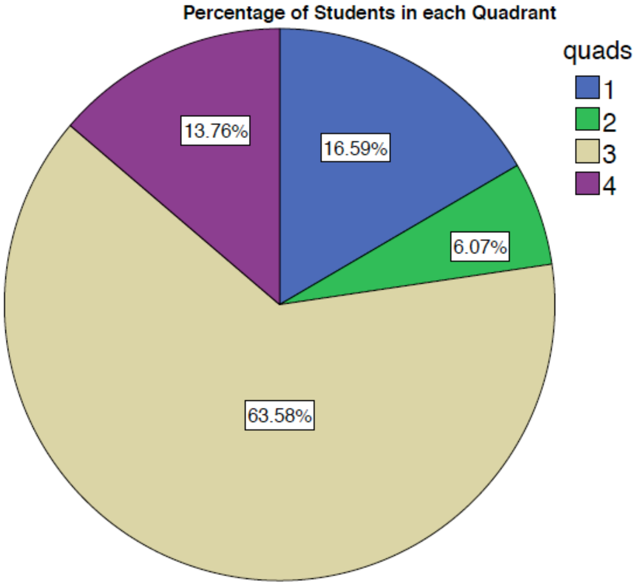


Figure 2 Percentage of Students with each Quad Score in How Much to Order

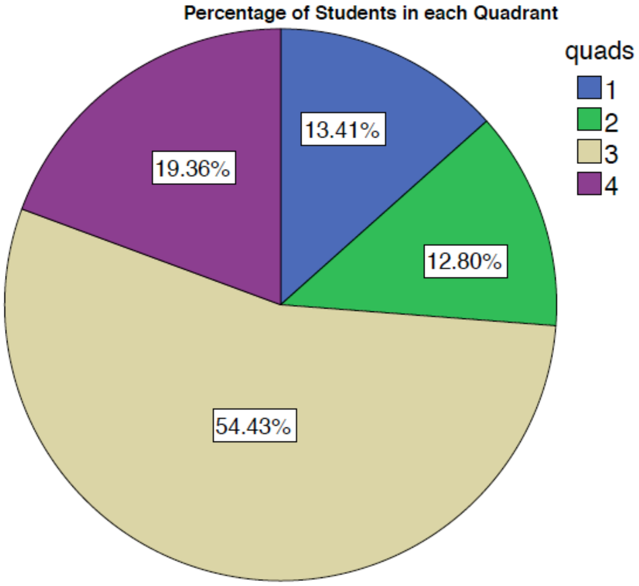


Figure 3 Percentage of Students with each Quad Score for Gas Laws on Planet Ardana

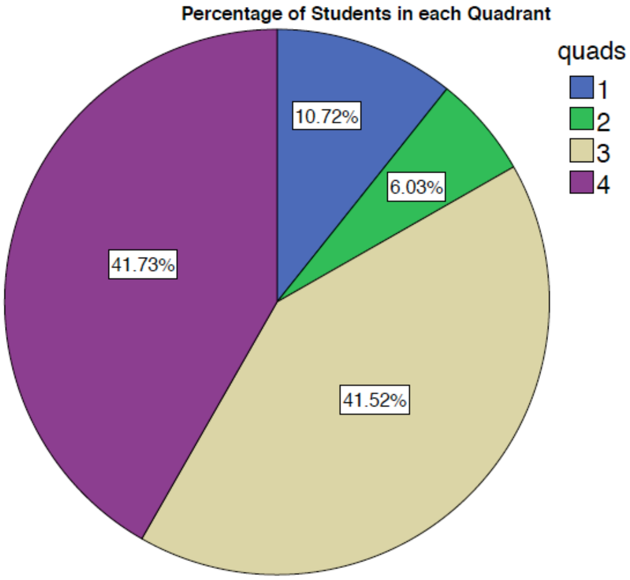


Figure 4 Percentage of Students with each Quad Score for RXN

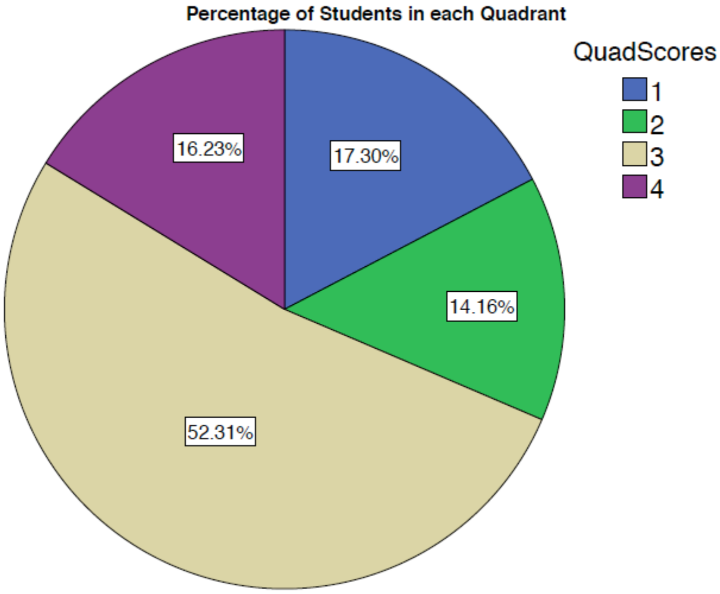


Figure 5 Percentage of Students with each Quad Score for Hazmat

Discussion of Quad Scores Within and Across Problem Sets

Due to the work of Stevens *et. al*. (2005), it has been shown that students' strategies for solving these complex real-world tasks tend to stabilize after about five examples within a given task. After that time, the students' strategies do not change appreciably. Based on this evidence, the results presented below represent real differences in student strategy use and not random variations in student performances.

In the Model Madness problem set students are categorizing substances as elements or compounds as well as the state of matter. For this problem set only 19.4% of the students are guessing (Quadrant 1) and 3.1% are transitioning to a more efficient and effective strategy (Quadrant 2). 46.0% of students are in Quadrant 3, meaning their problem solving strategy is effective but not effective as it could be. And finally, 31.5% of students are solving the problem correctly using the minimum amount of information to do so. The reason so many of the students are in the effective quadrants (Quadrants 3 and 4) is because this problem set is relatively simple and so students were able to quickly figure out what they need to do to solve the task.

In the How Much to Order problem set, students are working to calculate how much of a particular reactant would be needed to make a given amount of product using percent yield. For this problem set, 16.6% of students ended up in Quadrant 1 (guessing), 6.1% in Quadrant 2 (transitioning), 63.6% in Quadrant 3 (effective, but not totally efficient) and 13.8% in Quadrant 4 (effective and efficient). So students' problem solving strategies were less effective and efficient for this problem set than for Model Madness.

For the Gas Laws on the Planet Ardana problem set, students use data to determine what mixed up version of the ideal gas law is in effect in each case, the value for the gas constant and whether people could survive on that planet. For this problem set, again, we see an increase in the percentage of students in the effective quadrants compared to the How

Much to Order problem set. The percentage of students in each quadrant was as follows; Quadrant 1- 13.4%, Quadrant 2- 12.8%, Quadrant 3- 54.4%, and Quadrant 4- 19.4%.

For the RXN problem set, where students must determine the value of q, the heat of a reaction, as well as whether the reaction is exothermic or endothermic, there was the highest percentage of students getting the question right so far. The percentage of students guessing was 10.7% (Quadrant 1), the percentage transitioning to another strategy was 6.0% (Quadrant 2), the percentage getting the question correct while not being as efficient as they could be was 41.5%, and the most efficient and effective strategies used by students made up 41.7% of the strategies used on this problem set.

For the Hazmat problem set, students needed to use wet chemical tests and flame tests to identify a compound. For this problem set the data showed that students were moving back into less effective and efficient problem solving strategies as shown by the decease in the number of students in Quadrants 3 and 4, compared to the RXN task. There were 17.3% of students in Quadrant 1, 14.2% of students in Quadrant 2, 52.3% of students in Quadrant 3, and 16.2% of students in Quadrant 4. A partial explanation of this decrease in problem solving efficiency and effectiveness is due to lack of familiarity with the content of the question. Only about a third of students are in the laboratory and have an opportunity to work on a task like this, so many of the students resorted to less efficient problem solving strategies like guessing to try and answer this problem set. This is reflected in the relatively small percentage of students in Quadrant 4, as compared to the performance on the prior problem set (RXN).

The change in the percentage of students using the most effective and efficient problem solving strategies depends not only on the familiarity with the content of the task, but also on a students' familiarity with the problem solving process itself. In order for a task to be a problem, a student needs to be unfamiliar with either the content of the task, the process needed to solve the task, or both. The content for Model Madness, How Much to Order, Gas Laws on Planent Ardana, and RXN were covered in lecture before each assignment. Students were assumed to be familiar with the content of these tasks. The content for the Hazmat task was only covered in laboratory and since only about one third of students were co-enrolled in the laboratory, not all the students can be assumed to be familiar with the content of the Hazmat task. In order for a task to be a problem, a student needs to be unfamiliar with either the content of the task, the process needed to solve the task, or both.

Based on the percentage of students in Quadrant four for Model Madness and RXN, these tasks can be categorized as exercises for students. Based on the percentage of students in Quadrant four for How Much to Order, Gas Laws on Planet Ardana and Hazmat, these tasks can be categorized as problems for students. For the How Much to Order and Gas Laws on Planet Ardana tasks, the students were presumably familiar with the content as it was part of lecture, so the task became a problem because of the unfamiliarity with the process needed to solve the task. The percentage of students in Quadrant four does increase somewhat between these two tasks, indicating that when faced with complex real world tasks over the course of the semester, students' problem solving strategies are improving. When looking at the Hazmat data, it is also classified as a problem for students because of the small percentage of students in the Quadrant four, as well as the decrease in the overall percentage of students getting the question correct compared to the other four tasks. The Hazmat task is a problem for students not only because they are unfamiliar with the process to solve the task, but because they are also unfamiliar with the content of the task, due to a majority of students not taking the laboratory in the Engineering course.

Familiarity with the content and process of problem solving influences how effective and efficient the strategies used are when solving complex real-world tasks. It appears that trying to transfer those problem-solving skills to a task in a new content area lead to a decrease in the effectiveness and efficiency of those skills, particularly when the process was also unfamiliar. These results are supported by research in using analogies to solve two river crossing questions with similar premises but one additional constraint, and transfer of problem solving skills in statistics when taking into account cognitive load during training that have seen similar challenges for students in transferring skills (Reed, Ernst & Banerji, 1974;. Pass, 1992). When students are unfamiliar with the process needed to solve a task and/or the content of the task, the task becomes a problem for students, and they may have difficulty transferring problem solving strategies effectively to the new task.

CHAPTER 3: COURSE DESCRIPTIONS

Introduction

In order to answer the research questions posed in chapter one, a more lengthy description of the Engineering and Science Majors classes is required. Descriptions of the course and exam structures, including the amount of quantitative and conceptual content coverage will be discussed. In addition, the differences in chemistry content covered in the two courses will be addressed. The similarities and differences discussed in this chapter will set up a framework for understanding the analysis that follows in chapters four and five.

Course Descriptions

Two courses Iowa State University were compared, the Engineering class, a one-semester general chemistry course, and the Science Majors class, the first-semester of a two-semester sequence general chemistry course. Most of the students in both the Engineering and Science Majors classes are freshman. Of the three sections of the Engineering class that were taught in the fall 2010 semester, two were analyzed since they were taught by the same instructor in a 450 seat lecture hall. The other section was taught in a 200-seat lecture hall and met twice a week for 75 minutes. While the overall content coverage was the same, the pacing was faster in this section of the Engineering Class and therefore this section was dropped from the analysis.

The Engineering course was a survey course that discussed the topics in a two-semester sequence of general chemistry with a focus on engineering applications. The class meets three times a week for 50 minutes for lecture, as well as one time per week for 50 minutes in groups of about 25 students for recitation run by graduate student teaching assistants (TA's). There is a laboratory associated with the Engineering course, however it is not required for all students, only certain majors. As a result, only about one-third of the Engineering students were also enrolled in the laboratory. All assessments in the Engineering course were instructor-written, with the exception of the online-homework assignments. These assessments included four one-hour exams, quizzes given during recitation covering recently covered content, clicker quizzes given in lecture, and a comprehensive final exam.

The analysis of the Science Majors course included five sections that were taught by a total of three professors. These professors were all experienced and had taught the course together before. They coordinated their content coverage across the lecture sections by discussing the content that was to be covered in each chapter prior to that series of lectures. Lectures for the Science Majors course were held in either a 450 or a 250-seat lecture hall. The course met three times a week for 50 minutes for lecture, as well as one time per week for 50 minutes in groups of about 25 students for recitation run by graduate student TA's. The laboratory was a co-requisite for this course, so all students enroll in the laboratory. All assessments in the Science Majors course were instructor-written with the exception of the on-line homework and a portion of the final exam. These assessments included four one-hour exams, quizzes given during recitation covering recently covered content, clicker quizzes given in lecture, and the multiple choice algorithmic section of the comprehensive final exam. The other portion of the final exam in the Science Majors course was an American Chemical Society conceptual exam for the first semester of general chemistry.

The laboratory requirements, content coverage, and final exam structure were not the only differences between the two courses. The other major difference is how the students in the Engineering and Science Majors courses were motivated to relearn material they missed on exams throughout the semester. In the Science Majors course, an optional replacement exam was available between the fourth exam and the final that could replace the student's lowest hour exam score. The students had to indicate whether they wanted to take the replacement exam and which exam they would be replacing. In cases of tied low scores, the student could choose the exam to take. In the Engineering course, students are able to earn points back on any exam throughout the semester. The comprehensive final exam was broken down into five parts, one each covering content from the hour exams and one for material covered since the fourth exam. If a student earns a higher score on the final exam portion corresponding to a particular exam than they did on the exam, they earn a percentage of those points back. This idea is referred to as resurrection points, and allows students to earn credit for relearning material from the whole semester, instead of the topics on a single exam (Herschbach, "Resurrection" Points, 1997). A summary of the similarities and differences between the Engineering and Science Majors courses is given in Tables 2 and 3.

Table 2 Similarities between Engineering and Science Majors Classes

Large Lecture Courses – Multiple Sections	Comprehensive Final Exam
Experienced Professors	Clicker quizzes in Class
Recitations – Run by Graduate TA's	Online homework
Four one-hour Exams	Quizzes in Recitation

Table 3 Main Differences between Engineering and Science Majors Classes

Differences	Engineering Course	Science Majors Course
Content Coverage	Survey	First of two semester sequence
Motivational tools	Resurrection Points	Replacement Exam
Laboratory	Optional depending on Major	Required
Final Exam Structure	Instructor Written	ACS Conceptual Exam and Instructor Written Algorithmic Section

Lecture and Exam Content Coverage

Motivated by the differences described above between the one-semester survey course and the more in-depth two-semester course, a comparison of the content covered in the two courses was undertaken to control for the amount of time spent on task in each course. A comparison of the amount of time spent in lectures and hour exams on quantitative and conceptual tasks, as well as the types of questions asked on the hour exams is presented below. In addition, an analysis of the percentage of time spent in lecture on ten major concepts in general chemistry will be presented.

The data used for analysis was obtained from the video and audio recordings made of each lecture in the Engineering and Science Majors courses and made available online to the students of these courses. Only one lecture section was recorded in each course and must be assumed to be representative of the other sections of the course on a particular day. The first level of analysis used on these recordings was to note the amount of time spent in each course on quantitative and conceptual topics. A quantitative topic was defined as balancing an equation, using algorithms and/or quantitative problem solving, for example solving a gram-to-gram conversion question. Clicker questions that met these criteria were also counted in the time spent on a quantitative topic. All other topics were classified as conceptual, for example explaining electron configurations. A list of topics covered in each course, based on the subheadings of chapters is presented in Tables 3 and 4, and will be discussed later. The time spent on each topic was summed over the whole semester and converted to a percentage. The percentage of time on quantitative and conceptual topics is presented in Figure 6.

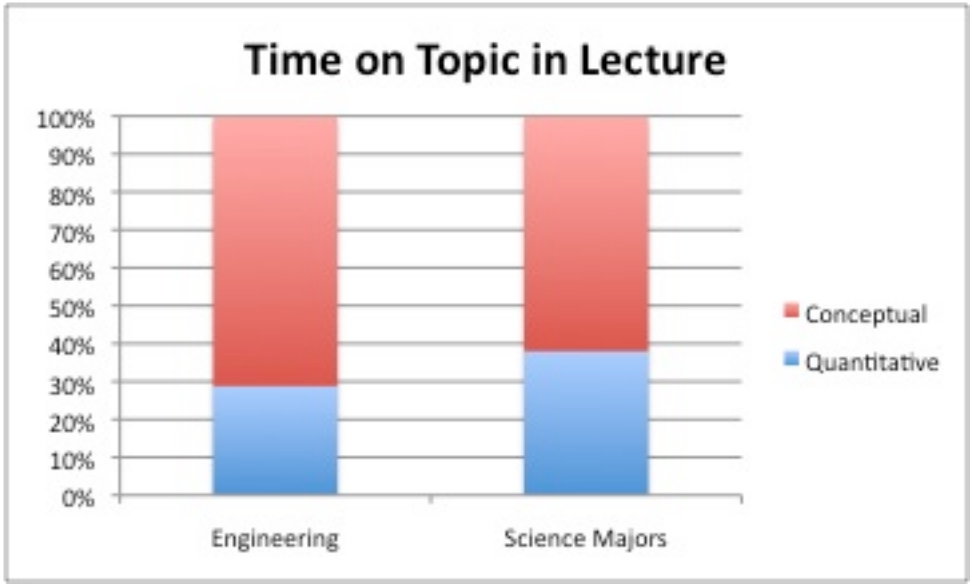


Figure 6 Time on Topic in Lecture by Course

A comparison between the Science Majors course and the Engineering course shows that more time was spent on quantitative topics in the former. This may be due to the fact that the Engineering students tend to have stronger math backgrounds than the Science Majors students. Most of the Engineering students have had calculus or were currently enrolled in calculus in the fall semester. Therefore, the Engineering professor could rely more on the students to be able to do the calculations outside of class, and could focus on presenting more conceptual material during lecture. In addition there was more material to be covered in the Engineering course, as the survey course covers the material from two semesters of general chemistry. The Science Majors professor spent more time working through the problems presented in class in a step-wise manner, to increase the likelihood that students would be able to carry out these calculations correctly. There are also fewer content areas to cover in the Science Majors course, though they may be covered in more depth when compared to the Engineering course.

Since the exam averages are going to be used as a data point for comparing the Engineering and Science Majors classes in the next chapter, an in-depth description of the exam conditions and content is in order. Both classes take four one-hour night exams over the course of the semester. These exams are taken in large lecture halls across campus with graduate students as proctors. The exams contained between 23 and 25 items in the Science Majors course and 25 items for the Engineering course. The exams had multiple choice parts, as well as long answer questions in both courses. These long answer questions were usually algorithmic or quantitative questions. There were some short answer questions on the Science Majors exams. There were on average three chapters of material covered on each exam for the Engineering and Science Majors courses. The topics covered on each exam are shown in Tables 4 and 5.

Table 4 Engineering Course Content by Exam

Exam One Coverage	Exam Two Coverage	Exam Three Coverage	Exam Four Coverage	Other Coverage
The Study of Chemistry	Limiting Reactants	The Ionic Bond	Spontaneity	Oxidation-Reduction Reactions and Galvanic Cells
The Science of Chemistry: Observations and Models	Theoretical and Percentage Yields	The Covalent Bond	Entropy	Cell Potentials
Numbers and Measurements in Chemistry	Solution Stoichiometry	Electronegativity and Bond Polarity	The Second Law of Thermodynamics	Cell Potentials and Equilibrium
Problem Solving in Chemistry and Engineering	Pressure	Keeping Track of Bonding: Lewis Structures	The Third Law of Thermodynamics	Batteries
Atomic Structure and Mass	History and Application of the Gas Law	Orbital Overlap and Chemical Bonding	Gibbs Free Energy	Electrolysis

Table 4 Engineering Course Content by Exam

Exam One Coverage	Exam Two Coverage	Exam Three Coverage	Exam Four Coverage	Other Coverage
Ions	Partial Pressure	Hybrid Orbitals	Free Energy and Chemical Reactions	Electrolysis and Stoichiometry
Compounds and Chemical Bonds	Stoichiometry of Reactions Involving Gases	Shapes of Molecules	Rates of Chemical Reactions	
The Periodic Table	Kinetic Molecular Theory and Ideal Versus Real Gases	Condensed Phases – Solids	Rate Laws and the Concentration Dependence of Rates	
Inorganic and Organic Chemistry	The Electromagnetic Spectrum	Bonding in Solids: Metals, Insulators, and Semi Conductors	Integrated Rate Laws	
Chemical Nomenclature	Atomic Spectra	Intermolecular Forces	Temperature and Kinetics	
Chemical Formulas and Equations	The Quantum Mechanical Model of the Atom	Condensed Phases – Liquids	Reaction Mechanisms	
Aqueous Solutions and Net Ionic Equations	The Pauli Exclusion Principle and Electron Configurations	Polymers	Catalysis	
Interpreting Equations and the Mole	The Periodic Table and Electron Configuration	Defining Energy	Chemical Equilibrium	
Calculations Using Moles and Molar Masses	Periodic Trends in Atomic Properties	Energy Transformation and Conservation of Energy	Equilibrium Constants	
Fundamentals of Stoichiometry		Heat Capacity and Calorimetry	Equilibrium Concentrations	
		Enthalpy	LeChatelier's Principle	
		Hess's Law and Heats of Reaction	Solubility Equilibria	

Table 4 Engineering Course Content by Exam ctd

Exam One Coverage	Exam Two Coverage	Exam Three Coverage	Exam Four Coverage	Other Coverage
		Energy and Stoichiometry	Acids and Bases	
			Free Energy and Chemical Equilibrium	

Table 5 Science Majors Course Content by Exam

Exam One Coverage	Exam Two Coverage	Exam Three Coverage	Exam Four Coverage	Other Coverage
The Study of Chemistry	Empirical Formulas from Analyses	The First Law of Thermodynamics	Electron Affinities	Characteristics of Gases
Classifications of Matter	Quantitative Information from Balanced Equations	Enthalpy	Metals, Nonmetals, and Metalloids	Pressure
Properties of Matter	Limiting Reactants	Enthalpies of Reaction	Trends for Group 1A and Group 2A Metals	The Gas Laws
Units of Measurement	General Properties of Aqueous Solutions	Calorimetry	Trends for Selected Nonmetals	The Ideal-Gas Equation
Uncertainty of Measurement	Precipitation Reactions	Hess's Law	Lewis Symbols and the Octet Rule	Further Applications of the Ideal-Gas Equation
Dimensional Analysis	Acids, Bases, and Neutralization Reactions	Enthalpies of Formation	Ionic Bonding	Gas Mixtures and Partial Pressures
The Atomic Theory of Matter	Oxidation-Reduction Reactions	Foods and Fuels	Covalent Bonding	The Kinetic Molecular Theory of Gases
The Discovery of Atomic Structure	Concentrations of Solutions	The Wave Nature of Light	Bond Polarity and Electronegativity	Molecular Effusion and Diffusion
The Modern View of Atomic Structure	Solution Stoichiometry and Chemical Analysis	Quantized Energy and Photons	Drawing Lewis Structures	Real Gases: Deviations from Ideal Behavior

Table 5 Science Majors Course Content by Exam ctd

Exam One Coverage	Exam Two Coverage	Exam Three Coverage	Exam Four Coverage	Other Coverage
Atomic Weights	The Nature of Energy	Line Spectra and the Bohr Model	Resonance Structures	A Molecular Comparison of Gases, Liquids, and Solids
The Periodic Table		The Wave Behavior of Matter	Exceptions to the Octet Rule	Intermolecular Forces
Molecules and Molecular Compounds		Quantum Mechanics and Atomic Orbitals	Strengths of Covalent Bonds	Select Properties of Liquids
Ions and Ionic Compounds		Representations of Orbitals	Molecular Shapes	Phase Changes
Naming Inorganic Compounds		Many-Electron Atoms	The VSEPR Model	Vapor Pressure
Some Simple Organic Compounds		Electron Configurations	Molecular Shape and Molecular Polarity	Phase Diagrams
Chemical Equations		Electron Configurations and the Periodic Table	Covalent Bonding and Orbital Overlap	Liquid Crystals
Some Simple Patterns of ChemicalReactivity		The Development of the Periodic Table	Hybrid Orbitals	The Solution Process
Formula Weights		Effective Nuclear Charge	Multiple Bonds	Saturated Solutions and Solubility
Avogadro’s Number and the Mole		Sizes of Atoms and Ions	Molecular Orbitals	Factors Affecting Solubility
		Ionization Energy	Period 2 Diatomic Molecules	Expressing Solution Concentration
				Colligative Properties
				Colloids

The content coverage for the first exam was similar across the two courses and mainly covered material students would have seen in high school chemistry. The ideas of measurements, atomic structure, ions, nomenclature, formulas for compounds, chemical equations and molar mass, as well as the fundamentals of stoichiometry including dimensional analysis were included on the first exam in both courses. There were some differences in content on the two exams, however. Prior to the first exam in the Science Majors course, students had covered classification and properties of matter and the some basic chemical reaction patterns. These topics were again topics that were likely covered in high school chemistry. Problem solving and the importance of observations and models were discussed in the Engineering course at the outset. In addition to these topics, the Engineering students had also talked about aqueous solutions and net ionic equations before the first exam. So while the content for the first exam was similar to high school chemistry content in both courses, there were differences between the two courses. These differences continued to increase throughout the semester as the Engineering students move into topics covered later in the first semester sequence and into the second semester general chemistry sequence.

The topics covered on the second exam in the Science Majors course included empirical formula calculations, general stoichiometry and solution stoichiometry, properties of solutions, reaction types, and information about the nature of energy. Many of these topics had already been covered in the Engineering course, so the content coverage for their second exam was different. For their second exam the Engineering students were being tested on the rest of stoichiometry not covered on the previous exam, including solution stoichiometry, properties of gases and gas laws mainly focusing on the ideal gas law, energy and electron configurations and finally, periodic trends in atomic properties. This wide-ranging exam content contrasts with the relatively closely related material on the second exam in the Science Majors course. The Science Majors focused on different types of reactions, and so stoichiometry, including balanced equations, can be shown throughout the different types of reactions. In contrast, while there is some discussion of stoichiometry on the Engineering second exam, in terms of solutions and gases in equations, other areas of focus include ideal gas laws and energy and its' relationship to electronic configurations and periodic trends in properties.

On the third one-hour exam in the Engineering course the topics include Lewis Structures, bonding and orbital overlap and shapes of molecules. In addition, topics about the properties and bonding in solids and liquids, including intermolecular forces, and finally energy changes, enthalpy and calorimetry. The Science Majors students are also being tested on enthalpy and calorimetry on the third exam. However, their exam also includes material from the second Engineering exam, namely electron configurations, atomic orbitals and periodic trends. The wave-particle duality of light is also a topic addressed in the Science Majors course before the third exam. At this point in the semester, the Engineering course has covered all the material that would normally be covered in a first semester of two-semester sequence. The next two exams will cover content from the second semester of general chemistry. Because the Science Majors course does not have to cover the entire general chemistry two-semester curriculum, the professors can slow down the coverage slightly and discuss topics in more depth than in the Engineering course. These differences in course content coverage by exam may lead to differing levels of challenge presented by a particular exam between the two courses. This difference in the level of difficulty will be addressed in the next chapter by using a method of analysis that looks at changes relative to the course mean to look at changes in student exam performance throughout the semester. The details of this comparison will be discussed in the next chapter.

The differences in the amount of topics to be covered is most apparent when looking at the content covered by the fourth hour exam in both courses. In the Engineering course the students are being tested on entropy and Gibbs free energy, kinetics and equilibrium topics. These are all topics that get little coverage in a high school chemistry course. In contrast, the Science Majors students are being tested on periodic trends that were not covered on the previous exam, bonding, Lewis structures and molecular shapes. The contrast between the topics covered in the two courses is striking. It is possible that the fourth exam is the most challenging one so far for the Engineering students. The timing of the exams is also important. The Science Majors fourth exam comes three weeks after the third exam and is before the weeklong fall break. The Engineering fourth exam occurs over a month after the third exam and it occurs after the weeklong fall break. The influences of the challenging material and extra long break between exams on the Engineering students performance will be discussed in the next chapter.

With the content coverage for the two-semester sequence completed by the fourth exam, the content covered only on the final exam in the Engineering course included electrolysis and batteries. The Science Majors class had about three weeks worth of content that was covered on the final exam only. These topics included gas laws, intermolecular forces and properties of the three states of matter and factors affecting solubility. While these topics are covered at the end of the semester that does not mean that they are not important. These topics may be a jumping off point for the second semester general chemistry course. As this is the only general chemistry course the Engineering students take, it makes sense to include electrochemistry in the curriculum, as it may be important to their future studies. With the topics on each exam compared between the two courses, it is now time to discuss comparing the types of questions being asked on each exam.

Having analyzed the amount of time spent on quantitative and conceptual topics in the two lectures, as well as discussing the different topics covered on the exams, one wonders what types of questions are being asked on these exams. To answer this question, the four one-hour exams in the Engineering and Science Majors classes were analyzed to determine the proportion of recall, algorithmic or problem solving and conceptual questions were on each exam. The questions were categorized based on the Raker and Towns (2010) article. In that article, recall questions were defined as requiring facts as answers without any procedure needed to call up the answer; problem solving/algorithmic questions as requiring a stepwise procedure to answer, and conceptual questions were defined as being answered by applying information to a novel chemical context (Raker & Towns, 2010). Answering a conceptual question might require applying multiple ideas, explaining how something happens, or predicting what would happen next. (Raker & Towns, 2010). Three raters were used to ensure inter-rater reliability of the categorizations and the inter-rater reliability was 0.821, which is a reasonable value for inter-rater reliability. The number of items in each category was calculated as a percentage of total items on each one-hour exam, which ranged from 23 to 25 items. The results of those analyses are presented in Figures 7 and 8.

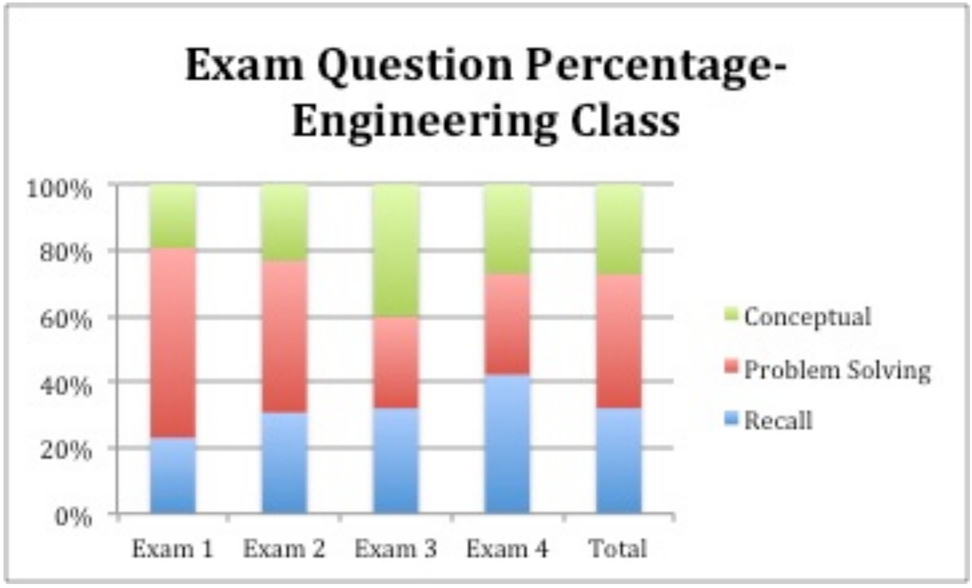


Figure 7 Exam Question Percentages by Type - Engineering Class

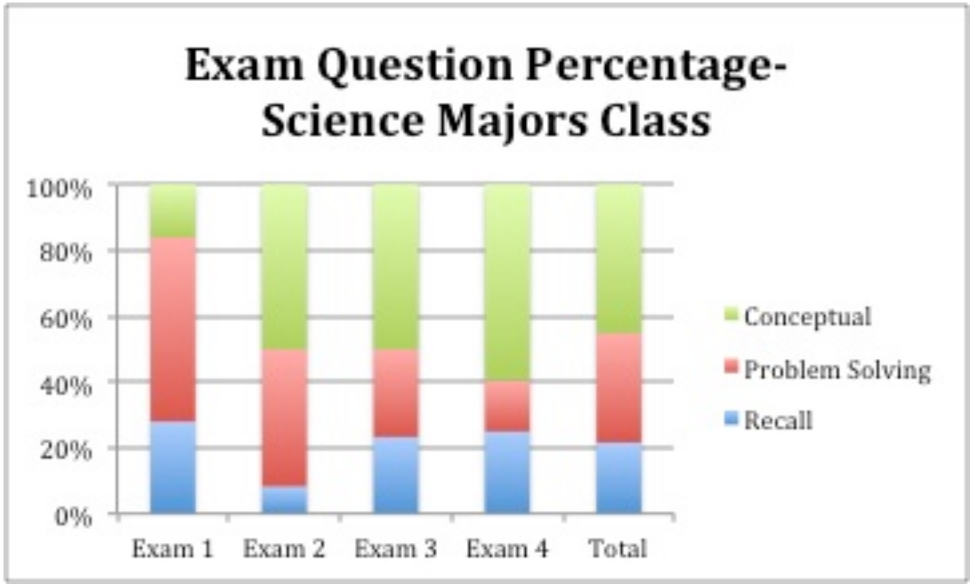


Figure 8 Exam Question Percentages by Type - Science Majors Class

While different content lends itself to more problem solving type questions than others, for example stoichiometry, others like bonding and intermolecular forces are more easily assessed using conceptual questions. While the percentage of questions in each category fluctuates from exam to exam, particularly for the Science Majors course, based on the overall trend, the Engineering course tends to ask more problem solving questions while the Science Majors course tends to ask more conceptual questions, as seen in the total columns in Figures 7 and 8. The Engineering professor is trying to prepare his students for future courses and possibility a career that is based on the students' ability to solve problems. To encourage students to develop these problem-solving skills, the Engineering professor asks problem-solving questions on his exams as well as on homework assignments. In the Science Majors course the focus is more on learning the chemical concepts, as they will be applied in the students' future course work and careers. Of course, the chemistry content is important for the Engineering students to learn, and problem-solving skills are important for the Science Majors students to learn as well. But overall the focus of the Engineering course is on teaching students problem solving skills in addition to chemistry content and the Science Majors course focusing on teaching students concepts.

When comparing the time spent in lecture on conceptual and quantitative topics to the percentage of conceptual and problem-solving questions on the exams throughout the semester, an interesting finding appears. During the lecture, the Engineering professor spends more time covering concepts, but on the exams there are more problem solving questions than conceptual questions. During the Science Majors lectures, the professors' focus on quantitative topics, but the focus of the exams is on conceptual questions. These opposite foci for the lectures and exams may be due to the amount of material being covered in each course as well as the mathematical backgrounds of the two populations of students. The Engineering course has a focus on problem solving skill development, which can be seen in the assessments used in the course, the exams in this case. The focus of the assessments in the Science Majors course is on ensuring that the students have learned the chemical concepts.

The final comparison of the content coverage in the Engineering and Science Majors courses will involve aligning the content coverage in the two courses with 10 Big Ideas or main concepts covered over the course of the undergraduate chemistry curriculum. The process for developing a content map for the undergraduate chemistry curriculum began in response to colleges moving to outcomes-based assessments that require professors within each discipline to determine the outcomes to be assessed (Murphy, Holme, Zelinsky, Caruthers, & Knaus, 2012). As the American Chemical Society Exams Institute is an independent entity that had experience developing exams that could be used to assess such outcomes, the Exams Institute took on the challenge of working with chemistry professors to develop the content map for undergraduate chemistry (Murphy, Holme, Zelinsky, Caruthers, & Knaus, 2012). Starting in March of 2008, professors got together at national American Chemical Society to meeting to decide on the four levels of the content map and align ACS exams to the map (Murphy, Holme, Zelinsky, Caruthers, & Knaus, 2012).

The bottom two levels of the map, Big Ideas and Enduring Understandings are generalizable to all areas of chemistry, while the top two levels, Sub-Disciplinary Articulations and Content-Level Details, are specific to the sub-disciplines of chemistry. The Big Ideas are the 10 main topics covered during all four years of chemistry coursework. The Enduring Understandings are the seven to 10 ideas about each Big Idea that professors want students to remember at the end of a chemistry course. The Sub-Disciplinary Articulations describe how each sub-discipline of chemistry talks about the Enduring Understandings. The Content-Level Details the most fine-grained level of the map.

The content coverage from the Engineering and Science Majors courses was aligned with the 10 Big Ideas of the content map for undergraduate chemistry, as a way to discuss differences in content coverage between the two courses, over the whole semester. The time on content in lecture converted to a percentage and normalized to 100% for both courses. The results of the analysis are shown in Figure 9.

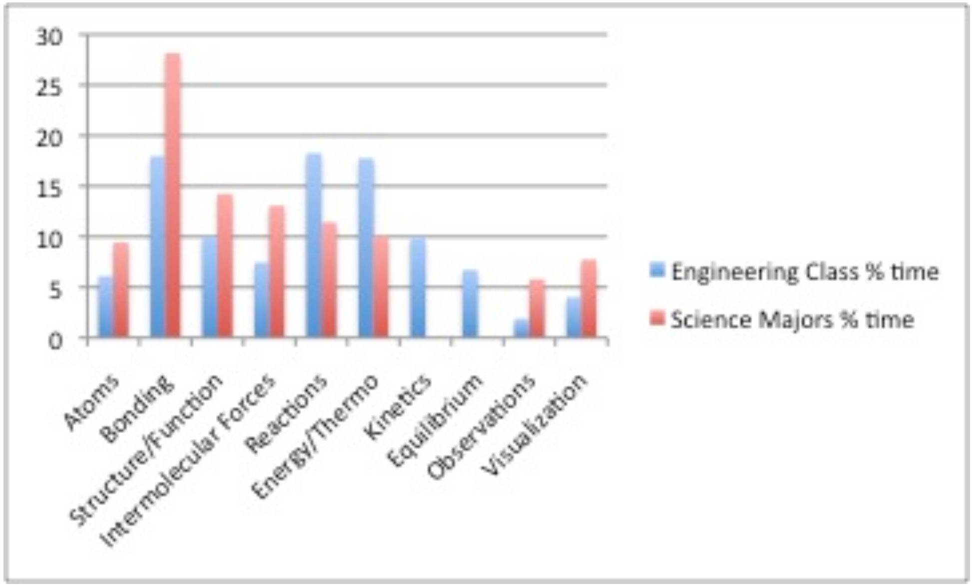


Figure 9 Percentage of Lecture Time Spent on each Big Idea (Normalized to 100%)

Some of the topics covered in the lecture overlapped with more than one Big Idea in the content map (20% in the Engineering course and 38% in the Science Majors course). These overlapping content was covered in each Big Idea represented. Less than 30% of the time was spent on any given topic in both courses. As is expected in a survey course, the Engineering course spends time on each of the 10 Big Ideas throughout the semester. The Science Majors course does not cover kinetics or equilibrium, as these are second semester topics. Besides kinetics and equilibrium, the Engineering course also focuses more time on reactions and energy and thermodynamics than the Science Majors course. This may be due to the fact that those topics are of particular concern to engineers, as they are important to the design, fabrication and breakdown of the materials they use to build. The Science Majors course spends more time than the Engineering course on atoms, bonding, structure and function relationships, intermolecular forces, observations and visualizations. Some of this extra time is a result of being able to go more in-depth on the topics covered in the course, because there are less of them to cover as compared to the Engineering survey course. Some of the extra time is a choice made by the professors in the course. Many of the Science Majors students go on to study organic chemistry where bonding and structure-function relationships are important. With this in mind, the professors focus more on these topics in the first semester general chemistry course to ensure that students have seen the material before they get to organic chemistry. Overall, there are differences in the alignment between the content coverage in the two courses and the Big Ideas level of the ACS content map for undergraduate chemistry. These differences reflect not only time constraints in the two courses, but also conscious decisions on the part of the professors in the courses. This in-depth analysis of the content coverage in the two courses will set up the analysis of the students' exam performance, attitudes about chemistry, as well as the analysis of their problem solving skills described in chapters four and five.

Chapter four details the comparison of students' exam performance in the two courses, based on exam averages and changes in exam performance as measured by delta z-scores. Within the exam performance analysis, particular attention is paid to the influence of the motivational tools in each course (resurrection points in the Engineering course, and a replacement exam in the Science Majors course) on the students' final exam performance. Changes in the students' attitude about chemistry were also assessed using the Attitude about the Subject of Chemistry Inventory version 2 (Xu & Lewis, Refinement of a Chemistry Attitude Measure for College Students, 2011). The influence of the laboratory in the

Engineering course as well as the level of TA experience on students' attitude about chemistry and exam performance will also be addressed in this chapter.

The analysis in chapter five focuses on interviews conducted with students from the two courses. The students were asked to talk-aloud while working on a total of six questions related to stoichiometry and thermochemistry. The analysis will focus on two of the stoichiometry questions and the problem solving strategies the students used to solve the questions. The strategies students' use will be compared with how stoichiometry tasks were solved in the lecture in the Engineering and Science Majors courses. The course descriptions and content coverage analysis presented in this chapter laid the groundwork for comparing the exam performance and problem solving strategies used by students as well as changes in their attitude about chemistry over the next two chapters.

CHAPTER 4: COURSE STRUCTURE EFFECTS ON EXAM PERFORMANCE AND ATTITUDE

Introduction

Now that we have a more detailed understanding of the content covered in the lectures and exams of both the classes we can begin to study how these differences influence students' exam performance and attitudes about chemistry. However, in order to be able to say that the effects are due to the course structure and not just the students' prior knowledge, we must establish equivalent groups of students before we begin the analysis. It would be best to be able to define this equivalent group based on an exam all the students took early on in the semester, such that all differences afterward could be attributed to the class itself (as much as that is possible).

Establishing Equivalence

In order to better understand the effects of the course format on student performance, it was important to start with a group that is equivalent across both courses. It was also helpful for this equivalent group to be based on a measure made early on in the semester, so that the effects for early assessments as well as later assessments could be addressed. With this in mind, the two variables that presented themselves as reasonable to use for equivalency were the Toledo Test scores and placement test scores. Both tests occurred prior to significant course instruction and both test for chemistry content knowledge of students at the outset of the course. The placement test scores for both courses are shown in Table 6. One possible advantage to the placement test score was its assessment of mathematical skills.

However, mathematical skills are not historically an issue for most students in the Engineering Class.

Within any equivalent group formed, the goal is to find students with means on the variable of interest that were as similar as possible. So much so in fact, that the means of the two groups were statistically the same. The equivalent group or groups formed will be selected from the students in each course who gave informed consent, received a grade in the course and complete both the departmental placement exam and the Toledo Tests given at the beginning of the semester. The equivalent groups will be determined based on either the placement test scores or the Toledo test scores. The process will be described below.

Table 6 Placement test scores and Toledo test scores for Engineering and Science Majors Classes

	Placement test scores (standard deviation)	Toledo test scores (standard deviation)
Engineering Class (N = 476)	23.73 (4.65)	23.70 (4.66)
Science Majors Class (N = 401)	35.00 (8.39)	24.71 (4.64)

Traditional hypothesis testing with t-tests suggested that the next step would be to break the two classes into groups somehow, starting perhaps at the largest group (i.e. all the students in each class (476 for Engineering Class and 401 for Science Majors Class) and comparing the means of each group with the null hypothesis being that they are the same. However, prior work by Lewis and Lewis (2005) suggests that traditional hypothesis testing, as a method of developing equivalent groups was insufficient for two reasons. First, the rejection of an alternative hypothesis is not the same statistically speaking as accepting the null hypothesis that the two means are the same (Type I error, measured by α). Second, it is possible to fail to reject the null hypothesis when in fact it is incorrect (Type II error, measured by β). While the Type I error can be controlled by setting α equal to some pre-determined level, it is not possible to control Type II error by setting β to a pre-determined value. The β value is very important when trying to determine if two groups are equivalent. It is imperative that two groups are accurately defined as equivalent or not, in a definite sense rather than a probabilistic sense.

With this in mind, Lewis and Lewis (2005), expanded on the prior work of others in the field of medicine, in putting forth a method for determining equivalence between groups in an educational setting. They suggest instead of using one hypothesis test to determine equivalence, that two tests be used. The concept being that if that the value for the difference in means between two groups was not in either region 1 or region 2 (see Figure 10 below), i. e. outside of the equivalence interval, then the only other place for the difference to be was within the equivalence interval, if region 1 and region 2 extend to infinity in either direction. In the analysis for this study, the Toledo Test scores and the departmental placement test scores will be used to attempt to determine equivalence.

H'_o θ_1 θ_2 H''_o

<------------------------|--|----------------->

region 1 equivalence interval region 2

Figure 10 Diagram of Equivalencing Testing

These two hypothesis tests are defined such that the difference in the means between the two groups you are testing are being tested as less than H'_o (i. e. less than θ_1) and as greater than H''_o(i. e. greater than θ_2).(Equation 3.1, 3.2).

$$H'_o : \mu_1 - \mu_2 < \theta_1 \qquad 3.1$$

$$H''_o : \mu_1 - \mu_2 > \theta_2 \qquad 3.2$$

The thetas define the two ends of the equivalence interval. In order to make the interval more like a traditional t-test around zero, the two theta values can be chosen to be the same value but with opposite signs. Therefore theta 1 is defined as the opposite of theta 2 as shown in Equation 3.3. Theta 2 is defined as the difference in means as determined from the Cohen's d calculation (Equation 3.4).

$$\theta_1 = - \theta_2 \qquad 3.3$$

$$\theta_2 = (\mu_1 - \mu_2) \qquad 3.4$$

Cohen's d normally allows you to calculate the effect size of your statistical results. However, the authors of the paper have suggested that Cohen's d (Equation 3.5) can also be used to calculate the values of the equivalence areas.

$$d = (\mu_1 - \mu_2) / \sigma \qquad 3.5$$

Cohen's d values greater than 0.2 are considered significant. Therefore a Cohen's d value of 0.2 and below is a measure of the noise in a measurement. If we set d equal to 0.2 and calculate sigma (Equation 3.6) by taking the square root of average of the variances, (sigma squared), we can solve for the difference in means that will be used as the endpoints of the equivalence interval.

$$\sigma = \text{square-root } \{(\sigma_1)^2 + (\sigma_2)^2 / 2\} \qquad 3.6$$

With these theta values defined, the new t-tests can be written. The t-tests are similar to the regular t-tests but instead of subtracting the observed difference in means from the population difference in means, the observed difference in means is subtracted from the population difference in means from the Cohen's d calculation. Each t-test determines if the difference is outside the equivalence interval, one of each side of the interval. The t-tests are show in Equations 3.7 and 3.8 and the pooled standard deviation equation is shown in Equation 3.9, but in practice this value comes from the square root of the mean square within groups variance from a one-way ANOVA calculation.

$$t_1 = [\{(\overline{X}_1 - \overline{X}_2) - \theta_1\} / s_p]^* (\text{square-root } \{1/N_1 + 1/N_2\}) \geq t(1-\alpha, (N_{total} - 2)) \tag{3.7}$$

$$t_1 = [\{ \theta_2 - (\overline{X}_1 - \overline{X}_2)\} / s_p]^* (\text{square-root } \{1/N_1 + 1/N_2\}) \geq t(1-\alpha, (N_{total} - 2)) \tag{3.8}$$

$$\begin{gathered} s_p = \text{square-root } \{(SS_1 + SS_2) / (N_1 + N_2 - 2)\} \\ \text{or } s_p = \text{square-root (Mean Square within groups)} \end{gathered} \tag{3.9}$$

The final consideration is how to determine if the groups are equivalent based on the results of the two t-tests. Since regions 1 and 2 in Figure 1 run to infinity in either direction, if the value is not in either of those regions, the value must be in the equivalence interval. Therefore a significant p-value (< 0.05) for both t-tests would indicate that the two groups you are testing using a given variable are equivalent. Both t-tests must give significant p-values in order to ensure the two groups are equivalent. The online placement test score was used first as the variable of interest for equivalencing, but any groups remaining in the equivalent group did not have sufficiently large N values for further analysis. Therefore the next option was to use the Toledo Test scores as the variable of interest for building equivalent groups.

Dividing the raw Toledo Test scores into various size groups, including thirds, quartiles, and deciles, as well as dividing Science Majors Class into engineering and non-engineering majors to compare with Engineering Class did not lead to the identification of equivalent groups across the two courses with sufficiently large N values for further analysis. Finally, z-scores were used as a mathematical way to adjust these data for use in equivalence testing.

Z-scores can be determined for any set of values, but the interpretation is based on the shape of the curve. The interpretations that follow are based on a bell curve. The curve can be broken into standard deviations, for example from -4 standard deviations on the left to +4 standard deviations on the right around a given mean value of a variable for a population. One would then be able to talk about performances that were within a given number of standard deviations of the mean. Z-scores are special cases of the standard deviations around a population mean. Instead of setting the middle of the normal distribution to the population mean, the center can be set to zero, such that the number of standard deviations away from the mean is the z-score of that variable (Rose & Sullivan, 1996). In fact, z-scores can be used to compare individual scores on different variables (Rose & Sullivan, 1996). Quantitatively, taking an individual score and subtracting the mean and dividing by the standard deviation leads to a z-score. Z-scores of the Toledo Test scores for the 877 students were converted in this fashion using the appropriate mean and standard deviation depending on the class the student attended. A more detailed description of how and why z-scores were chosen as the metric for determining equivalent groups is given below.

Toledo Test scores for the 877 students were converted to z-scores using Stata version 12.0 using the method described above. A group (tsgrp) was designated that would include the z-scores of the Toledo test scores as long as they were larger than or equal to a pre-set negative z-score (z_neg) and smaller than a pre-set positive z-score (z_pos), (i. e. if your z-scores were in the range, you were assigned tsgrp==1 else, you were assigned (tsgrp==0). These pre-determined z-scores could then be varied from either positive or negative three to zero by 0.1 intervals until an equivalent group had been found.

Table 7 Toledo Scores of Engineering Class versus Science Majors Class before equivalencing * = < 0.10, ** = <0.05, * = < 0.01**

	Toledo Score Mean (standard deviation)	t-test (p-value)	z-score Toledo score Mean (standard deviation)	t-test (p-value)
Engineering Class (N = 476)	23.70 (4.67)		-0.992 (1.00)	
Science Majors Class (N = 401)	24.71 (4.64)	-3.2201(0.0013)***	0.118 (0.99)	-3.2201(0.0013)***

Table 7 shows the Toledo test scores for the two classes of students, along with the t-test results. The students mean Toledo tests scores are statistically significantly different before the two groups have been equivalenced. Also, the z-scores of the Toledo Test scores are statistically significantly different before an equivalent group has been formed of students in both the Engineering and Science Majors Classes.

Since the sigma value for calculating theta 2 and theta 1 comes from a one-way ANOVA, three variables had to be assigned to determine equivalency. The first as discussed earlier was the grouping variable (tsgrp) as determined by z-scores of the students' Toledo Test score (variable name = toledoscore). The other two variables were the toledoscore itself and the course (Engineering Class or Science Majors Class with the caveat that the student had to be included in the first grouping variable (i. e. tsgrp==1). In addition to the theta values, t-test statistics and p-values, the values of z_neg and z_pos were displayed along with a table with the N values for each case of tsgrp (i. e. number of people in the equivalent group if the equivalence test came back with an appropriate p-value (tsgrp==1) and the number of people excluded, (tsgrp==0). A minimum number of people in the equivalent group were determined to be 100, as it would allow for reasonable future analysis. Several z_neg and z_pos combinations led to a sufficient number of people in the equivalent group, so the z_neg and z_pos values where chosen to maximize the number of people in the equivalent group. The final values for the z_neg and z_pos where -2.1 and 1.3 respectively and they lead to a total of 792 students in the equivalent group overall. Upon further inspection, seven students were found to have missed at least one exam and were therefore removed from the equivalent group by changing their tsgrp value from one to zero manually. Once these students were removed there was a total of 782 students in the equivalent group overall with 428 of them being from Engineering Class and 354 from Science Majors Class. All future analyses will focus on these 782 students. The Toledo Test Scores for the two courses, as well as the t-test values after equivalencing are shown in Table 8.

Table 8 Toledo Scores Engineering Class versus Science Majors Class after equivalencing * = < 0.10, ** = 0.05, * = < 0.01**

	Toledo Score Mean (standard deviation)	z-score Toledo score Mean (standard deviation)	t-test 1 (p-value)	t-test 2 (p-value)
Engineering Class (N = 428)	23.46 (3.70)	-0.149 (0.79)		
Science Majors Class (N = 354)	23.73 (3.64)	-0.093 (0.78)	1.849 (0.0324)**	3.753 (0.0001)***

The p-values of both t-tests after equivalencing are significant. Since the two t-tests are testing whether the values are outside of the equivalence region between the means of the two groups, two statistically significant values for the t-tests means that the z-scores of the Toledo Test scores for the two classes are within the equivalence region. If the Toledo Test scores are a measure of the students' prior knowledge of chemistry, then these students have the same distribution of prior knowledge of chemistry. The future differences in student performance are determined by the course in which the students were enrolled.

Overarching Theme

The overall question that is under investigation is what is the impact of a one-semester general chemistry course versus the first semester of a two-semester general chemistry course on student learning? More specific questions about the effects of the overall course structure on student exam performance and attitudes about chemistry, the effects of how professors reward the relearning of missed material and it's effect on final exam performance, as well as the level of TA experience teaching general chemistry and the effect of being enrolled in the laboratory in the Engineering course. These questions are all trying to understand the effect of the course structure on the students' exam performance and attitudes about chemistry. The answers to these questions will determine what changes to course structure will be suggested to improve student understanding of chemistry content.

Research Question

What is the impact of accounting for effort points in calculating students overall performance versus only counting their exam grades?

Hypothesis

Effort points are defined in this study as the homework points, and points on clicker quizzes. These clicker and homework points account for 20.2% of points available in the Engineering course and 18.8% of the available points in the Science Majors course. When effort points are accounted for, students' performance will be more positive than when overall performance is calculated based only on exam scores. With the effort points the students will get credit for spending more time with the material by doing homework and taking part in clicker questions. It is assumed that when taking effort points into account a student may end up in a different quartile than when their exam grades are used to place them in a quartile in the course.

Method

One research question that can be addressed by these data is there a difference in student performance when grouping students based on exam scores alone versus grouping based on the inclusion of effort points? A hypothesis was developed about differences between student exam performance when accounting for non-exam points (grouping based on course percent) versus accounting only for points earned on exams (grouping based on exams only). The exam points only account for 71% of the points in the Engineering course and 68.8% of the points in the Science Majors course. The hypothesis stated that student performance as measured by delta z-scores was expected to be more positive when grouping based on course percent versus when grouping based on exams only, because students would be "getting credit" for familiarizing themselves with content on exams in more low-stakes environments, i. e. on homework, clicker questions and quizzes. This hypothesis is predicated on the idea that students would be moving to different quartiles when grouping based on exams only as opposed to grouping by course percent. So the first step in testing this hypothesis is to see how many students do change quartiles when grouping based on exams only. A two-way table was produced to compare the quartiles students are in when grouped by course percent versus by exams only overall, show in Table 9 below.

Table 9 Movement between Quartiles using Effort Points versus Exam Scores only

Course Percent	Exams Only Quartile				
Quartile	Q1	Q2	Q3	Q4	Total
Q1	181	17	0	0	198
Q2	18	153	21	1	193
Q3	0	22	158	16	196
Q4	0	0	14	181	195
Total	199	192	193	198	782

There were a total of 109 students who changed quartiles, that's 13.9% of the students in the equivalent group. With the exception of the one student who moved from quartile two to quartile four when calculating based on exams only, most students did not move far from the quartile they were assigned based on course percent.

The next step is to test for collinearity between course percent and the exams only scores. The exam scores were totaled for the four one-hour exams and added to the final exam score. These values were then divided by the appropriate maximum value of points available (850 points for Engineering Class and 800 points for Science Majors Class) and multiplied by 100 to convert to a percentage. These exam percentages for each class were combined together into one exam percentage variable. The course percent variable underwent the same treatment Z-scores of the course percent variable and the exam percentage variable were calculated to be able to directly compare performance across courses. Some collinearity is expected as the exam scores from the four one-hour exams and the final exam make up a large portion of the points in each class (70.6% in Engineering Class and 68.8% in Science Majors Class). However, collinearity above these values indicates that the two variables are in fact measuring the same thing. A two way scatter plot was generated by course to assess visually whether there was a high level of collinearity, see Figure 11. A Pearson r statistic and linear coefficients and R^2 values were also calculated.

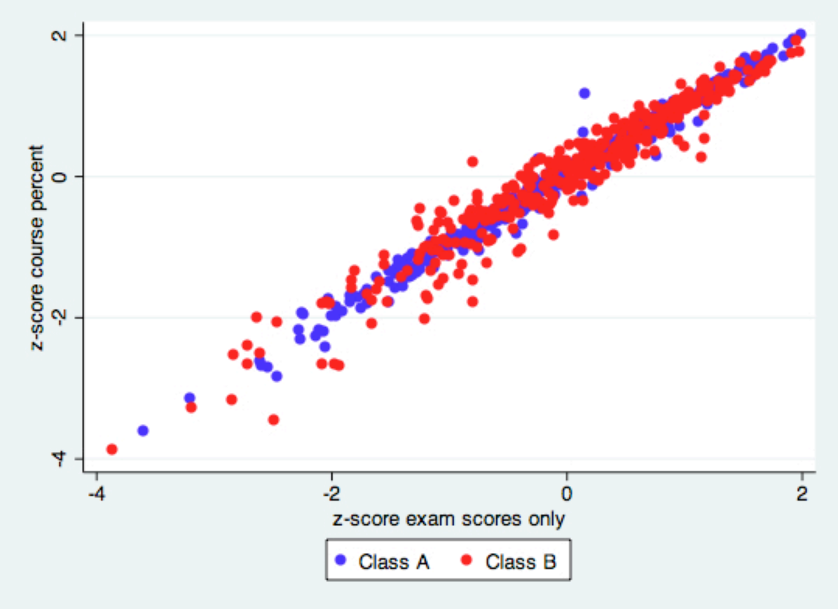


Figure 11 Effort Points versus Exam Scores Only z-scores graph

The Pearson correlation coefficient (r) value for this graph is 0.9793 with an p-value of less than or equal to 0.0000. The coefficient between these two variables is 0.9793 with a p-value of less than or equal to 0.0000. The intercept of this line is 5.22 x 10^-10 with p-value of 1.000. The line of the two variables is z-score of course performance = 0.9793 (z-score of exam only performance) + 0. We have removed the shift in the y-axis by taking the z-scores of the two variables. The R^2 value of this line is 0.9591. This means that 95.91% of the variability in the z-score of course performance is explained by the z-score of the exams only performance. If two variables have a coefficient that is 0.90 or above, statistically speaking the two variables are measuring the same thing. Since exams are a measure of student learning, then because of this collinearity, a measure of course percent and a measure of exam performance both measure student learning (Kline, 2011). Since exam scores have traditionally been used to as measures of student learning, exam score only data will be used in all future analysis.

Research Question

Are there differences in student movement relative to the mean in the two courses due to differences in course structure?

Hypothesis

The movement of students relative to the mean is not expected to be different in the two classes due to differences in course structure.

Sub Hypothesis

The differences between the exam averages and the class average on a given exam will be the same for each subgroup of students in the two classes.

Method

Students were grouped into Quartiles based on their overall exam performance. These groups were then used to calculate average scores as percentages on each of the four one-hour exams and the final exam in each class. These percentages are compared across the two courses in order to start to get an idea of the effect of the overall course structure on the student exam performance. The medians and means are presented because the data are not normally distributed and variance is not the same across the Engineering and Science Majors

Classes. Non-parametric versions of the individual samples t-test and ANOVA, the Wilcoxon rank-sum test and the Kruskal-Wallis equality-of-populations rank test, were used to analyze whether or not the difference in medians was statistically significantly different between the two classes at a particular level of Quartiles or Grade and if there was a statistically significant difference in the medians between levels of the grouping variables within each class. All chi-squared results in the Kruskal-Wallis equality-of-populations rank test will be presented with ties, which occurs when the score received the same rank as another one. This rarely makes a difference in the interpretation of the results. It did not make a difference in any of the analyses that follow.

These non-parametric tests are similar to parametric two group t-tests and ANOVAs in that they are looking for differences between groups and how likely it is that those differences are due to chance. Since multiple comparisons are being made with the Kruskal-Wallis tests and any subsequent Wilcoxon rank-sum tests look at differences between each pair of groups, the Bonferroni correction was used to adjust the alpha value to determine significance of the resulting p-values. The Bonferroni correction is used to prevent results that are not actually statistically significant from being declared so. The number of comparisons made divides the desired alpha value for the overall comparisons. Now, in order for a particular results to be declared statistically significant, the p-value must be less than this smaller alpha value. The Bonferroni corrected alpha values are shown above the relevant tables.

Exam Averages: Quartiles

The exam averages for both courses divided by quartiles based on overall exam scores are shown in Table 10 and Table 11. Graphs depicting the changes in exam averages by quartiles for the Engineering and Science Majors courses are shown in Figures 12 and 13.

Table 10 Exam Averages for the Engineering Class - Quartiles

Mean (Median) (N = 428)	Exam 1 (%)	Exam 2 (%)	Exam 3 (%)	Exam 4 (%)	Final Exam (%)
Top Quartile (N = 110)	90.1 (90.5)	83.0 (84.0)	86.3 (87.5)	78.2 (79.0)	90.3 (90.5)
Middle Top Quartile (N = 104)	83.5 (84.0)	75.2 (76.0)	77.4 (77.0)	65.6 (65.5)	81.6 (81.5)
Middle Bottom Quartile (N = 106)	76.8 (78.0)	66.0 (67.0)	70.0 (70.0)	55.3 (56)	75.6 (75.8)
Bottom Quartile (N = 108)	65.2 (67.0)	53.4 (53.0)	60.5 (60.0)	40.2 (42.0)	64.3 (65.0)
Class Average	78.9 (82.0)	69.4 (71.0)	73.6 (75.0)	59.9 (61.0)	78 (79.3)

Table 11 Exam Averages for the Science Majors - Quartiles

Mean (Median) (N = 354)	Exam 1 (%)	Exam 2 (%)	Exam 3 (%)	Exam 4 (%)	Final Exam (%)
Top Quartile (N = 89)	87.5 (89.0)	88.9 (90)	89.8 (90.0)	89.1 (89.0)	83.7 (84.7)
Middle Top Quartile (N = 88)	79.1 (79.0)	78.7 (79.0)	83.7 (84.0)	80.8 (82.0)	74.8 (75.3)
Middle Bottom Quartile (N = 87)	74.3 (75.0)	69.7 (71.0)	76.9 (78.0)	76.1 (77.0)	66.1 (66.0)
Bottom Quartile (N = 90)	62.8 (64.0)	51.7 (54.5)	67.9 (69.5)	64.8 (65.0)	54.9 (56.7)
Class Average	75.9 (76.5)	72.2 (75.5)	79.5 (81.0)	77.7 (80.0)	69.8 (70.7)

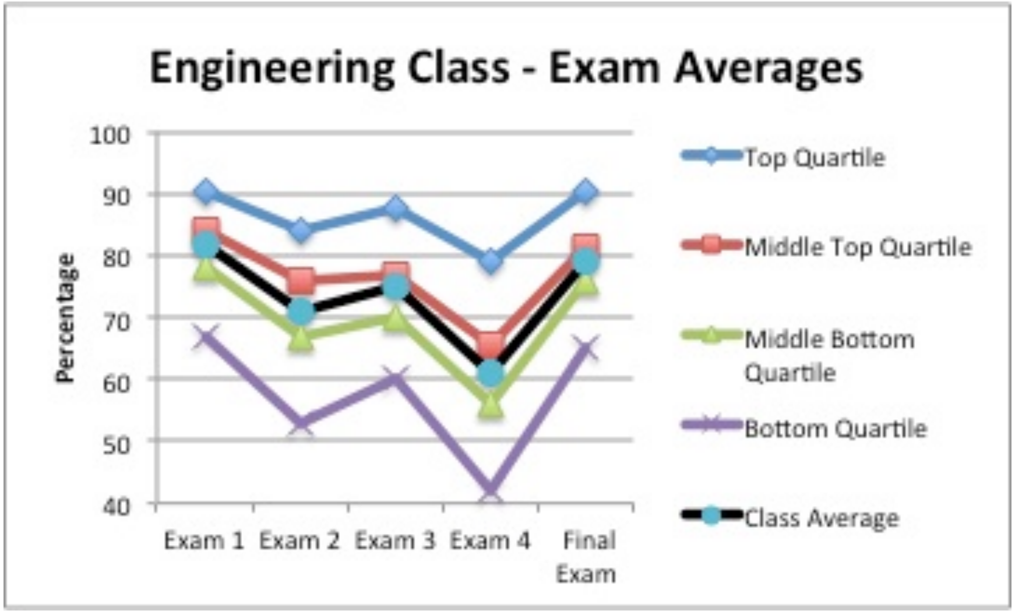


Figure 12 Exam Averages by Quartiles for Engineering Class

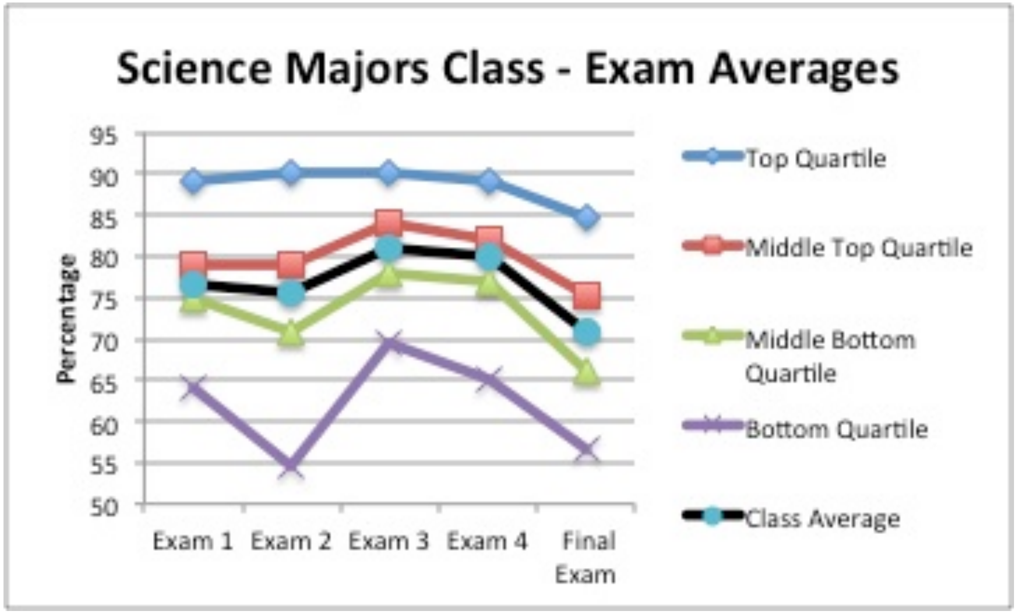


Figure 13 Exam Averages by Quartiles for Science Majors Class

As expected, in both the Science Majors and Engineering classes the Top and Middle Top Quartiles perform above the class average and the Middle Bottom and Bottom Quartiles perform below the class average. The Top and Middle Top quartiles of students in both classes perform fairly consistently across the semester, while there is more variability in the performance in the Middle Bottom and especially the Bottom quartiles of students in both classes. There is a large drop in the class average and student performance on the fourth exam in the Engineering class. This is most likely due to the fact that the exam happened after the weeklong fall break, and therefore students had not studied as much as they normally would have, or they may have forgotten what they studied over the break. This drop did not occur in the Science Majors class because their fourth exam occurred before the break.

The students' performance on each exam will be compared at each quartile level across the two courses, to look for differences between the two courses, using the Wilcoxon rank-sum test, in Table 12. After that, exam scores will be compared across quartiles within each class using the Kruskal-Wallis equality of medians test. Any significant results of that test will be followed up with Wilcoxon rank-sum tests to determine which pairs of quartiles are different. Differences between the two courses would be expected due to the differences in the exam averages for the two courses.

Table 12 Wilcoxon ranksum test results of analysis of median differences between Engineering and Science Majors Classes * = < 0.10, ** = < 0.05, * = < 0.01**

z-statistic (p-value) mean difference (E – SM)	Exam 1	Exam 2	Exam 3	Exam 4	Final Exam
Top Quartile	2.519 (0.0118)** 1.5	-5.421 (0.0000)*** -6	-3.939 (0.0001)*** -2.5	-8.548 (0.0000)*** -10	7.551 (0.0000)*** 5.8
Middle Top Quartile	4.517 (0.0000)*** 5	-3.176 (0.0015)*** -3	-6.085 (0.0000)*** -7	-9.768 (0.0000)*** -16.5	7.237 (0.0000)*** 6.2

Table 12 Wilcoxon ranksum test results of analysis of median differences between Engineering and Science Majors Classes * = < 0.10, ** = < 0.05, * = < 0.01 ctd**

z-statistic (p-value) mean difference (E – SM)	Exam 1	Exam 2	Exam 3	Exam 4	Final Exam
Middle Bottom Quartile	2.484 (0.0130)** 3	-2.671 (0.0076)*** 5	-5.623 (0.0000)*** -8	-10.664 (0.0000)*** -21	8.965 (0.0000)*** 9.8
Bottom Quartile	2.165 (0.0304)** 3	0.404 (0.6864) -1.5	-4.807 (0.0000)*** -9.5	-10.441 (0.0000)*** -23	6.585 (0.0000)*** 8.3

For the first exam of the semester, the Engineering students had higher exam averages, but for the next three exams, the Science Majors students had higher exam scores. However, when comparing the final exam scores, the Engineering students once again have higher scores than the Science Majors students. The quartiles in each course out perform the groups below them, as expected on all four one-hour exams and the final exam. These results are shown in the Appendix.

These statistics indicate that from the first exams students are on a trajectory that leads them to their quartile at the end of the semester based on their exam scores only. While some students may move out of their trajectory, students in each quartile perform consistently with their ultimate place at the end of the semester. It is possible that these exam average results are due to the level of difficulty of the exams in the two courses, instead of students' understanding of chemistry content. This question will be addressed using delta z-scores later in the chapter. The impact of laboratory on the Engineering students' exam performance will be discussed next.

Sub Hypothesis

Within the Engineering class, the students in the laboratory will have higher scores on the exams relative to the mean as compared to the students not in the laboratory.

Method

Before answering the question of the influence of laboratory on student exam performance, it is important to know if the students in the laboratory and non-laboratory groups were equivalent to begin with. The equivalencing discussed above was used to determine if the students in the two courses were equivalent. The same tests of the z-scores of their Toledo Test scores will be used to determine if the students in the Engineering class were equivalent when there were broken down into laboratory and non-laboratory students. There were a total of 38 sets of z-scores that lead to an equivalent group between the lab and no lab students containing between 398 and 448 students. The amount of overlap between the existing equivalent group (tsgrp) and the new equivalent group (lab_nolabgrp) was calculated using a two-way table. There were four sets of z-scores that included all the Engineering tsgrp students in the new lab_nolabgrp equivalent group; -2.1, 1.4; -2.0, 1.4; -2.1, 1.3; and -2.0, 1.3. The set of z-score that produced the widest interval was chosen as the set of z-scores of the Toledo Test Scores that would determine the equivalent group lab_nolabgrp, -2.1, 1.3. The next step was to determine if the results of the analysis that follows was the same using the two groups. The exam averages, show in Table 13 and Table 14,are essentially the same for the two groups, when comparing two different equivalencing groups. Further evidence of the similarities between the two equivalent groups is shown in Table 15, which shows the ranksum results between the Lab and No Lab groups for the exam averages using the two different equivalent groups. The conclusions one would draw from the data from the two equivalent groups is same, because both equivalencing groups leads to significant results in for the same exams. For this reason, the tsgrp equivalent group will be used while comparing the performance of the laboratory and non-laboratory students in the Engineering course from here on out.

Table 13 Exam Averages for Lab - No Lab group with tsgrp

Median	E1	E2	E3	E4	FE
No Lab (N = 298)	81.5	70.0	74.0	59.0	78.8
Lab (N = 130)	82.0	74.0	78.5	63.0	80.0
Average	82.0	71.0	75.0	61.0	79.3

Table 14 Exam Averages for Lab - No Lab group with lab_nolabgrp

Median	E1	E2	E3	E4	FE
No Lab (N = 305)	81.0	70.0	74.0	59.0	78.8
Lab (N = 130)	82.0	74.0	78.5	63.0	80.0
Average	82.0	71.0	75.0	61.0	79.3

Table 15 Ranksum between Lab/No Lab group with tsgrp and lab_nolabgrp

* = < 0.10, ** = < 0.05, *** = < 0.01

z-statistic (p-value)	E1	E2	E3	E4	FE
No Lab/Lab with tsgrp	-1.034 (0.3011)	-2.304 (0.0212)**	-2.609 (0.0091)***	-1.394 (0.1634)	-1.392 (0.1638)
No Lab/Lab with lab_nolabgrp	-1.019 (0.3080)	-2.274 (0.0230)**	-2.625 (0.0087)***	-1.533 (0.1254)	-1.418 (0.1563)

To test this sub-hypothesis, the Engineering students were broken into two groups based on whether or not they were enrolled in the laboratory. The medians and means were calculated for each exam and the Wilcoxon rank-sum test was used to see if there was a statistically significant difference between the medians in the laboratory and no laboratory group. The exam averages by laboratory are shown in Table 16. A graph of these data is shown in Figure 14.

Table 16 Exam Averages by Laboratory in Engineering Class

Mean (Median) (N = 428)	Exam 1 (%)	Exam 2 (%)	Exam 3 (%)	Exam 4 (%)	Final Exam (%)
No Lab (N = 298)	78.4 (81.5)	68.4 (70.0)	72.7 (74.0)	59.2 (59.0)	77.4 (78.8)
Lab (N = 130)	80.2 (82.0)	71.7 (74.0)	75.8 (78.5)	61.4 (63.0)	79.3 (80.0)
Class Average	78.9 (82.0)	69.4 (71.0)	73.6 (75.0)	59.9 (61.0)	78 (79.3)

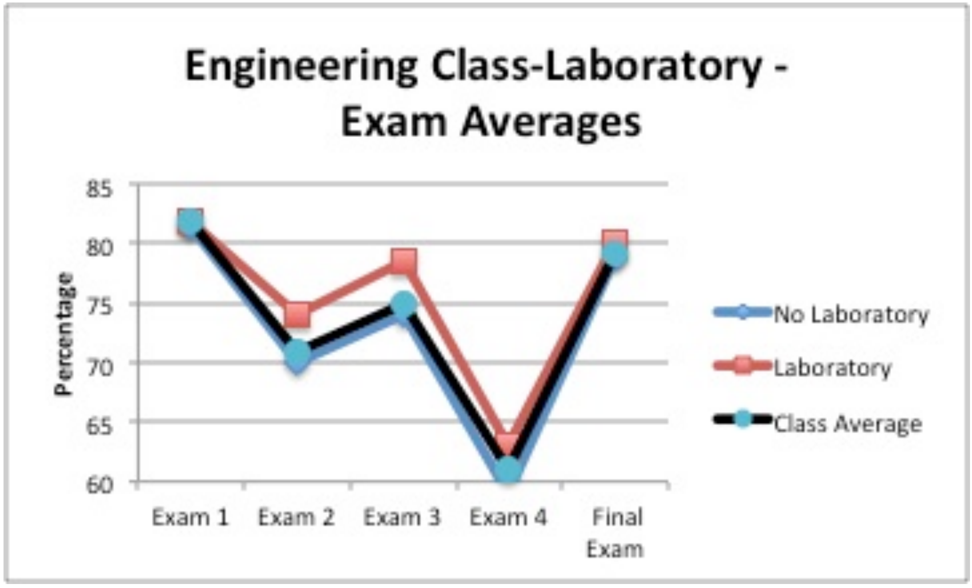


Figure 14 Exam Averages by Laboratory in Engineering Class

Overall, it appears that the students who are taking laboratory in the Engineering class always outperform the non-laboratory students, as well as having above average performance on each exam. A rank-sum test was run to determine if these median exam averages were statistically significantly different between the laboratory and non-laboratory students. The results of the rank-sum test are shown in Table 17.

Table 17 Wilcoxon ranksum test results of analysis of median differences between No Laboratory and Laboratory students in Engineering Class

* = < 0.10, ** = < 0.05, *** = < 0.01

z-statistic (p-value)	Exam 1	Exam 2	Exam 3	Exam 4	Final Exam
No Lab/ Lab	-1.034 (0.3011)	-2.304 (0.0212)**	-2.609 (0.0091)***	-1.394 (0.1634)	-1.392 (0.1638)

The only statistically significant differences between the laboratory and non-lab are for the exam performance on the second and third exams. In both of these cases, the students in the laboratory have higher median exam averages than the non-laboratory students. For the second and third exam there was a large overlap between the laboratory and exam content coverage, and for these exams, the laboratory students' exam averages were higher than the non-laboratory students. This positive effect of taking laboratory may be due to the extra amount of time spent working on the material while in laboratory. For the other exams there was not an appreciable overlap between the exam and laboratory content coverage. This suggests that enrolling the laboratory can improve students' exam performance when the coverage matches exam content coverage. Matz and colleagues (2012) have presented similar research indicating that concurrent enrollment in the laboratory improves student exam performance.

Sub Hypothesis

There is a positive relationship between the amount of experience a teaching assistant has teaching general chemistry and the students' exam averages.

Method

The number of general chemistry courses taught in either the fall or spring semesters since fall 2006 were counted for each teaching assistant in both courses. This value became their TA experience. This value was then assigned to students in their sections in the fall 2010 data. In the Engineering class the TA experience values were 0, 1, 4, and 5, while in the Science Majors class they were 0, 1.5, 3, and 7. The sections with the 1.5 level of TA experience has two TA's, one with three semesters of teaching experience and one with none, so an average was taken for the TA experience in those sections. The statistical results are the same if this recitation section is grouped in with the TA experience level 3 group based on the more experienced TA's teaching experience. TA's with no prior teaching experience are in charge of most of the recitations in both courses (71.3% in the Engineering class and 86.4% in the Science Majors class). The correlation was calculated between students' exam performance and their teaching assistant's general chemistry teaching experience. The results of this analysis are presented in Table 18 for the two courses.

Results

Table 18 Correlations between TA experience and Exam Averages in Engineering and Science Majors classes * = < 0.02, ** = < 0.01, * = < 0.002**

correlation coefficient (p-value)	Exam 1 (%)	Exam 2 (%)	Exam 3 (%)	Exam 4 (%)	Final Exam (%)
Engineering Class (N = 428)	-0.1189 (0.2074)	-0.1248 (0.1466)	-0.0355 (1.0000)	-0.0550 (1.0000)	-0.1235 (0.1586)
Science Majors Class (N = 354)	0.0929 (1.0000)	0.0605 (1.0000)	0.0302 (1.0000)	0.0212 (1.0000)	0.0451 (1.0000)

There are no statistically significant correlations between TA experience and students' exam performances in either the Engineering or the Science Majors class. This means the number of general chemistry courses that a teaching assistant has taught has no apparent influence on students' exam performance.

Findings/Conclusions about Exam Averages

When each class is divided into quartiles based on overall exam performance, all the exams have statistically significant differences in performance between the Engineering and Science Majors classes for each quartile level except for the Bottom quartile on exam two. The Engineering students have higher median exam averages for all the quartiles on exam one and the final exam, while the Science Majors students out perform the Engineering students on exams two through four except for the bottom quartile on exam two. These differences may be due to the fact that the first exam is mainly material that is very similar to high school chemistry content, and the Engineering students may have a better grasp of high school chemistry content than the Science Major students. Although the two classes had similar distributions of Toledo scores after equivalencing, that does not rule out the idea that the Engineering students, by virtue of their better mathematics background, have an advantage when it comes to the first exam, as a strong mathematics background is a predictor for success in chemistry courses in college (Nordstrom, 1990). As mentioned in the Chapter 3, there were more quantitative questions on the first exam in the Engineering course as compared to the Science Majors class.

Engineering students' higher median exam averages on the final exam across the quartiles may be a result of the reward system in each course for relearning previously missed material. In the Science Majors course, a replacement exam is available for students to take between the fourth exam and the final exam. The students may choose only the exam that they scored the lowest on, to replace with this exam. This may lead them to focus on relearning the material on that exam, as opposed to studying all the material covered during the semester for the final exam that occurs shortly after the replacement exam. If students choose not to take the replacement exam, they may also be under the false impression that they know the material from the semester better than they actually do, and therefore they don't study as much for the final exam as they would have if they had taken the replacement exam. In the Engineering class, relearning missed material is rewarded on the final exam itself. There are sections that cover material similar to the content on the first four exams along with a section that contains material covered since the fourth exam. Students can earn all of the points back on each section on which they out perform their score on the original exam. In this way, the instructor of the Engineering course encourages students to restudy material from the whole semester, throughout the whole course, in order to prove that they have learned the material by the end of the semester. This is shown to be effective in that the Engineering students' performance on the final exam, on a numerical basis is higher than Science Majors students across all the quartiles.

The Science Majors students out performed the Engineering students on exams two through four across all the quartile levels except on exam two, where both classes Bottom quartiles students had median exam averages that were not statistically significantly different. It is possible that in both classes, the second exam is the place where the lowest performing students realize that their high school chemistry knowledge is not going to sustain them through the class. These students may not have a very strong high school chemistry background to start with and so, they "run out" of that knowledge after the first exam, while other students do not. The content of exams two through four are also increasingly different between the two courses, as the Engineering class more through two-semesters worth of content, while the Science Majors class spends more time on bonding and structure-function relationships near the beginning of the semester. An alternate explanation for these results is that the second through fourth exams may have been more difficult in the Engineering course as compared to the Science Majors course, leading to a decrease in Engineering students' performance relative to the Science Majors.

Overall, each quartile outperformed the quartiles below them in terms of their median exam average, as one would expect. The fourth exam in the Engineering course is the one where all the students performed poorly, but it may be that the Middle Top and Top students' performance was coming into line with how the Middle Bottom and Bottom students' performance would have been anyway. Within the Science Majors class, all the quartile pairs had statistically significantly different exam performance on every exam and each quartile outperformed the quartiles below them.

When the material covered on the exams and in the laboratory is closely aligned, students in the laboratory performed better on the exam when compared to the non-laboratory students, as occurred with the second and third exam in the Engineering course. There were no statistically significant correlations between teaching assistants level of experience teaching general chemistry and students' exam performance. This is more than likely due to the fact that 71.3% of the Engineering students and 86.4% of the Science Majors students are being taught by teaching assistants who are teaching their first general chemistry course. There may not have been enough students with more experienced teaching assistants to make a difference statistically.

Delta Z-Scores as Measures of Exam Performance

Delta z-scores are a way of quantifying the differences in slopes between each exam pair for a particular group of students. Since delta z-scores are the differences between the z-scores for the two exams in the exam pairs and because z-scores put all the performances on the same scale, these values are easily compared across the Engineering and Science Majors classes. We can use these delta z-scores for the exam pairs throughout the semester to see if the overall course structure makes a difference on student exam performance when grouped by quartiles, deciles and grades, as well as the effect of TA experience in the two classes and the laboratory experience in the Engineering course.

Quantitatively, taking an individual score and subtracting the mean and dividing by the standard deviation leads to a z-score. A delta z-score is the difference between the z-scores for two exams. The z-scores were calculated at an individual student level for each exam and then these z-scores were subtracted from each other, the later exam minus the earlier exam, to get a delta z-score for each exam pair for each student. The delta z-scores presented in the following tables are the mean delta z-scores for the students in a particular group, (quartile, TA experience or lab/no lab in the Engineering class). As these values have already been set to the same scale, with a mean of zero and a standard deviation of one, exam performance between exam pairs can be compared across the Engineering and Science Majors classes, as well as across the semester. These delta z-scores can also be thought of as a measure of the slope between the two exam scores in the exam pair in the Exam Average graphs. The class average for the delta z-scores was calculated by taking the average of the amount of change for each exam pair in each course. Tables 19 and 20 show the delta z-scores for exam pairs for students divided into quartiles in both the Engineering and the Science Majors classes.

Sub Hypothesis

The mean delta z-scores for any exam pair will be the same for each quartile in the Engineering and Science Majors class.

Delta Z-Scores by Quartiles

The delta z-scores for the exam pairs for the two courses are shown in Tables 19 and 20. Graphs of these data are shown in Figures 15 and 16. The goal of the graphs is to look at trends for changes in exam performance. Table 21 shows the results of an analysis of the differences in delta z-scores comparing the two courses across the quartiles. As delta z-scores are not commonly used tools for comparisons there are no statistical tests to determine if the values are do to random chance or real differences between the object of the measurements besides comparing the values for the two courses to each other.

Table 19 Delta z-scores for Engineering Class

Mean (Median) (N = 428)	Exam1, Exam2 delta z-score	Exam2, Exam3 delta z-score	Exam3, Exam4 delta z-score	Exam4, Final Exam delta z-score
Top Quartile (N = 110)	0.0717 (0.1202)	0.0497 (-0.0353)	0.0877 (0.1489)	0.0119 (0.0181)
Middle Top (N = 104)	0.0410 (-0.0029)	-0.1060 (-0.0733)	0.0427 (-0.0381)	-0.0181 (0.0173)
Middle Bottom (N = 106)	-0.0742 (-0.1695)	-0.047 (-0.0311)	0.0178 (0.0013)	0.0566 (-0.0279)
Bottom (N = 90)	-0.0417 (-0.0490)	0.0845 (0.0107)	-0.1302 (-0.1560)	-0.0479 (-0.0585)
Class Average	-0.0005 (0.0185)	-0.0033 (-0.0246)	0.0045 (0.0011)	0.0006 (-0.0054)

Table 20 Delta z-scores for Science Majors Class

Mean (Median) (N = 354)	Exam1, Exam2 delta z-score	Exam2, Exam3 delta z-score	Exam3, Exam4 delta z-score	Exam4, Final Exam delta z-score
Top Quartile (N = 89)	0.0252 (-0.1320)	-0.0346 (-0.0625)	-0.0445 (-0.1283)	0.1513 (0.1971)

Table 20 Delta z-scores for Science Majors Class ctd

Mean (Median) (N = 354)	Exam1, Exam2 delta z-score	Exam2, Exam3 delta z-score	Exam3, Exam4 delta z-score	Exam4, Final Exam delta z-score
Middle Top (N = 88)	0.1151 (0.1306)	0.0040 (-0.0188)	-0.1374 (-0.1838)	0.1331 (0.0985)
Middle Bottom (N = 87)	-0.0267 (-0.0378)	-0.0936 (-0.0522)	0.1199 (-0.0051)	-0.1573 (-0.1831)
Bottom (N = 90)	-0.1339 (-0.0600)	0.1371 (0.0956)	0.0570 (0.0733)	-0.1205 (-0.1932)
Class Average	-0.0057 (-0.0308)	0.0042 (-0.0105)	-0.0014 (-0.0588)	0.0018 (0.0303)

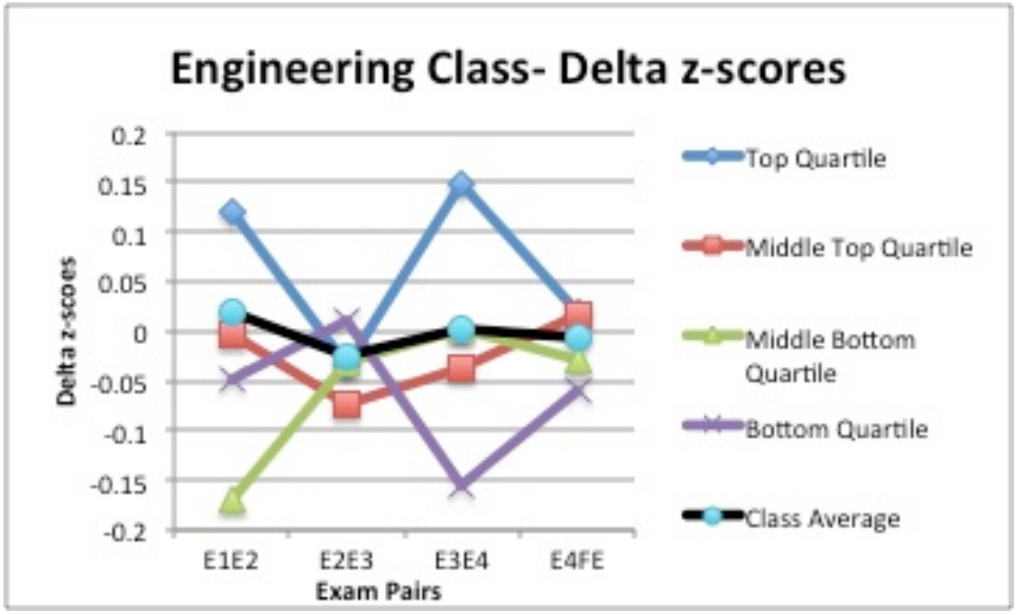


Figure 15 Delta z-scores for Exam Pairs for Engineering Class

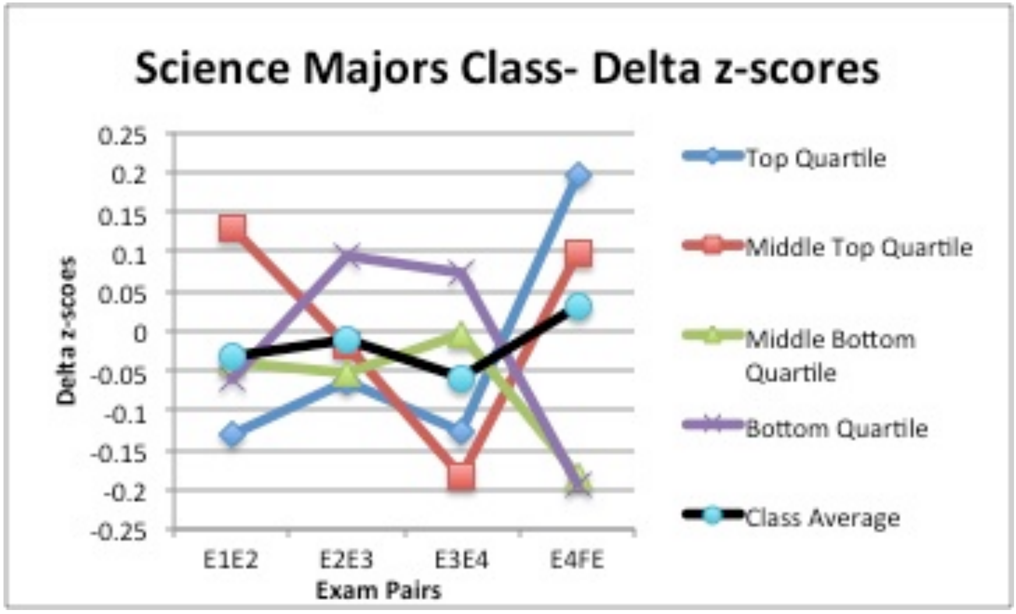


Figure 16 Delta z-scores for Science Majors Class

The graphs of the delta z-scores give an idea of which groups doing better on each exam pair in each course. On the first exam pair in the Engineering course the Top quartile of students did much better than the others, and the Middle Bottom quartile had the worst drop in performance. For the first exam pair in the Science Majors course, the Middle Top students were the high performers, while the Top quartile had the lowest performance. For the next exam pair, Exam two and Exam three, the Bottom quartile in the Engineering and Science Majors course both had the most improvement. Moving to the third exam pair, the Top quartile in the Engineering course had a large improvement compared to the rest of the course, and the Bottom quartile's performance in the Science Majors class decreased slightly, but was still far above the rest of the quartiles.

But the most interesting differences occur during the last exam pair, between the fourth one-hour exam and the Final exam. In the Science Majors course, the Top quartile of students' performance improves drastically from the prior exam pair. The Middle Top quartile also improves to a lesser degree. When you compare that to the same quartiles in the

Engineering course, the Engineering students’ performance improves slightly for the Middle Top quartile and decreases for the Top quartile. This may be because these students are content with their scores and are focusing their studying elsewhere.

Focusing on the two lower performing quartiles of students again leads to interesting differences between the two courses on the last exam pair. The Science Majors performance drops quite a bit for both the Middle Bottom and Bottom quartile of students on the last exam pair. This same trend is not observed in the Engineering course. In that course, the students in both the Middle Bottom and Bottom quartile improve moving from the fourth exam to the final. These differences suggest that the lower performing students in the Engineering course are making use of the ability to earn back points throughout the semester to improve their final exam performance, while the lower performing students in the Science Majors course are not performing as well on the final exam, possibly due to a lack of studying for the final exam. Their studying may be focused on the material on the replacement exam. It appears that this focus does not help them improve their final exam performance as much as the lower performing Engineering students. Another way to analyze this is to compare the differences in delta z-scores for the exam pairs between the two courses, shown in Table 21.

Table 21 Wilcoxon ranksum test results of analysis of median differences between Engineering and Science Majors Classes * = < 0.10, ** = < 0.05, * = < 0.01**

z-statistic (p-value)	Exam1, Exam2 delta z-score	Exam2, Exam3 delta z-score	Exam3, Exam4 delta z-score	Exam4, Final Exam delta z-score
Top	1.245 (0.2130)	0.822 (0.4111)	1.627 (0.1038)	-1.649 (0.0992)*
Middle Top	-1.103 (0.2702)	-0.931 (0.3521)	1.277 (0.2015)	-1.290 (0.1970)
Middle Bottom	-0.435 (0.6635)	0.399 (0.6900)	-0.585 (0.5583)	1.748 (0.0804)*
Bottom	0.498 (0.6184)	-0.371 (0.7105)	-1.704 (0.0884)*	1.088 (0.2764)

While the Science Majors students outperform the Engineering students for the Top Quartile on the last exam pair, and for the Bottom quartile for the second to last exam pair, the most interesting results come when looking at comparing the Middle Bottom quartile performance on the last exam pair. For this exam pair, the Middle Bottom quartile of students in the Engineering course's performance is much better than the Science Majors students. The Middle Bottom quartile from the Engineering course ended up near the class average for change in delta z-score for that exam pair, while the Science Majors students ended up far below the class average for that exam pair. This indicates that using resurrection points on a comprehensive final exam leads the Middle Bottom quartile of students in particular to improve their performance on the final exam as compared to students who use a replacement exam as a way to earn points for relearning missed material. As noted above the Middle Bottom quartile of students appears to perform better on the final exam in the Engineering course than the Science Majors course. This may be due to the fact that the Engineering students can earn back points on content throughout the semester, and so the Middle Bottom students are using studying habits that improve their performance on the whole final overall, as opposed to studying once particular set of information, as in the Science Majors course. This assumed difference in studying methods may have lead to the difference in exam performance for the Middle Bottom students in the Engineering course. Assessing students' studying habits throughout the semester is suggested as a project for future work in the Conclusion chapter. The results of the Kruskal-Wallis test presented in Table 22, show if there is any statistically significant difference in the delta z-scores for each exam pair across all four quartiles in a course. If there is a statistically significant difference for a particular exam pair, a rank-sum test is used to show what groups within that course are causing the significant difference. These tests are similar to an ANOVA and the pairwise t-tests that follow, to determine what groups are leading to the significant overall ANOVA test.

Table 22 Kruskal-Wallis equality-of-populations rank test results of analysis of median differences within the Engineering and Science Majors Classes * = < 0.10, ** = < 0.05, * = <0.01**

χ^2 statistic (p-value)	Exam1, Exam2 delta z-score	Exam2, Exam3 delta z-score	Exam3, Exam4 delta z-score	Exam4, Final Exam delta z-score
Engineering Class (df = 3)	3.744 (0.2904)	2.251 (0.5221)	3.899 (0.2725)	0.610 (0.8942)
Science Majors Class (df = 3)	4.073 (0.2536)	2.920 (0.4041)	4.404 (0.2210)	11.159 (0.0109)**

Table 23 Ranksum results for Exam 4 Final Exam delta z-score for Science Majors Class * = < 0.02, ** = < 0.008, ** = < 0.002

Group Compared	z-statistic (p-value)	Group Compared	z-statistic (p-value)
Top / Middle Top	1.144 (0.2527)	Middle Top/ Middle Bottom	1.726 (0.0843)
Top/ Middle Bottom	2.336 (0.0195)*	Middle Top/ Bottom	1.984 (0.0473)
Top/ Bottom	2.448 (0.0144)*	Middle Bottom/ Bottom	-0.781 (0.4347)

The only statistically significant difference within either class's delta z-score performance is for the Science Majors Class for the Exam four – Final Exam exam pair, shown in Table 23. The difference is due to the statistically significant difference between the Top and Middle Bottom and the Top and Bottom students' performance. The Top quartile of students had a large positive median delta z-score, while both the Middle Bottom and Bottom students had large negative median delta z-scores for the Exam four – Final Exam exam pair.

Delta Z-Scores by No Laboratory/Laboratory in Engineering Class

The results for the comparison of delta z-scores by laboratory in the Engineering course are shown in Table 24. A graph of these data is shown in Figure 17.

Table 24 Delta z-scores for Engineering Class - Laboratory

Mean (Median) (N = 428)	Exam1, Exam2 delta z-score	Exam2, Exam3 delta z-score	Exam3, Exam4 delta z-score	Exam4, Final Exam delta z-score
No Laboratory (N = 298)	-0.0273 (-0.0490)	-0.0078 (-0.0211)	0.0407 (0.0408)	-0.0089 (-0.0356)
Laboratory (N = 130)	0.0610 (0.1251)	0.0069 (-0.0272)	-0.0785 (-0.1175)	0.0224 (0.0327)
Class Average	-0.0005 (0.0185)	-0.0033 (-0.0246)	0.0045 (0.0011)	0.0006 (-0.0054)

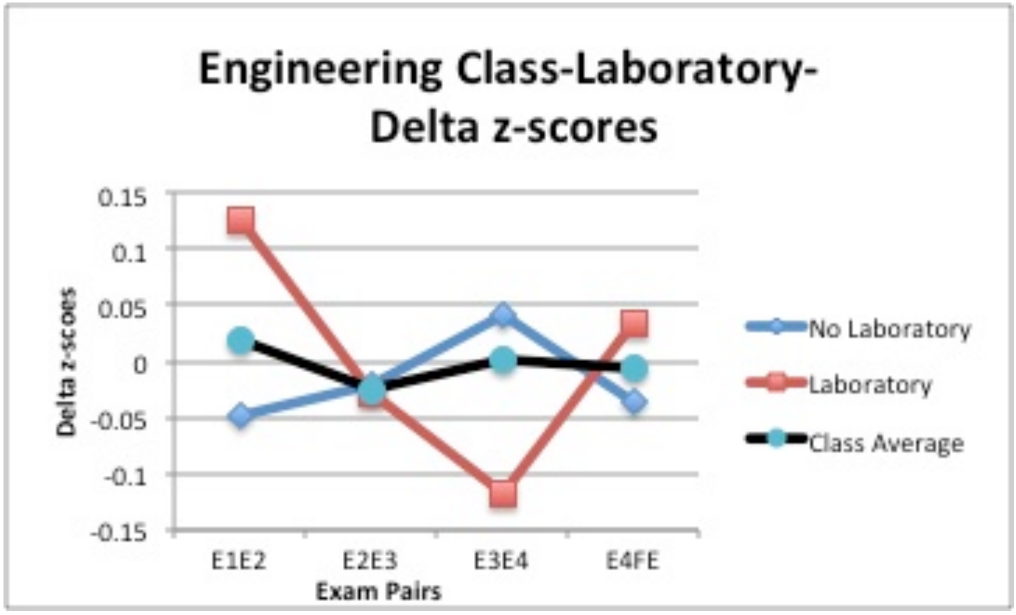


Figure 17 Delta z-scores for Engineering Class - Laboratory

Table 25 Ranksum results for Delta z-scores for Engineering Class - Laboratory * = < 0.10, ** = < 0.05, * = < 0.01**

z-statistic (p-value) (N = 428)	Exam1, Exam2 delta z-score	Exam2, Exam3 delta z-score	Exam3, Exam4 delta z-score	Exam4, Final Exam delta z-score
No Laboratory/ Laboratory	-1.511 (0.1308)	-0.238 (0.8123)	1.755 (0.0793)**	-0.566 (0.5712)

There is only one statistically significant difference in median delta z-scores for the Exam three – Exam four exam pair, shown in Table 25. For this exam pair, the students in the laboratory have a negative median delta z-score, indicating that their exam performance on the fourth exam is worse than their Exam three performance relative to the mean. The students not enrolled in the laboratory have a positive median delta z-score indicating that their exam performance relative to the mean, improves on Exam four over their performance on Exam three. This does not change the importance of the laboratory's influence on the students' performance on the second and third exams. The content on the fourth exam in the Engineering course did not match the laboratory content between the third and fourth exam. The impact of the laboratory only occurs when the content in the laboratory aligns closely with the exam content, so it is not surprising that the non-laboratory student improved more than the laboratory students between the third and fourth exam.

Delta Z-Scores by TA Experience

The last area of analysis for the delta z-scores for each exam pair throughout the semester is to look at the correlation between teaching assistants' experience teaching general chemistry courses and the students' delta z-scores. The results of this analysis are shown in Table 26.

Table 26 Correlations between TA experience and delta z-scores
*** = < 0.10, ** = < 0.05, *** = < 0.01**

Correlation Coefficients (p-value)	Exam1, Exam2 delta z-score	Exam2, Exam3 delta z-score	Exam3, Exam4 delta z-score	Exam4, Final Exam delta z-score
Engineering Class (N = 428)	-0.0069 (1.0000)	0.0969 (0.4515)	-0.0268 (1.0000)	-0.0946 (0.5047)
Science Majors Class (N = 354)	-0.0404 (1.0000)	-0.0378 (1.0000)	-0.0105 (1.0000)	0.0276 (1.0000)

There are no statistically significant correlations between a teaching assistants' amount of general chemistry teaching experience and the students' median delta z-scores for any exam pair throughout the semester.

Findings/Conclusions for Delta Z-Scores

By comparing equivalent groups in the Engineering and Science Majors courses, differences in students' level of prior knowledge has been accounted for as much as possible. This means the comparisons made above should be the result of course level differences, not prior knowledge differences. One would expect that there would be differences in how students perform on exams in courses with such wide differences in course coverage on each exam. However, the delta z-scores between the two courses are not statistically significantly different. This indicates that in general, students in the two courses are learning to meet the course expectations in both classes at the same speed. Hidden curriculum (what we test is what students' think is important), and pedagogical ecology (idea that the set up of a traditional classroom (or any classroom) give cues as to the teacher and student role and their level of interaction.

Not having a lot of statistically significant differences between the Engineering and Science Majors class when testing the general differences between the two courses is not surprising as there are so many similarities between the two courses. The only major differences between the two courses are the amount of content covered in one semester and the reward system for relearning missed material. The reward system seems to have some effect on how the Middle Bottom quartile of students prepares for the final exam, leading to higher averages than the Science Majors Middle Bottom students. Overall, the students are on an identifiable trajectory early on the course, and their performance does not deviate from that path much over the course of the semester.

There are some differences in the changes in student movement across the semester, in particular for the last exam pair. Overall, the Middle Top and Bottom quartiles are moving similarly in the two courses. The Top quartile in the Science Majors course is moving more positively going into the final compared to the Engineering course. But the most interesting finding is that the Middle Bottom quartile in the Engineering course is significantly outperforming the Science Majors students in the same quartile, moving from the fourth exam to the final. This positive movement for the Middle Bottom Engineering students can be attributed to the resurrection points on the final exam in that course. Having the ability to earn points back on material throughout the semester appears to be a better motivator for these lower performing students to learn missed material when compared to a replacement exam.

Attitude Data

The Attitude toward the Subject of Chemistry Inventory version two (ASCI v2) is an eight item survey that asks students to use a seven point scale between to adjectives (ex. easy/hard) to describe their feelings about the subject of chemistry (Xu & Lewis, 2011). The eight items load onto two factors, intellectual accessibility and emotional satisfaction, as seen in Table 27. The individual item scores range from one to seven, for example, one being hard and seven being easy. The range of scores for the two factors are from four to 28, again with higher values indicating more intellectual accessibility or emotional satisfaction. The inventory was administered during the first week of classes and after the semester break to get an idea of where students' attitudes started and how much their attitudes changed over the course of the semester. The two courses will be compared to determine what changes in attitude occurred over the course of the semester, as will the laboratory and non-laboratory students in the Engineering course. Finally, correlations will be calculated between the amount of teaching experience each TA has in the two courses, and the students attitudes about the subject of chemistry. The subcategories of the two factors, intellectual accessibility and emotional satisfaction, are shown in Table 27.

Table 27 Subcategories of Intellectual Accessibility and Emotional Satisfaction in ASCI v2

Emotional Satisfaction	Intellectual Accessibility
Uncomfortable/Comfortable	Hard/Easy
Frustrating/Satisfying	Complicated/Simple
Unpleasant/Pleasant	Confusing/Clear
Chaotic/Organized	Challenging/Not Challenging

There were no major differences between the quartiles in either course with regard to the individual items on the ASCI version 2. There were also no major differences between the courses, or between the laboratory and non-laboratory students in the Engineering course for the individual items. Most of the items only showed that there were no differences pre/post in the students' attitudes about chemistry. There were also no statistically significant correlations between the level of TA experience and the change in students' attitudes over the course of the semester. As a result, the individual items on the ASCI version 2 will not be discussed further. There were differences within and between the courses, as well as between the laboratory and non-laboratory students, when the two factors within the ASCI version 2, intellectual accessibility and emotional satisfaction were analyzed. One of these factors also showed a correlation with TA experience in the Engineering course.

In order to study the changes in students' intellectual accessibility and emotional satisfaction over the course of the semester, the students' final score for the factor was plotted versus the amount of change that had occurred in their attitude over the course of the semester. The axes intersect at the class average post score for the factor of interest on the horizontal axis and at zero on the vertical axis. Figure 18 is a general picture of how the graph is drawn and how it can be interpreted. The final scores more from negative to positive from left to right, and the negative scores are on the bottom of the change axis. In this way the four quadrants of the graph can be labeled as positive or negative changes over the course of the semester. The upper half of the graph notes a positive change as the students have a more positive view of chemistry at the end of the semester, while the bottom half of the graph indicates a more negative view of the subject of chemistry at the end of the semester.

As the amount of overlap between the quartiles, particularly at the origin, it's difficult to tell what the trends between quartiles or even between the same quartiles in the Engineering and Science Majors courses. Because of this, the analysis will focus on the tables displaying the percentage of students in each quartile that ended up in the four quadrants of each graph. Each graph was designed in the same way with the origin at zero change on the y-axis, and the median final score for that factor on the x-axis. So starting in the upper right corner that would be Quadrant one. Students in this quadrant have an overall positive change in their attitude from the beginning of the semester, as well as an above median final score for the factor at hand. Quadrant two, in the upper left corner, again has this positive change in attitude over the course of the semester, but the students' overall final scores for the factor are below the median. Students in Quadrant three in the lower left corner of the graph also have below median final scores, but their views about chemistry have become more negative over the semester. Finally, students in Quadrant four, in the lower right corner, have more negative views of chemistry at the end of the semester, but they also have final scores that are above the median.

So overall, students in Quadrants one and three do not deviate from the trend they presented at the beginning of the semester. Students in Quadrant two, are ones who had relatively negative views of chemistry to start out with, but they have improved somewhat over the semester, even if they are not above the median yet. This is the quadrant; one would ideally like students to be moving into. Students in Quadrant four, had relatively positive views of chemistry at the beginning of the semester, but they have decreased over the course of the semester somewhat. Quadrant four is the least desirable quadrant for students to move into over time in a chemistry course.

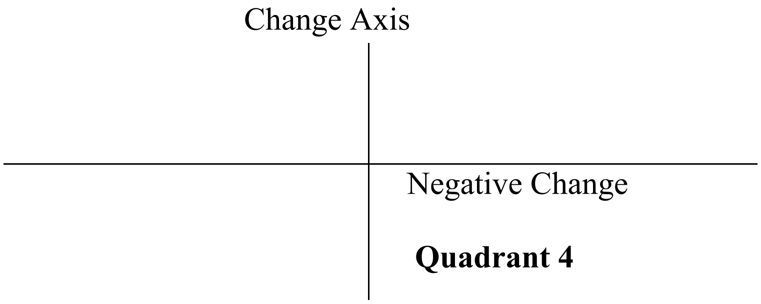


Figure 18 Description of Quadrants in Change versus Final Score graphs

Emotional Satisfaction

The Emotional Satisfaction information for the Engineering course is shown in Figure 19 and Table 28. For the Top Quartile of students 60% of the students were showing that they had followed the trend set at the beginning of the semester regarding how emotionally satisfying they felt chemistry was. Of the other 40% of the students, most of them ended up in Quadrant two, indicating that while their final scores were below the median, they had improved over the course of the semester, and the students believed chemistry was more satisfying at the end of the semester. While these students end up with scores that are below the median at the end of the semester, their views have become more positive, as indicated by the larger Emotional satisfaction value at the end of the semester compared to the beginning of the semester. As a larger value indicates a more positive view of the emotional satisfaction provided by the course, these students find chemistry more satisfying at the end of the semester compared to their views at the beginning of the semester. Only a small portion of students felt that chemistry was more frustrating at the end of the semester from the Top quartile in the Engineering course. The Middle Top students had 67.5% of their students maintain the trend of their views of the satisfaction they got from chemistry over the semester. A smaller percentage of students felt that chemistry was more satisfying at the end of the semester, and a slightly higher percentage of Middle Top students felt chemistry was more frustrating at the end of the semester. Sixty two and a third percent of the Middle Bottom students maintained their trajectory for how satisfying the feel chemistry is. Fewer of these students felt that chemistry was less satisfying at the end of the semester as compared to the other quartiles, and they have almost the same percentage of students as the Top quartile that felt that chemistry was more satisfying at end of the semester. Finally, 73.1% of the Bottom quartile of students maintained their trajectory for feeling satisfied with chemistry over the semester. The Bottom quartile also had the most students feel less satisfied at the end of the semester, and the least amount of students feel more satisfied at the semester.

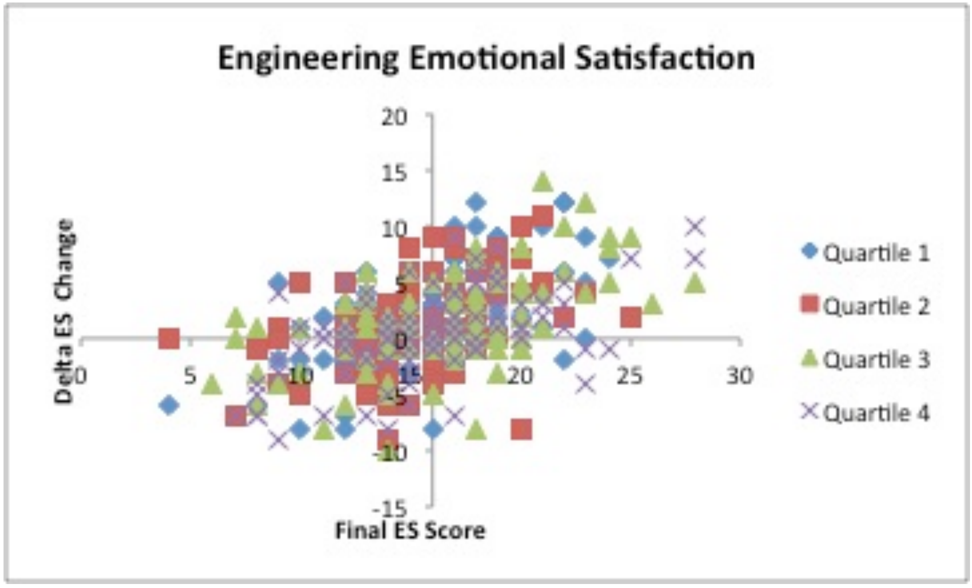


Figure 19 Engineering Emotional Satisfaction Change versus Final score

Table 28 Engineering Emotional Satisfaction Percentages of Quartiles in each Quadrant

Quartiles	Quadrant 1 (%)	Quadrant 2 (%)	Quadrant 3 (%)	Quadrant 4 (%)
Top (1)	32.7	30.9	27.3	9.1
Middle Top (2)	38.5	23.1	28.8	9.6
Middle Bottom (3)	42.5	29.2	19.8	8.5
Bottom (4)	56.4	16.7	16.7	10.2

The changes in Emotional Satisfaction for the Science Majors class are presented in Figure 20 and Table 29. For the Top Quartile of students, 65.2% of students maintained their trajectory about their satisfaction of chemistry from the beginning of the semester. A small percentage of students feel less satisfied at the end of the semester with compared to the beginning, while maintaining an above median final score. A little under a quarter of the students in the Top Quartile in the Science Majors course ended up in Quadrant two at the end of the semester, indicating that their initially negative views of chemistry were improved over the course of the semester.

Fifty seven point nine percent of students in the Middle Top Quartile did not have changes in emotional satisfaction that deviated from the trajectory established at the beginning of the semester. Less of these students ended up feeling more frustrated at the end of the semester, while still having final scores that were above the median. The Middle Top Quartile students had the highest percentage of students who were positively influenced in their view of how satisfying chemistry is of all the Science Majors students.

Sixty-nine percent of the Middle Bottom students did not change the trajectory of their views of how satisfying chemistry was over the course of the semester. The amount of Middle Bottom students who feel more frustrated with chemistry at the end of the semester was slightly more than the Top quartile of students, while there were less students who were positively influenced in their satisfaction with chemistry over the course of the semester.

Finally, 77.8% of students in the Bottom quartile did not change the trajectory of their satisfaction with chemistry over the course of the semester. The Bottom quartile of students also had the smallest percentage of students who were positively influenced by the course to feel more satisfied with chemistry at the end of the semester, and a medium sized portion of students who were more frustrated at the end of the semester, compared with the other quartiles in the Science Majors course.

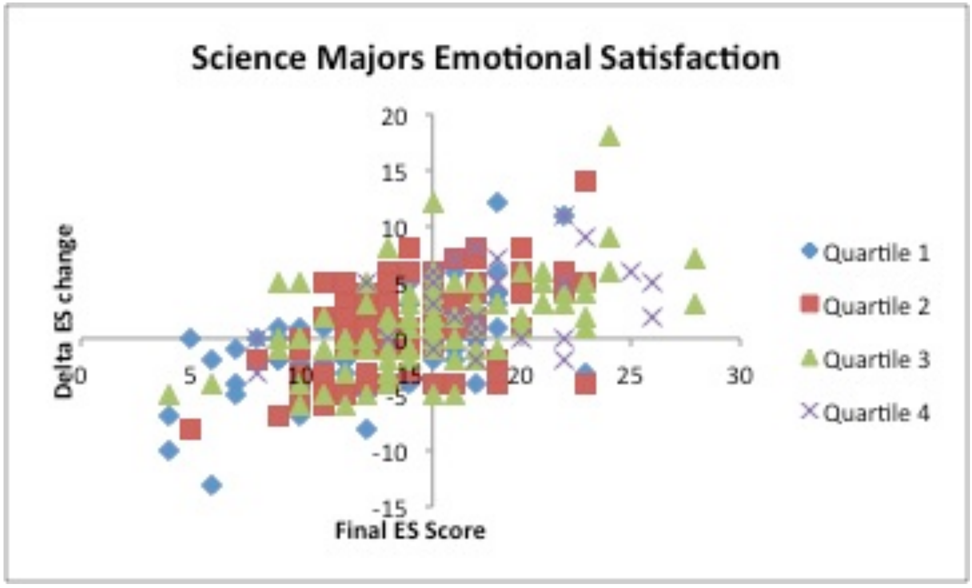


Figure 20 Science Majors Emotional Satisfaction Change versus Final Score

Table 29 Science Majors Emotional Satisfaction Percentage of Quartiles in Quadrants

Quartiles	Quadrant 1 (%)	Quadrant 2 (%)	Quadrant 3 (%)	Quadrant 4 (%)
Top (1)	31.5	24.7	33.7	10.1
Middle Top (2)	29.5	34.1	28.4	8.0
Middle Bottom (3)	44.8	20.7	24.2	10.3
Bottom (4)	66.7	13.3	11.1	8.9

More students in the Top Quartile in the Science Majors course did not change their views of chemistry drastically over the course of the semester. Comparing the percentage of students in the Top Quartile across the four quadrants in the two courses, there were fewer students in the Engineering course in the Top Quartile who were more frustrated, and more students who were positively influenced by their time in the course, as to their satisfaction with chemistry over the course of the semester. More students in the Middle Top quartile in the Engineering course were unchanged in their views, but more students were frustrated, and fewer students were positively influenced by the Engineering course to see chemistry as satisfying at the end of the semester. The larger changes come from the Middle Bottom and Bottom quartiles in both courses. The Science Majors students in the Middle Bottom quartile are less likely to change their opinions about the satisfaction they get from chemistry between the beginning and the end of the semester. More of the Middle Bottom quartile students in the Science Majors course are more frustrated with chemistry at the end of the semester, and less of them are more satisfied, when compared to the Engineering course. We see similar trend with the Bottom Quartile of students in the Science Majors course. This indicates that the Middle Bottom and Bottom quartile Engineering students are positively influenced by the course to feel more satisfied with chemistry at the end of the semester, compared to students in the Science Majors course. The Bottom Quartile of students in each course has the most positive views of the Emotional Satisfaction from the chemistry course. This is surprising given these students' low performance in the course. On would expect these students to have lower Emotional Satisfaction in the course, particularly by the end of the course.

Intellectual Accessibility

The changes in the Engineering students views of the intellectual accessibility of chemistry are shown in

Figure 21 and Table 30. Fifty-eight and one-tenth percent of the Top Quartile of students in the Engineering course trajectory of views of the intellectual accessibility of chemistry were unchanged over the course of the semester. Almost 15% of these students find chemistry more challenging at the end of the semester. Finally, over a quarter of the Top Quartile of students were positively influenced by the Engineering course to see chemistry as easier over the course of the semester.

The Middle Top quartile of Engineering students have a higher percentage of students that maintained their trajectory of feeling chemistry was challenging or not over the course of the semester. This trend of maintaining the same trajectory for the semester continues for the other quartiles in the Engineering course as well. More of the Middle Bottom students see chemistry as more challenging at the end, and only 23.1% of the students feel chemistry is more accessible at the end of the semester.

For the Middle Bottom students, 66.1% of the students' views were not influenced by the course, other than to make them more extreme over the course of the semester. Slightly more of these students were positively influenced by the course as compared to the Middle Top students, and less of them feel that chemistry is more frustrating at the end of the semester, while maintaining final scores above the median.

Finally, for the Bottom Quartile of students in the Engineering course, almost 70% have not changed their views of how challenging chemistry is, other than become more extreme over the semester. There are more students from this quartile who feel chemistry is challenging than the Middle Bottom quartile, and they are the students least likely to see chemistry as less challenging at the end of the semester.

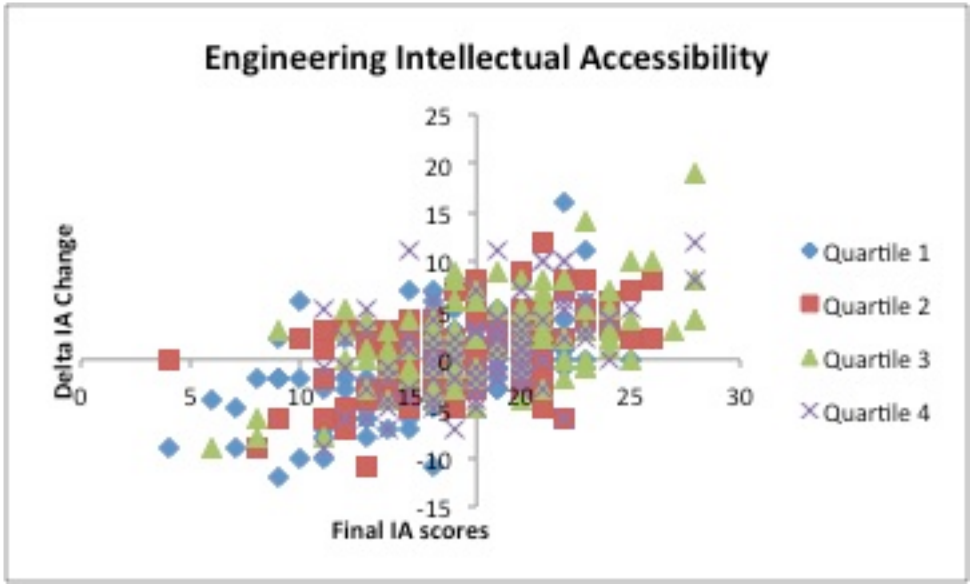


Figure 21 Engineering Intellectual Accessibility Change versus Final Score

Table 30 Engineering Intellectual Accessibility Percentages of Quartiles in Quadrants

Quartiles	Quadrant 1 (%)	Quadrant 2 (%)	Quadrant 3 (%)	Quadrant 4 (%)
Top (1)	23.6	27.3	34.5	14.6
Middle Top (2)	31.7	23.1	28.8	16.3
Middle Bottom (3)	45.3	23.6	20.8	10.3
Bottom (4)	55.6	20.4	12.9	11.1

The Science Majors graph and table about the changes in their views of the intellectual accessibility of chemistry over the semester are shown in Table 31 and Figure 22. Almost 63% of the Top Quartile of Science Majors students have only become more extreme in their views of how challenging chemistry is by the end of the semester. There are 11.3%

of Top Quartile students who believe that chemistry is harder at the end of the semester, and about a quarter who were positively influenced by the course who believe that chemistry is easier than they expected at the end of the semester.

Sixty-seven and one-tenth percent of the Middle Top students have not changed their trajectory from the beginning of the semester about how challenging chemistry is. Less of these students find chemistry to be more challenging at the end of the semester compared to the Top Quartile of students. In addition, only 22.7% of the Middle Top students find chemistry is to be less challenging at the end of the semester.

Over 80% of the Middle Bottom students in the Science Majors course have become more extreme in their views of the level of challenge presented by chemistry by the end of the semester. Only 3.5% of these students think chemistry is more challenging while having final scores that are above the median. Almost 15% of these students think chemistry is more difficult at the end of the semester compared to where they were at the beginning of the semester.

The amount of students who held more extreme positions about the level of challenge presented by chemistry at the end of the semester was slightly lower for the Bottom Quartile of students. There were almost twice as many students, percentage wise, in the Bottom Quartile who found chemistry to be more challenging by the end of the semester. A little over 14% of the Bottom quartile of students in the Science Majors course found the topic of chemistry to be less challenging at the end of the semester.

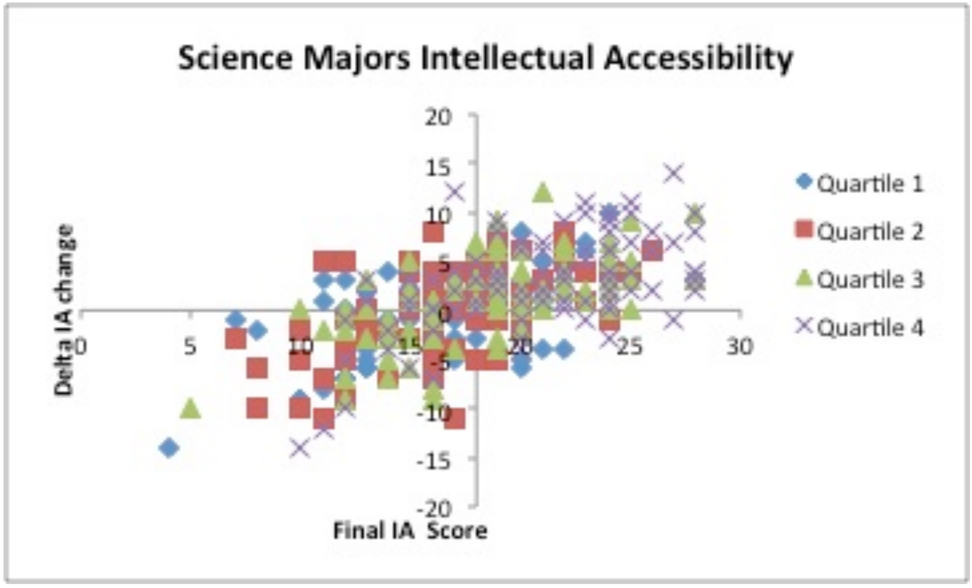


Figure 22 Science Majors Intellectual Accessibility Change versus Final Score

Table 31 Science Majors Intellectual Accessibility Percentage of Quartile in each Quadrant

Quartiles	Quadrant 1 (%)	Quadrant 2 (%)	Quadrant 3 (%)	Quadrant 4 (%)
Top (1)	23.6	25.8	39.3	11.3
Middle Top (2)	34.1	22.7	33.0	10.2
Middle Bottom (3)	57.5	14.9	24.1	3.5
Bottom (4)	67.8	14.4	11.1	6.7

Comparing the two courses in terms of the level of challenge presented by chemistry leads to some interesting findings. There are more students in the Science Majors course who stay of the same trajectory about how challenging chemistry is for all four quartiles. There are also a higher percentage of students in the Engineering course who find chemistry to be more challenging at the end of the semester across all four quartiles. However, the percentage of students in the Engineering course that were positively influenced by the course to see the chemistry content as easier by the end of the semester is also higher than the percentage of students in the same quadrant (Quadrant 2) in the Science Majors class, for all four quartiles. So overall, while the amount of content covered in the Engineering course may have led some students to feel that chemistry was more difficult than they expected, there were more students whose views of the difficulty level presented that were still below the median at the end of the semester, but that were improving. The Bottom Quartile of students in each course has the most positive views of the Intellectual Accessibility of the chemistry course. This is surprising given these students' low performance in the course. On would expect these students to have lower Intellectual Accessibility scores in the course, particularly by the end of the semester.

Attitudes about Chemistry Lab versus No Lab Engineering Course

The changes in students' views of the emotional satisfaction they get from chemistry in the Engineering course, when divided into students enrolled in laboratory and students who are not enrolled in laboratory are shown in Figure 23 and Table 32. Over 65% of both the laboratory and non-laboratory students' views of how satisfying chemistry is did not change other than to become more extreme over the course of the semester. There were slightly more students in the non-laboratory group who felt that chemistry was more frustrating at the end of the semester as compared to the laboratory students. Slightly more than 25% of the laboratory students were positively influenced by the course in their views of the satisfaction gained by studying chemistry, even if their final scores were below the median value for the course. A similar value can be found for the number of non-laboratory students who found chemistry to be more satisfying at the end of the semester. Overall, it does not appear that the laboratory had a large influence on changing how satisfied students' felt studying chemistry over the course of the semester.

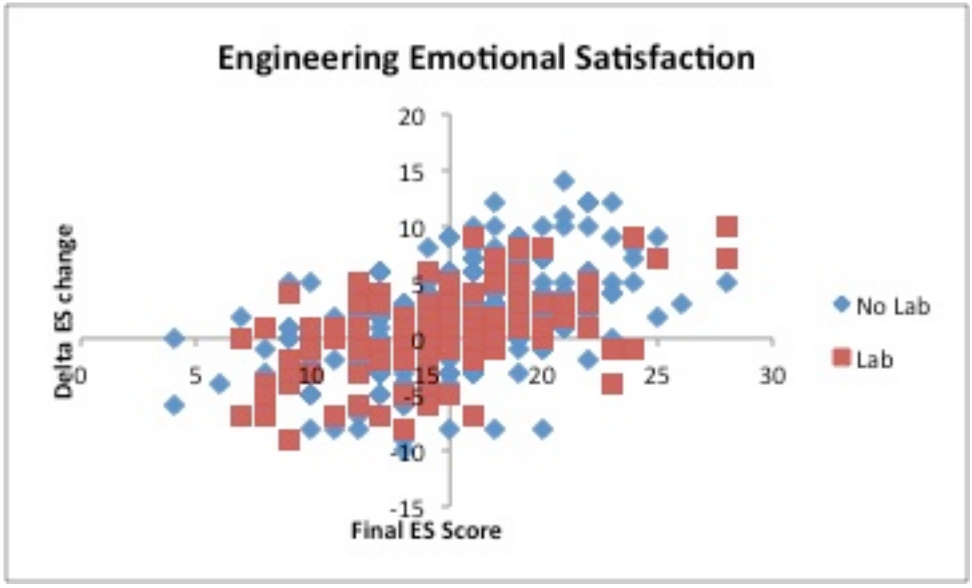


Figure 23 Engineering Emotional Satisfaction Change versus Final Score-Lab/No Lab

Table 32 Engineering Emotional Satisfaction Percentage of Lab/No Lab in Quadrants

Group	Quadrant 1 (%)	Quadrant 2 (%)	Quadrant 3 (%)	Quadrant 4 (%)
No Lab	44.3	24.8	21.5	9.4
Lab	38.5	25.4	26.9	9.2

The changes in the laboratory and non-laboratory students in the Engineering course views of the intellectual accessibility of chemistry are shown in Table 33 and Figure 24. In this case the percentage of students whose views of how difficult chemistry was, that didn't change except to become more extreme over the semester is higher in the non-laboratory group. The non-laboratory group also has a higher percentage of students who felt at the end of the semester that chemistry was easier to learn, even is their final scores are below the median for the class. Finally, only 11.7% of the non-laboratory students said that chemistry was more difficult at the end of the semester while having final scores above the median. These values suggest that taking the laboratory is leading students to believe that chemistry is more challenging at the end of the semester as compared to the students who did not take the laboratory.

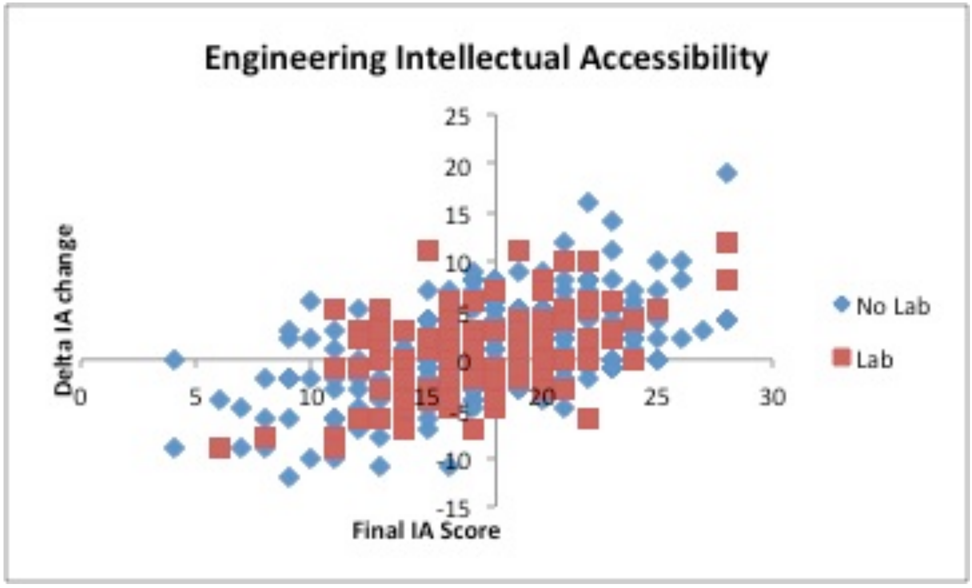


Figure 24 Engineering Intellectual Accessibility Change versus Final Score - Lab/No Lab

Table 33 Engineering Intellectual Accessibility Percentage of Lab/No Lab group in each Quadrant

Group	Quadrant 1 (%)	Quadrant 2 (%)	Quadrant 3 (%)	Quadrant 4 (%)
No Lab	38.6	24.2	25.5	11.7
Lab	40.0	22.3	21.5	16.2

TA Experience Correlations with Students' Attitudes

The correlations between the levels of TA experience in the two courses and students' attitudes about chemistry were analyzed. The correlations between the TA experience and the two factors, Emotional Satisfaction and Intellectual Accessibility, are presented in Table 34.

The only statistically significant correlation to TA experience was the change in Intellectual Accessibility for the Engineering course. To investigate the correlation further, the median change in the Intellectual Accessibility Factor for the Engineering students for each level of TA experience is presented in
Table 35. There is a positive relationship between students' perception of their ability to learn chemistry when they had less experienced TA's, while the students who had more experienced TA's tended to have neutral or negative perceptions about their ability to learn chemistry at the end of the semester. This may be due to the less experienced being able to

articulate the content to the students in a way that they can understand, as opposed to the more experienced TA's who may not be able to articulate how or why they are presenting a particular idea. This is related to Vygotsky's (1978) zone-of-proximal development as it relates to the novice-expert spectrum. When discussing experts they are frequently characterized as having a large networks of connections between topics in their area of expertise, as well as being unable to articulate (Bransford, Brown, & Cocking, 2000). Another explanation for the positive relationship between less experienced TA's and students' improved perception of their ability to learn chemistry at the end of the semester may be related to the more experienced TA's challenging their students more to learn the content while the less experienced TA's are just working to cover all the material in the semester. In either case, there is a positive relationship between the less experienced TA's and their students' beliefs about their ability to learn chemistry at the end of the semester. While this connection is interesting, due to the low numbers of experienced TA's, the results may not be generalizable to other chemistry courses.

Table 34 Correlations between TA experience and Two Factors in ASCI v2 * = < 0.05, ** = < 0.025, * = < 0.005**

Correlation Coefficients (p-value)	Intellectual Accessibility Change	Emotional Satisfaction Change
Engineering (N = 428)	-0.1157 (0.0491)*	-0.0663 (0.5126)
Science Majors (N = 354)	0.0160 (1.0000)	-0.0516 (0.9996)

Table 35 Engineering median change in Intellectual Accessibility by TA experience Level

	TA exp = 0 (N = 307)	TA exp = 1 (N = 60)	TA exp = 4 (N = 26)	TA exp = 5 (N = 37)
Median IA change	1	2	-1	0

Attitude Findings

The major findings from the attitude data collected about the two courses are that the Science Majors course does a better job improving the Top and Middle Top quartiles of students' views of the emotional satisfaction one can get from studying chemistry, while the Engineering course does a better job of this for the Middle Bottom and Bottom quartiles of students. As far as the intellectual accessibility of chemistry, overall the Engineering course had more students who were negatively influenced by taking chemistry, indicating that they had positive views of how challenging chemistry would be at the beginning of the semester, but they found chemistry to be more challenging by the end of the semester. This trend held across all four quartiles. However, the Engineering students were also the ones who had the highest percentages of students who held negative views about how difficult chemistry would be coming in, that improved over the course of the semester. So, while the amount of content covered may have led students to feel overwhelmed and that chemistry was challenging, something about the Engineering course, perhaps the resurrection points as a motivational tool, allowed students to improve their view of the level of challenge presented by the subject of chemistry at the end of the semester.

Analysis of the views of the Engineering students, who were and were not enrolled in laboratory in the fall, indicates that the laboratory did not influence the students' perceptions of how satisfying it was to study chemistry. This is due to minor differences in the percentage of laboratory and non-laboratory students in each quadrant for the emotional satisfaction variable. Based on differences between the percentage of students who felt chemistry was more challenging at the end of the semester, one would say that students who took the laboratory in addition to the lecture for the Engineering course, found the content to be harder at the end of the semester compared to the non-laboratory students. This may be due to spending more time on the content because of the laboratory course. While this extra time on content seems to benefit students when they are testing on that material, it does not lead them to think chemistry is less challenging.

There was a correlation between the amount of TA experience in the Engineering course and the change how difficult to understand a student thought chemistry was. Relatively inexperienced TA's tended to have students with more positive views of how challenging the chemistry content was. This may be due to less experienced TA's being better able to articulate to students the content to be covered and how it relates to them and what they already know. The more experienced TA's may be too far removed from the experience of learning general chemistry for the first time to be able to explain it in a way that students can understand.

CHAPTER 5: ANALYSIS OF INTERVIEW DATA

Introduction

When studying problem solving it is important to differentiate between an exercise and a problem. There are several ways to define a problem. Hayes (1981) defined a problem as "whenever there is a gap between where you are now and where you want to be, and you don't know how to find a way across the gap". This is referred to as the gap idea. Another way to define a problem is as a path. A problem is a task for which a student does not see a direct path between what was given and the answer (Bodner G. M., 2003). Another way to think about problem solving is as "what you do, when you don't know what to do"

(Wheatley, 1984). This could mean working backwards, or drawing a picture, or any number of things. An exercise, in contrast, is a task for which a student knows a way to get an answer, and just needs to carry out the steps to get there (Bodner G. M., 2003). A key question for chemistry instructors is how to turn an exercise into a problem for students, to allow them to practice or test their problem solving skills. A discussion of this process can be found at the end of the chapter.

The problem solving behaviors of chemistry students has been studied for a long time (Bodner, 2000; Chandrasegaran, 2009; Nakhleh, 1993; Nurrenbern, 1987). Most analyses discuss comparing algorithmic and conceptual questions (Chandrasegaran, 2009; Nakhleh, 1993; Nurrenbern, 1987). These observed differences between the knowledge students demonstrate conceptually and algorithmically on chemistry questions has lead to changes in how chemistry textbooks are written and how chemistry is taught. Recent research suggests that these differences between the knowledge students display when answering algorithmic and conceptual questions is no longer valid, because students have developed algorithms for answering conceptual questions (Holme & Murphy, 2011). However, there have been no studies comparing students' problem solving behaviors when faced with two algorithmic questions, one that is more traditional, requiring dimensional analysis to solve and another that does not require dimensional analysis to solve. That comparison will be the focus of this chapter.

Population Description

In the Engineering course, 428 students gave consent and were included in the quantitative analysis. Of these students, 91.1% of them are male, and only 30.4% of them are taking the laboratory. In the Science Majors course, 354 students gave consent and were included in the quantitative analysis. Of these students, 57.1% of them are male, and all of them are enrolled in the laboratory as a co-requisite course. This is the population of students that was sampled from for the interviews. Only students who had given consent previously were selected for interviews. The selected students were given pseudonyms to protect their identities when discussing their work. A total of 20 students were interviewed from both the Engineering and Science Majors class. These interviews were video and audio recorded.

Interview Description

Forty volunteers were solicited from the two courses to take part in a one and a half hour to two-hour interview at a time of mutual convenience for the interviewer and interviewee. The students were offered free food for taking part in the interview. In order to get at the students' thoughts while they were working through the problems, a talk-aloud protocol was used during the interviews (Bowen C. W., 1994). This protocol asks students to verbalize what they are thinking about doing or why they are doing a given behavior while they are doing it. The interviews were video and audio recorded for data collection purposes. The interviews were transcribed as part of the data analysis process. The interview guide is given in the Appendix.

During the interview students were asked about their chemistry background, including high school and whether they enjoyed chemistry in high school, as well as the course they were taking now. If the student was in the Engineering course, they were asked if they were also enrolled in the laboratory. The students were provided with a computer with access to the Internet, their textbook, class notes, periodic tables if in the Engineering course, and any lab materials they desired to bring to the interview. The students were allowed to use any resources they desired to get information to answer the questions, but the interviewer didn't provide any assistance except for definitions of words to students for whom English was not their first language.

As shown in the interview guide, three stoichiometry and three thermochemistry questions were asked during the interview. One question for each of the topics was a simulation that the students were asked to use and work through some tasks related to simulation. The focus of the analysis in this chapter will be on the two-stoichiometry problems that are not simulations. Stoichiometry is a topic that permeates much of general chemistry and so it is important to understand how students solve these types of problems. It is possible that when students struggle with other concepts in chemistry, like equilibrium, kinetics, and thermochemistry, that difficulties students have solving stoichiometry tasks may be contributing to these difficulties in other subject areas. If strategies can be identified in either the Engineering or Science Majors courses that lead students to be more proficient at solving stoichiometry tasks, then these strategies may improve students' performance on other topics of chemistry as well.

Description of Participants in Analysis

Of the 40 interviews conducted, 11 students from the Engineering course and 12 students from the Science Majors course were selected for further analysis. In order to be included in the analysis, the students must have completed both the familiar and the unfamiliar stoichiometry tasks, as well as have taken the Toledo Test at the beginning of the semester. The Toledo Test given at the beginning of the semester as a measure of students' prior knowledge of chemistry. Four questions on that test covered the topic of stoichiometry and these were scored for correctness to give an idea of a students' prior knowledge of stoichiometry. The percentage of the four questions that each student got correct in both courses is included in Table 36 and Table 37. There is a range of scores for both of the courses, but the Science Majors course has more students who correctly solved at least one of the Toledo Test questions. There was a variety in the level of experience with stoichiometry tasks as measured by the stoichiometry questions on the Toledo Test taken at the beginning of the semester.

Table 36 Percentage Correct for Four Stoichiometry Problems - Engineering Interviews

Pseudonym	% Correct
Jeff	0
Sean	0
CC	0
Feng	0
Laura	25
Lindsey	25
Seth	25
Tom	50
Matt	50
Ryan	100
Randy	100

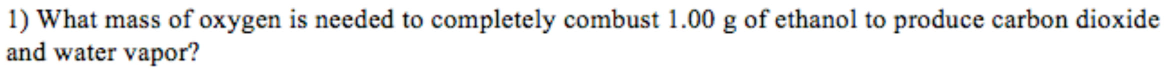
Table 37 Percentage Correct for Four Stoichiometry Problems - Science Majors Interviews

Pseudonym	% Correct
Tahir	0
Dave	25
Cullen	25
Jake	25
Cole	25
Joseph	50

Bill	50
Clark	50
Alice	50
Lin	50
Paul	50
Geoff	100

Interview Questions used for Analysis

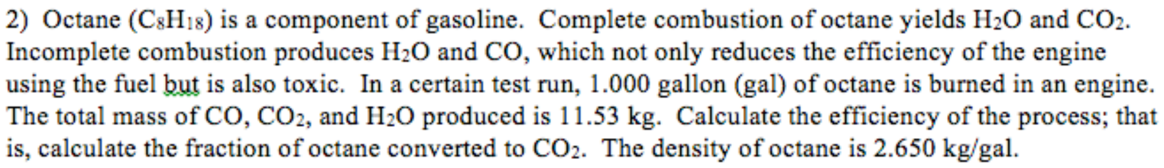
The two-stoichiometry problems that were analyzed are presented below in Figure 25 and Figure 26.



1) What mass of oxygen is needed to completely combust 1.00 g of ethanol to produce carbon dioxide and water vapor?

Figure 25 Familiar Stoichiometry Task From Rapid Knowledge Assessment*:* Correlating Student Reported Immediate First Steps and Problem Solving Efficiency. *Abstracts of Papers of the American Chemical Society, 239,* March 21, 2010.

The first stoichiometry problem dealt with gram-to-gram conversion, which is a common topic in general chemistry, and therefore one that students were expected to be familiar with. The question is also similar to the type of question seen in the lecture and on homework problems. This task requires the use of dimensional analysis for successful completion. Analysis of the students' problem solving behaviors while working on this task will be presented later.



2) Octane (C_8H_{18}) is a component of gasoline. Complete combustion of octane yields H_2O and CO_2. Incomplete combustion produces H_2O and CO, which not only reduces the efficiency of the engine using the fuel but is also toxic. In a certain test run, 1.000 gallon (gal) of octane is burned in an engine. The total mass of CO, CO_2, and H_2O produced is 11.53 kg. Calculate the efficiency of the process; that is, calculate the fraction of octane converted to CO_2. The density of octane is 2.650 kg/gal.

Figure 26 Unfamiliar Stoichiometry Task From *Chemistry 2nd Edition* (p. 114), by J. Burdge, 2009, New York, NY: McGraw Hill

The second stoichiometry task that students were asked to solve was an unfamiliar task. This task also comes from the stoichiometry chapter of the *Chemistry 2nd Edition* textbook, but this task is different from the familiar task described above. On the surface this question looks like a limiting reactant task, as the question discusses producing carbon monoxide and carbon dioxide. An incomplete combustion leading to carbon monoxide production is caused by a limited supply of oxygen. This information may lead students to treat this task as a limiting reactant problem, at least initially. However, this task can actually be solved without worrying about the amount of oxygen available. In fact this task does not require dimensional analysis at all, and can be solved using algebra. By comparing the two algorithmic questions, one requiring dimensional analysis and the other not, a hole in the problem solving literature will be filled.

Research Questions

1. Are students' problem-solving behaviors influenced by how problem solving is presented in the lecture when the task is familiar?
2. Are students' problem-solving behaviors influenced by how problem solving is presented in the lecture when the task is unfamiliar/novel?
3. Are there components of a task that can be identified that lead the task to be a problem for students instead of an exercise?

These questions will be answered by analyzing students problem solving behaviors, time on each behavior, as well as detailed descriptions of how they work on problems for both stoichiometry problems and comparing how the students in each course solve the task.

Analysis

To study the problem solving behaviors of the students, the transcribed interviews were coded, first using the seven problem solving behaviors described by Calimsiz (2003) in his thesis, then using a more detailed description of what the students are doing. The seven problem solving behaviors described by Calimsiz are shown in

Table 38.

Table 38 Description of Seven Problem Solving Behaviors

Problem Solving Behavior Number	Calimsiz Description	How used in this study	Example quote
1	Familiarizing themselves with the problem	Reading the problem	"What mass of oxygen needed to completely combust 1 gram of ethanol to produce carbon dioxide and water vapor" –Laura
2	Restating the problem	Restating the goal of the problem	"Question asks the complete combustion of octane give you water and carbon dioxide gases but an incomplete combustion produces ahh water vapor and carbon monoxide. And so, it's asking if it's a combination of combustion, of incomplete combustion, what's the efficiency of, of the combustion of octane"-Bill

Table 38 Description of Seven Problem Solving Behaviors ctd

Problem Solving Behavior Number	Calimsiz Description	How used in this study	Example quote
3	Trying to find a place to start	Trying to find a place to start	"So first we're gonna (sic) set up the stoichiometric equations"-Tom
4	Trying to minimize and if possible, bridge the gap (find a solution), if possible, by purely working forward, if not	Writing equations, doing calculations	"All right so you have ethanol which is C_2H_6O so plus oxygen which yields carbon dioxide, water and I just balance the equation its

	by working forward and working backward at the same time, or by working backwards		2 carbons there, uh you have to balance the equation" – Tahir
5	Consulting various sources	Looking up information in textbooks, online, or notes	"I'm gonna (sic) look up the formula for ethanol. I should know it but I don't. (types on computer) So I put in ethanol chemical formula. And it's, well, I guess I'll use Wikipedia, I don't like to use it but I'm going to."-Dave
6	Modifying or abandoning a particular step or complete route	Changing the route they are using to solve the problem	"So, I uh, (erases) so 1 divided by, it's a 4, (erases), flip those"-CC
7	Evaluating their work	Double checking their work, determining if the work they are doing/did was appropriate, planning their strategy for solving the task	"Now I just like to stare at it a little bit and make sure it's right."-Cole

Since step four of the list of behaviors, working on the problem, is very vague, and part of the project is to look at what exactly the students are doing when they solve these tasks, their problem solving behaviors were coded in more detail, especially the parts that had been coded as a step four previously. Time on each behavior was also recorded for each interview. With these data available, graphs of the movement between problem solving behaviors, time on each behavior, as well as qualitative descriptions of how students solved the problems and any common errors that were made, can be generated and discussed.

In chapter three, the content coverage and time on quantitative and conceptual topics were discussed. As mentioned in that chapter, the data for that analysis came from the lecture capture from the Engineering and Science Majors courses. These same lecture captured video and audio files have been transcribed and are available to look at how problem solving is presented in each course within the context of specific topics.

Description of Problem Solving using Gram-to-Gram Conversion

The set up for this discussion will include a general description of the question presented to the class, followed by a detailed description of how the problem solving process is presented to the students. The example problem used in the Engineering course was the combustion of phosphine to produce tetraphosphorous decoxide. The formula for phosphine and an amount were given and the students were asked to calculate the amount of product that will be produced. The equation was already balanced. The professor starts by identifying what type of question it is, as well as talking about what's given and requested in the question. *"Ok, that's the kind of question. So it's a mass of a chemical, and we want to know the mass of some other chemical that gets formed. Ok, so how do we do that, so that I digesting the problem is the first thing I do."* Next the professor notes that there are chemicals in the equation that won't affect the outcome. *"I look at usually problems are just like this, there are several chemicals, so of which are going to have no influence on our goal of this particular problem."* The next step in the Engineering course is to calculate the molecular weights of the reactant and the product, as they will be needed to calculate the answer. From there, the professor lays out the plan for the dimensional analysis. He talks about the steps of converting the mass of the reactant to moles using the molecular weight calculated before, converting to moles of product using the coefficients from the balanced chemical equations, and finally, converting the moles of product into the mass of product using the molecular weight previously calculated.

This professor is explicitly showing students the step-by-step directions of how to solve this type of problem. This particular example is from early on in the stoichiometry discussion as shown by the fact that the equation was already balanced in the question. In other tasks later in the lecture, the students are expected to balance the equations as part of solving the questions. The description of how gram-to-gram stoichiometry conversions are presented in the Science Majors class is very similar to how it is presented in the Engineering course.

The example used in the Science Majors class involves the combustion of methanol, where the mass of methanol is given and the question asks for the amount of water produced. The equation is balanced as given. The first thing the professor asks to do is identify what type of reaction is being described in the problem. "*We're going to burn methanol in air. If we had to classify this reaction, what type of reaction is that? It's combustion. The product of any CHO compound reaction with oxygen is just* CO_2 *and water."* The professor notes what the question gives and what the question is asking for. *"we're gonna (sic) start with 209 grams of methanol, this is methanol over here. We want to know, if, if, that is, if that is used up, what mass of water is produced?"* The first step the professor suggests is to calculate the molecular weights of methanol and water, as they will be important later. The Science Majors students have experience working with mole-to-mole conversions and molecular weights and so the professor gives the students time to work on the problem by themselves before they work it out as class. "*Again, it's better if you do this before I show you how we get the answer to this. So I'm gonna (sic) walk, I'm gonna (sic) give you about 2 minutes."*

Once the students have worked on the problem by themselves, the professor suggests that they work on the problem as a class. While working through the problem as a class, the professor walks through the expected steps to use for a dimensional analysis procedure. He discusses the steps in much the same way as the Engineering professor, except that he makes explicit the idea that the molecular weights and stoichiometric coefficients are needed to solve the problem, but are not given in the question. *"Those are the three things you need to get from the problem, that aren't told."* This explicit notation of the information that's needed and not given in the question, is the only major difference between the way gram-to-gram stoichiometry is presented in the Engineering and Science Majors classes.

Overall the presentation of how to do gram-to-gram stoichiometry problems is very similar in the Engineering and Science Majors class. There were 12 examples of questions presented in the Science Majors lecture during the stoichiometry portion, and 15 examples in the Engineering course the require dimensional analysis to solve. These examples where shown over several course periods where the topic of stoichiometry was covered. There were no stoichiometry questions discussed in either lecture that could have been solved without using dimensional analysis. With this many dimensional analysis stoichiometry problems, the students in both courses have been trained to recognize a stoichiometry task and to apply a dimensional analysis algorithm to solve the task.

One would expect the students in each course to be able to successfully solve the familiar stoichiometry task, and that they would solve the task in similar ways. The data that follows will test that hypothesis, by analyzing the problem solving behaviors from Table 38, looking at the time spent on each of those tasks, as well as an in depth description of how the students solved the task and any errors they made in the process.

Analysis of Familiar Stoichiometry Task

Problem Solving Behaviors

Below are graphs showing how students moved between the problem solving behaviors over the course of the interview. On the vertical axis are the seven problem solving behaviors from Table 38. The horizontal axis represents time while working on the task during the interview. The number of instances range up to the maximum number of different behaviors students in each course used to solve the task. The number of instances was used to order the behaviors students used during the interview. A different line represents each student's movements. Coding all of the students' comments and work during their interviews and then listing the behaviors in order produced the graphs. If the same behavior occurred multiple times in a row it was only counted once. These graphs are shown in Figure 27 and Figure 28.

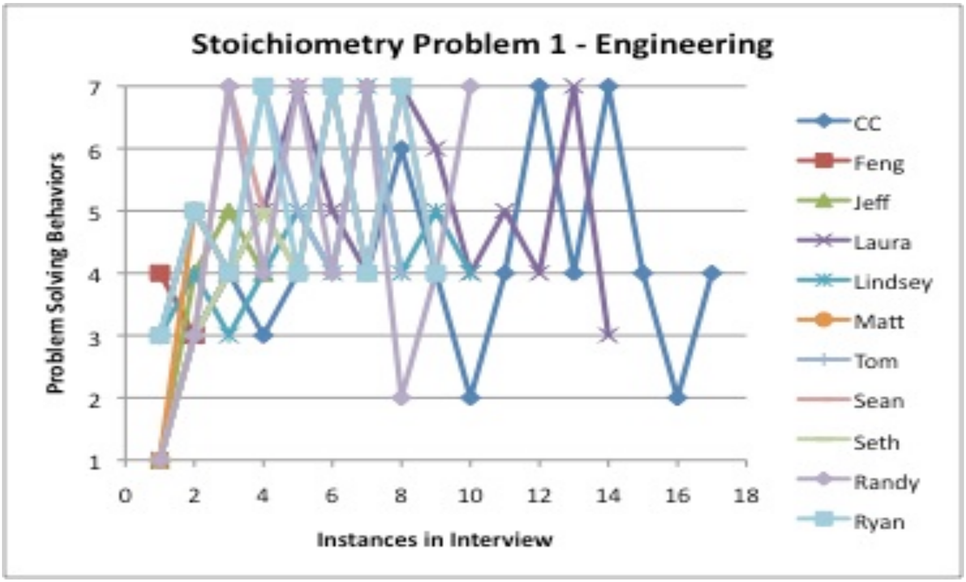


Figure 27 Graph of Movement between Problem Solving Behaviors for Familiar Task - Engineering Students

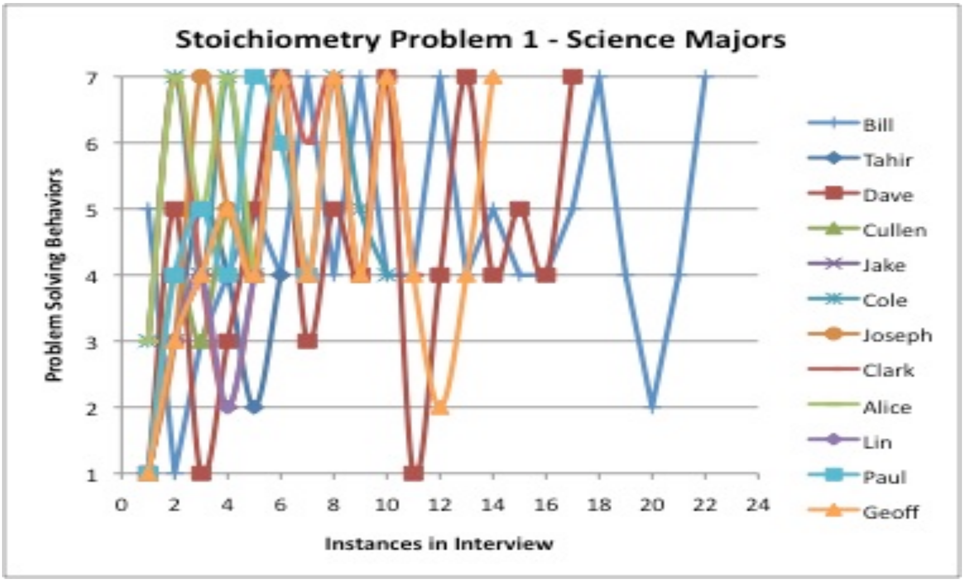


Figure 28 Graph of Movement between Problem Solving Behaviors on Familiar Task- Science Majors Students

In both graphs you can see quite a bit of bouncing between behaviors four and seven. Behavior four is working on the problem, whether that is carrying out a calculation or writing the chemical equations. Behavior seven is where students are evaluating their work, be it, double-checking the balancing of an equation, or determining if the next step their planning is appropriate. In both classes, for this task, the students are doing a lot of checking of their work as they move through the problem. Near the beginning of the interview, students are reading the question and trying to determine how to begin, behaviors one through three. But after about the fourth instance in the Engineering course and the seventh instance in the Science Majors course, the students have determined what to do, and have moved on to the working of the problem. There is some re-reading and restating of the problem, behaviors one and two, in both courses, but it seems to occur more in the Science Majors course. The other major behavior that occurs in both sets of interviews is checking resources, or looking for information, behavior five. This behavior tends to happen at the beginning of the interview, and as will be seen later, is almost exclusively a search for the formula for ethanol. Overall, despite a few differences between the two courses, all the students interviewed, tended to read the question, decide where to start and spend most of their time working on the problem and/or checking their work.

Time on Behaviors

Another way to study how students solved the task of gram-to-gram stoichiometry is to look at the percentage of time each student spent of a particular behavior. Each behavior coded in the interviews also had a time stamp associated, so as to be able to calculate the amount of time spent on the behavior. Times were totaled for each behavior, taking into account multiple instances in a row of the same behavior. These totals were divided by the total amount of time spent working the task to determine the percentage of time spent on each behavior. These graphs are shown in Figure 29 and Figure 30.

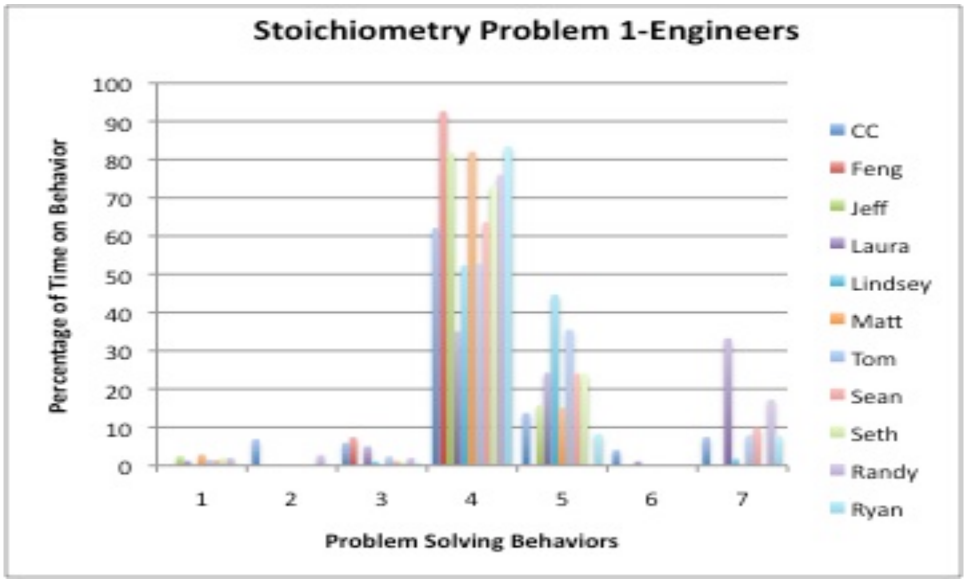


Figure 29 Graph of Percentage of Time on Behavior – Familiar Task - Engineering Students

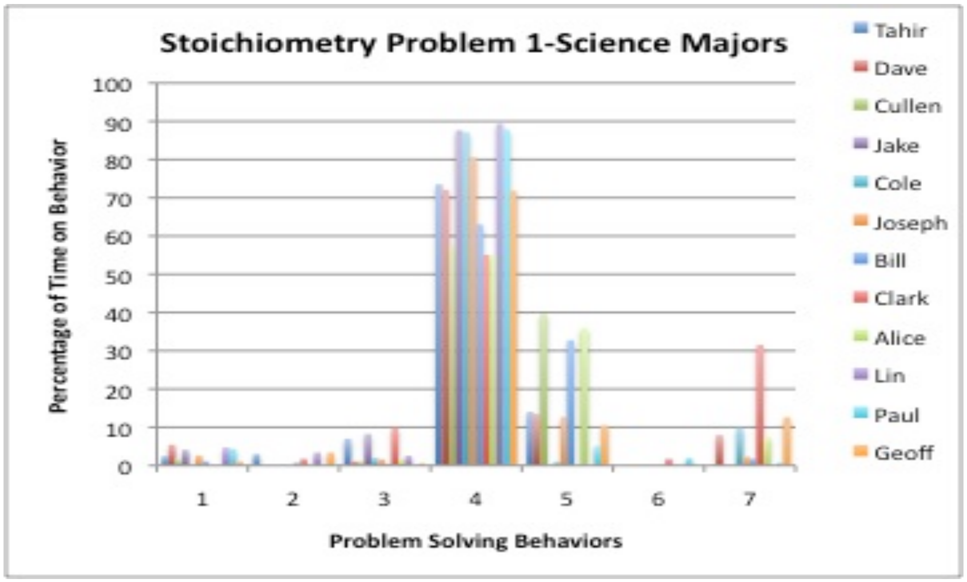


Figure 30 Graph of Percentage of Time on Behavior – Familiar Task - Science Majors Students

The graphs of time on task show similar information, as was shown in the movement between behaviors graphs. The majority of the behaviors that students in both courses spent their time on were working on the problems, looking up information and checking their work, behaviors four, five and seven. Once again, there is not much difference in the relative amounts of time students in both classes spent on these behaviors. The fact that students in both classes spent almost no time, trying to figure out a place to start, behavior 3, indicates that in both classes, students knew how to solve this task, and it was just a question of setting up the proper conversions and running the calculations. This indicates that this task was more of an exercise for the students than a problem.

Another possible indication that this task was an exercise for students was the amount of time spent working on this task. The average amount of time spent on the first stoichiometry problem was calculated for both courses, presented in Table 39. As can be seen in this table, the students are spending less than six minutes on this problem on average.

This relatively small amount of time spent on the familiar task indicates that this task was a relatively straightforward exercise for the students. There is a difference between the amounts of time spent working on this task between the two courses. The Engineering students spent a little over three minutes working on this task, while the Science Majors students spent almost six minutes working on the problem. While the students in each class spent proportionally the same amount of time on each behavior, it appears that the Science Majors spent more time in an absolute sense working on the task.

Table 39 - Average Amount of Time on Familiar Task

Course	Average Time on Familiar Task (min.sec)
Engineering	3.17
Science Majors	5.49

Common Paths to Solving Familiar Task

From the detailed descriptions of how students solved this task, and the detailed descriptions of what students were doing when working on the task specifically, a list of the common behaviors used to solve task one was developed see Table 40. The steps in this list are not necessarily done in a particular order, besides reading the question first.

Table 40 Steps to Solve Familiar Task with Quotes

Step to Solve Familiar Task	Example Quote
Read problem	"Alright, what mass of oxygen is needed to completely combust 1 gram of ethanol to produce carbon dioxide and water vapor"- Jeff
Write chemical equation	"So the reaction would be, ethanol, C_2H_5OH plus O_2 since it's combustion, and that would give us the products, CO_2 and H_2O."- Randy
Look up formula for ethanol	"A: I guess ethanol would be in here. No probably not…do we have a periodic table here somewhere? (Looks through the material) um ok am I allowed to use internet I don't know the.." – Alice "I: That's fine! A: ok all right, its weird. Ok so C_2H_6O"- Alice
Add ethanol to chemical equation	"So I have O_2 plus C_2H_5OH reacts to form carbon dioxide and water vapor, CO_2 plus H_2O" – Dave

Balance chemical equation	"Basically now we have to balance this. There is going to be 2 in front of the carbon, that balance out the carbons on both sides, but there is an odd number of oxygen's, now there is 3 oxygen's on both sides, except now there is 5 oxygen's. So let's make a deal and get rid of this 2 coefficient, I just threw in, so we have 3 atoms to worry about the carbon, hydrogen and oxygen. Initially we have 3 oxygen on the left hand side 6 hydrogen and 2 carbon on the right hand side we have one carbon 3 oxygen and one hydrogen, so trying to get rid of the one; so there is going to be a 6 coefficient on the very least on the right hand side. So we have to recalculate that stuff changes that to 3 times 1 is 3 plus 2, 4 5, oxygen"- Geoff
Calculate molecular weights of oxygen and ethanol	"So, umm, we have about like umm 1 C_2H_5OH umm so that's 1 gram so umm we need to figure out how many moles per gram of this substance we need when with the oxygen. So, umm, 15.999 moles of, err grams of the oxygen (erases) O2, umm, ok, so I think we're going to, so, C2 so that's 24 plus 5 plus 16 plus 1 so that's going to be 22 plus 24, it's going to be 46."-Laura
Use dimensional analysis to convert grams of ethanol to grams of oxygen	"Want to mess with 1 gram of ethanol, (writing) times 1 mole of that over 6, 24, (calculator noises) 46. Which means we're given 0.02167 moles, which I misspelled. 3 to 1 ratio, this is ethanol times 1 mole (writing) Except there's another thing. (calculator noises) So we need 0.02652 O_2 times 32, total mass is 2.08. Tricky, 2.09. (writing) grams O_2"-Jake

As we have seen throughout the analysis of the first stoichiometry task, the steps taken by students in both courses when solving this task were similar. The similarity in how students went about solving this task is not that surprising considering how similar the presentation of gram-to-gram conversion stoichiometry was in both the Engineering and Science Majors courses. Overall, all the students followed the method that had been presented in class, where in they identify what's given and what's needed to answer the question and write and balance a chemical equation. Frequently the students had to look up the formula for ethanol, to varying levels of success. With a formula in hand, they balanced their equations and calculated the necessary molecular weights. The last step was to set up the dimensional analysis to convert the mass of ethanol into the mass of oxygen.

The dimensional analysis step should be explained in more detail. Using Jake, a Science Majors student, as an example. Jake was chosen because of his clear explanation of the steps he took while working on this task. Jake talks about converting the given mass of ethanol to moles using the molecular weight of ethanol, which he calculates. *“1 gram of ethanol, (writing) times 1 mole of that over 6, 24, (calculator noises) 46. Which means we're given 0.02167 moles.”* The next step in the dimensional analysis is to convert to moles of oxygen. While Jake doesn’t explicitly name the compound he is converting to, from his balanced equation shown in Figure 31, as well as the fact that the question asks for the mass of oxygen, one assumes, the ratio he discusses is the one between oxygen and ethanol. “*3 to 1 ratio.*” Another clue to the relationship he discusses comes from his next step, which involves converting the moles he calculated into grams, using the molecular weight of oxygen. *“So we need 0.02652* O_2 *times 32, total mass is 2.08. Tricky, 2.09. (writing) grams* O_2*”.*

Overall, Jake’s work represents how students worked through the dimensional analysis for this task. He converted from grams of ethanol to moles using the molecular weight of ethanol. From there he converted to moles of oxygen using the coefficients from the balanced equation. Finally, he converted to the grams of oxygen using the molecular weight of oxygen.

$3\ O_2 + C_2H_6O \text{ -> } 2\ CO_2 + 3\ H_2O$

Figure 31 Jake's Balanced Equation for Familiar Task

As shown in Table 40 there are similar numbers of students in both the Engineering and Science Majors class using the steps described above to solve task one. The counts are totals of the behaviors exhibited from the interview transcripts and may not take into account students who, read the question for example, but did so silently. As has been seen throughout the analysis of the first task, the behaviors exhibited by the students in both classes are very similar.

Table 40 Counts of Behaviors Evidenced in Interview Familiar Task

Step to Solve Familiar Task	Engineering Counts (N = 11)	Science Majors Counts (N = 12)
Read problem	7	10
Write chemical equation	10	12
Look up formula for ethanol	10	9
Add ethanol to chemical equation	10	12
Balance chemical equation	11	11
Calculate molecular weights of oxygen and ethanol	7	11
Use dimensional analysis to convert grams of ethanol to grams of oxygen	9	10

Common Errors When Solving Familiar Task

There were a total of four common errors made by students in both courses. These errors are shown in Table 41 along with examples of the quotes of students displaying these errors. The errors were mainly related to the formula or molecular weight for the two compounds of interest in the task. If a student used the wrong formula for ethanol and used the wrong molecular weight because they had the wrong formula, they were only marked as using the wrong formula for ethanol, as long as the molecular weight matched the formula they had chosen. However, when one looks at the prevalence of these errors across students in the two courses, the more common errors of the ones listed are in fact using the wrong molecular weight of oxygen and making an error when balancing the chemical equation.

Table 41 Common Errors Familiar Task with Quotes

Common Errors	Example quotes
Wrong formula for ethanol	"I just did the actual structure for the carbon Uh for the molecule so 2C, 2 carbon atoms and then 6 Hydrogen. So C_2H_6"-Clark
Wrong molecular weight for ethanol	"And then, a gram of ethanol times, would be (writing) is 16, so 24, 6, 16, 33, 6. And then, 32 times 33 is 96."- Lindsey
Wrong molecular weight of oxygen	"so you know the molecular weight of oxygen is UH 16, yeah 16 grams"-Tahir
Improper balancing of chemical equation	"So then it just yields, (erases) it asks for carbon dioxide which is CO_2 and water vapor which is H_2O, both in the gas phase. And so you need to know, make sure it's balanced, which it's not, so 2, 3, 3,2, 5,6, just counting up all the molecules. Six and three, three of them, 4 and 3 is 7"-Bill

Table shows that there were a total of five students, three Engineers and two Science Majors, who used the wrong molecular weight for oxygen and two Engineering and three Science Majors who could not properly balance the chemical equation. Two students used the wrong formula for ethanol and two students used the wrong molecular formula. The wrong formula and molecular weight used were equally divided between the courses. So overall students had more difficulties, not with the formula for ethanol but with the molecular weight of oxygen, a compound they've seen frequently, and with the act of balancing the equation. It's possible that students didn't take in account the oxygen in ethanol when they balanced the equation, leading to the error.

Table 42 Counts of Errors Exhibited in Interview Familiar Task

Common Errors	Engineering Count (N = 11)	Science Majors Count (N = 12)
Wrong formula for ethanol	1	1
Wrong molecular weight for ethanol	1	1
Wrong molecular weight of oxygen	3	2
Improper balancing of chemical equation	2	3

Students' performance on this task, in terms of completion, can be grouped into three categories, the students who got the question correct, wrong, and who gave up on the task. Table 43 shows the breakdown of students in each class, as well as overall, into these three categories. From this chart we can see that eight of the 23 students (35%) got the task correct, while 14 of the 23 students (61%) got the question wrong, and one student in the Engineering course gave up on the task. Along with the common errors discussed in Tables 41 and 42 above, calculation errors may have led to 61% of the students in the interviews coming up with the wrong answer to this task. It should be noted that the number of students getting the question wrong is equally divided between the two courses, while the number of Science Majors getting the question correct outnumbers the Engineering students. It is important to remember that whether or not the students got the correct answer to the task or not, they are all mainly following the same set of steps to solve the task. Finally, the steps the students' take are very similar to how they are instructed in both classes to go about solving gram-to-gram conversion tasks.

Table 43 Categorization of Student Performance on Familiar Task

Group	Right	Wrong	Gave Up
Engineering (N = 11)	3	7	1
Science Majors (N = 12)	5	7	0
Overall (N = 23)	8	14	1

Findings for Familiar Task

Because of similarities between the way gram-to-gram stoichiometry problems are presented in the two courses, and the fact that the stoichiometry problems presented in both classes were almost exclusively solved using the dimensional analysis method, we are training students to recognize stoichiometry problems and apply this dimensional analysis algorithm to solve those tasks. Students are practicing using this process all the time. So when they are presented with a familiar task that can be solved using dimensional analysis, they tend to use what they've been presented with in class.

Despite using the dimensional analysis process that they are very familiar with, over 60% of the students got the familiar task wrong. When looking at the error students made in Table 42 many of them could be explained by the fact that the students were not paying attention. The most likely errors for both courses were using the wrong molecular weight for oxygen, despite using the proper formula, and errors with misbalanced chemical equations. These types of errors likely indicate that students are so used to using the dimensional analysis process to solve stoichiometry tasks that they stop paying attention to the details of solving the task, like ensuring that the equations are balanced or that they've taken into account the diatomic nature of oxygen when calculating the molecular weight of oxygen.

Analysis of the Unfamiliar Stoichiometry Task

Problem Solving Behaviors

The problem solving behaviors graphs for the unfamiliar task for the two courses are shown in Figure 32 and Figure 33.

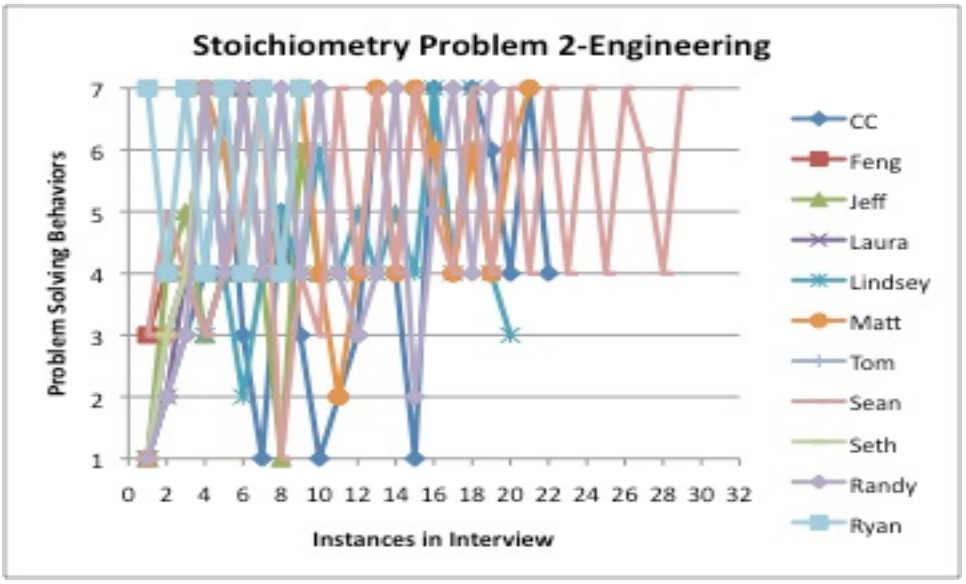


Figure 32 Graph of Movement Between Problem Solving Behaviors – Unfamiliar Task-Engineering Students

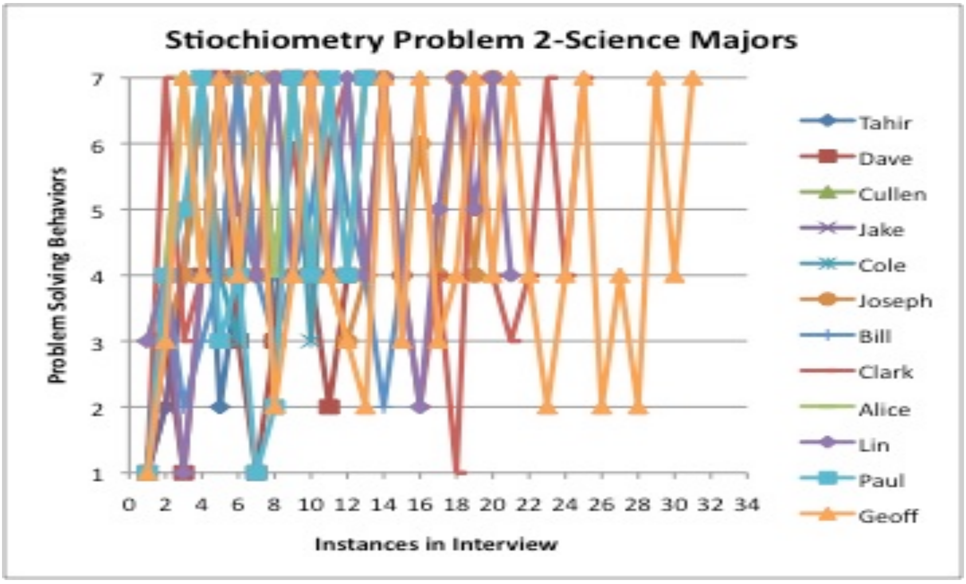


Figure 33 Graph of Movement Between Problem Solving Behaviors – Unfamiliar Task - Science Majors Students

When looking at the movement between problem solving behaviors graphs for the two courses, we some similarities to the familiar task in the students movements between behaviors. Once, again, there is a lot of movement between working on the problem and checking their work, between behaviors four and seven, in both courses. As compared to the other stoichiometry task, there is more re-reading/restating of the problem when students are working on this problem as evidenced by the increased number of students in behaviors one and two, particularly later in the task. This re-reading/restating the problem is more pronounced in the Science Majors class. This may lead students to restart the problem when they are working, re-reading/restating the problem as they move to the second calculation.

There is also an increase in the number of students displaying behavior three, trying to find a place to start. There is more movement to behavior five, looking up information, particularly in the Engineering course. The students are spending more time on thinking about how to start the problem and looking up information, usually about the efficiency of the combustion of octane. This indicates that students' do not immediately know how to answer this question. Another source of information that indicates that students are less sure what to do when trying to solve task two is the fact that it takes them much longer to solve this task as compared to task one. In Table 41, the average times spent by students in both courses on task two. One of the first things we notice is that the average amount of time is much longer, almost four times longer for the Engineering course and almost twice as long for the Science Majors class, for both courses as compared to the average times from the familiar task, see Table 44. This lends more evidence to the idea that students are more uncertain about how to solve this problem as compared to relatively familiar task one. Furthermore, the average times for the two courses are much closer for this task as compared with the previous stoichiometry task, but the Science Majors students still spent more time solving the task. This may mean that the students in both classes struggled with what to do when solving this task. The final indication that students struggled with this task is that only one student out of 23 was able to successfully solve this task. The rest of the students either gave up or got the question wrong as seen in Table 45 In fact, 14 of the 23 students gave up on this task, evenly divided between the two courses. This same even division of students can be seen in the percentage of students who got the answer to the task wrong.

Table 41 Average Amount of Time on Unfamiliar Task

Course	Average Time on Unfamiliar Task (min.sec)
Engineering	11.43
Science Majors	13.48

Table 45 Categorization of Student Performance on Unfamiliar Task

Group	Right	Wrong	Gave Up
Engineering (N = 11)	0	4	7
Science Majors (N = 12)	1	4	7
Overall (N = 23)	1	8	14

Time on Task

The percentage of time students in each course spent on each of the seven problem solving behaviors is shown in Figures 34 and 35.

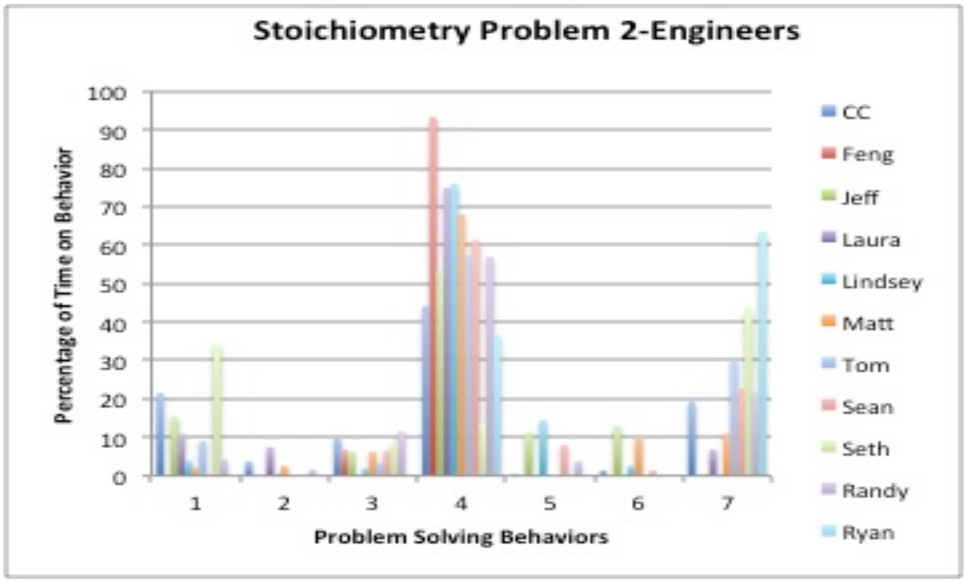


Figure 34 Percentage of Time on Behaviors – Unfamiliar Task – Engineering Students

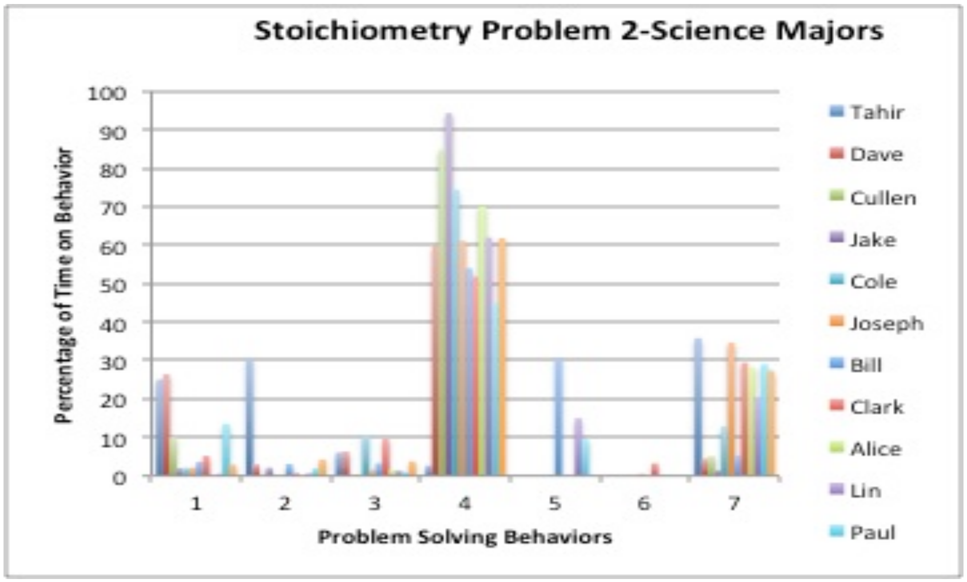


Figure 35 Percentage of Time on Behavior - Unfamiliar Task - Science Majors Students

When looking at the percentage of time spent on each problem solving behavior by students in both courses, overall we see that there is a great deal of time spent working on the task, and on evaluating and checking their work, behaviors four and seven. This was also seen in the first stoichiometry task. There is more restating of the problem and more reading of the problem, behaviors one and two, by students in both courses. There is more time being spent on behavior five, looking up information in the Science Majors course. The amount of time on double-checking the reasonableness of work is higher in the Engineering course. This may be due to how problem solving has been presented over the course of the semester in the Engineering course. There is more decision making expected, of what the right next step is, what information is needed, as compared to the Science Majors classroom, where the pragmatic desire to be able to produce the answer is more common in the instruction.

Common Paths to Solving Unfamiliar Task

The first stoichiometry task was relatively familiar to students and they knew what to do to solve the task. As an analogy to the signal to noise ratio, the ratio was high for the first stoichiometry task, because students all solve the task in a similar fashion, and had similar mistakes that led them to the incorrect answer. In contrast, the second task has a very low signal to noise ratio, in that the students took lots of different paths while trying to solve this task. Five students were selected to show the breadth of responses to the unfamiliar stoichiometry task. These students were selected to represent both courses, and to represent one of five relatively common behavior patterns within the larger data set. The behaviors represented were students who used one chemical reaction equation and gave up, used one chemical reaction equation and got the answer wrong, used two chemical reaction equations and gave up early, used two chemical reaction equations and gave up fairly late in the process and finally, someone who set up a relationship between the amount of carbon dioxide and carbon monoxide but still gave up. The last representative sample, writing a relationship between the carbon dioxide and carbon monoxide was chosen because there was a subset of six students, who had made it most of the way through the correct algebraic process to solve the task. The work on these students will be discussed in detail later in the chapter. Writing one chemical equation means that the student tried to have the water, carbon dioxide and carbon monoxide in one reaction. Students who wrote one chemical equation tended to have trouble with assigning the stoichiometric coefficients for the carbon containing compounds. Writing two chemical equations meant that the students wrote separate reaction for the complete and the incomplete combustion.

Bill represents students who calculated a percentage based on one chemical equation and got the answer wrong. Seth gave up on the task early on but had written two chemical equations. He was an example of someone who was working in the right direction but gave up. Cullen used one chemical equation but gave up on the task. Jeff correctly wrote two for the complete and incomplete combustion, but who got the question wrong. Joseph represented the students who noted a relationship between the amounts of the two carbon containing products, but ultimately gave up on the task. The discussion of Joseph's work lead into a discussion of the six students who wrote about a relationship between the amounts of the carbon dioxide and carbon monoxide. The number of students who wrote mathematical equations about the relationships in the task was evenly divided between the courses.

Bill – Wrong Answer with One Equation

Bill is a Science Major student and who had gotten 50% of the four-stoichiometry tasks correct on the Toledo Test at the beginning of the semester. Bill starts out as most of the students did by planning to convert the given volume of octane into a unit he's familiar with, grams, and notes that the total mass of the three products are given. He starts working on the task, as would be expected if he were using the dimensional analysis procedure to solve the task. He writes and balances two chemical equations, one for the complete combustion and one for the incomplete combustion. Bill's next step, to add the two equations together, may seem unusual, but he has a reason for this.

"*I'm not 100% sure, but I'm thinking about adding the complete and incomplete cause I think, you're gonna (sic) start off with the same amount of octane and* *oxygen, the only difference you're gonna (sic) get is the percentage of the difference of, of product…. So I was, so I'm just thinking that since the reactants for both complete and incomplete combustion are the same and the water's the same as a product, to take the complete and the incomplete equations and add them together."*

Bill's reasoning is based on the fact that the total mass of the products was given for the three products and therefore the three products must be in one equation. Several other students also either combined two equations together, or only wrote one equation for this reaction, possibly because of the way the masses were given for the three compounds. His equations are shown in Figure 36.

$C_8H_{18}(l) + 12.5\ O_2\ (g) \rightarrow 9\ H_2O + 8\ CO_2$
complete: $2\ C_8H_{18}(l) + 25\ O_2\ (g) \rightarrow 18\ H_2O\ (g) + 16\ CO_2\ (g)$
incomplete: $C_8H_{18}(l) + 5\ O_2\ (g) \rightarrow 9\ H_2O + 8\ CO\ (g)$
combine: $3\ C_8H_{18}(l) + 30\ O_2\ (g) \rightarrow 27\ H_2O + 16\ CO_2\ (g) + 8\ CO\ (g)$

Figure 36 Bill's chemical equations

The next steps for Bill are to find the mole ratios for the reaction and convert the given gallons of octane into kilograms using the given density. These steps are in keeping with using a dimensional analysis procedure to solve this task. Bill then begins to talk about what he expects the answer to be.

"And so that's, the only thing I'm concerned about is that the complete combustion yields 100% CO_2 and H_2O and that incomplete yields 100% H_2O and CO and so if I add them it would be roughly 50/50% and so I don't think that would be an accurate way to calculate the efficiency of the process."

Bill is trying to make a guess as to what the percent efficiency for the reaction will be, but he seems unsure of how to calculate this value. He searches for more information about octane combustion and carbon monoxide in his textbook. Bill doesn't appear to find anything useful in the textbook to help him solve this task. The interviewer then asks him what information he's looking to find in the textbook. Bill mentions trying to find information about the efficiency of combustion of octane.

"I'm hoping to find something on the combustion of octane to figure out a percent efficiency for carbon dioxide and carbon monoxide. That way I could apply that to the mole ratio to figure out."

Bill appears to be confused about what he's looking for in this task. The task asks a student to solve for the efficiency of the process by calculating the faction of octane converted to carbon dioxide. It appears that Bill thinks he needs to find the efficiency of the combustion in general before he can solve this task. This indicates that Bill doesn't realize that the question is asking him to calculate the efficiency of the reaction.

Once Bill has made it clear that he's looking for information about the general efficiency of the reaction, he continues to work on the task using a dimensional analysis approach. He converts the gallons given into kilograms using the density given in the question and converts the kilograms into grams. He then converts the mass of octane into moles using the molecular weight he calculated. He converts the moles of octane into moles of carbon dioxide using the coefficients from the combined equation. From there he converts to mass of carbon dioxide and into kilograms. Bill's last step is to take the mass of carbon dioxide produced in kilograms, divide it by the total mass of products given in the question, and multiply by 100 to turn the answer into a percent.

Bill disregarded the idea that he needed to compare the amount of carbon dioxide produced with the amount of octane available and instead compared the amount of carbon dioxide produced with the total mass of products. Other than when he combined the two equations together, Bill followed the steps a student would be expected to use to solve this task if they used the dimensional analysis method. While some of the steps Bill used were necessary to solve the problem successfully, writing and balancing two chemical equations, and converting the gallons given to kilograms of octane, overall his inability to see the task as not requiring dimensional analysis and his confusion about what efficiency means in the context of the question, led to his incorrect answer on this task.

Jeff – Wrong Answer with Two Equations

Jeff is an Engineering student who did not get any of the stoichiometry questions correct on the Toledo Test taken at the beginning of the semester. Jeff also starts out by writing two balanced chemical equations for the reaction and converting the given gallons of octane to kilograms using the density given in the question. He is undecided about whether or not to combine the equations together. *"but the mass produced is. See that's where I'd get confused cause that says that's produced with all three of these. But all of three of these isn't in the same equation."* It appears that Jeff is also struggling with whether or not to combine the equations together, as Bill was. Ultimately he decides to use the two equations, but he admits that more explicit instructions would have been helpful. *"Umm, whether or not it had been a complete or incomplete combustion would have been helpful."* The possibility that the reaction could be a combination of complete and incomplete combustion does not seem to occur to Jeff. The equations Jeff used to solve the unfamiliar task are given in Figure 37.

$C_8H_{18} + 25/2\ O_2 \text{ -> } 8\ CO_2 + 9\ H_2O$
$C_8H_{18} + 13\ O_2 \text{ -> } 8\ CO + 9\ H_2O$

Figure 37 Jeff's chemical equations

Jeff continues working on the task using the dimensional analysis process He continues to calculate the number of moles of octane produced from the kilograms of octane calculated earlier. He converts to the moles of octane to the moles of carbon dioxide using the coefficients from the chemical equation for the complete combustion. From there he converts into mass of carbon dioxide in kilograms, as would be expected when using the dimensional analysis approach to solving the task.

Jeff notes that the value seems to be too large, but that it is theoretical yield of carbon dioxide for the reaction. "*Number seems too big. ... So that's hypothetically is produced using the equation.*" From his written work, it appears that Jeff has given up on using his theoretical yield of carbon dioxide he previously calculated. He instead sets up a ratio between the total mass of the products given in the equation and the mole ratios for water and carbon monoxide versus carbon dioxide see Figure 38. The 17 in the equation is the sum of the coefficients for water and carbon monoxide in the incomplete chemical reaction, while the 8 is the coefficient for carbon dioxide in the complete chemical reaction equation.

$\boxed{11.53\text{kg} / X = 17/8}$

Figure 38 Jeff's equation for calculating efficiency

"*In the actual equation it's, well 11.53 total, well I have to find out how much is carbon dioxide.*" Jeff seems to have abandoned his theoretical yield calculation from before is simply setting up a ratio with the total mass of products. He solves this proportion for X and multiplies by 100 to get an answer to the unfamiliar task.

Jeff was also using the dimensional analysis process to solve this task, as Bill had done. He saw the total mass of the three products being given as signal to write one equation for the chemical reaction, but he chose to use the two equations instead. While Jeff used the dimensional analysis approach to solving this task, his final answer came by setting up a proportion between the total mass of the products given and the combined coefficients for the incomplete reaction products and the coefficient for the carbon dioxide produced. While he used a relationship between the complete and incomplete products, he was ultimately unsuccessful in solving this task correctly.

Cullen – Gave Up with One Equation

Cullen is a Science Majors student who got 25% of the stoichiometry questions correct on the Toledo Test at the beginning of the semester. Cullen starts off by converting the gallons given to kilograms using the given density. His next step is to write and balance one equation with the three products, carbon dioxide, carbon monoxide, and water. *"And then it's the balanced equation, is what I'd want to figure out next. So,* C_8H_{18} *plus* O_2 *goes to* H_2O *and* CO_2*, but sometimes CO."* It seems that he was misled along the way while reading the question. However, after trying to balance the single equation, he decides to write two equations instead. "*Alright, I don't think we can do that (erases)." "So, set up two equations."* Because Cullen was having a hard time balancing the one equation, he moved to two separate equations for the complete and incomplete combustion. The equations Cullen wrote are shown in Figure 39.

$2\ C_8H_{18} + 25\ O_2 \rightarrow 18\ H_2O + 16\ CO_2$
$C_8H_{18} +\ O_2 \rightarrow 9\ H_2O + 8\ CO$
$3\ C_8H_{18} +\ 26\ O_2 \rightarrow 27\ H_2O + 16\ CO_2 + 8\ CO$

Figure 39 Cullen's chemical equations

Cullen's next step, in keeping with using a dimensional analysis procedure, was to convert the mass of octane calculated earlier into moles. After looking at the question more closely, Cullen decides to combine his equations for the complete and incomplete combustion together into one equation, seen in the third line of Figure 39. "*Well, could I add these together? I believe. You get the amount of product here."* Cullen appears to be using the same logic as Bill for combining his equations together, namely that the masses of all three products were given in the question. After combining the two equations together, Cullen discussed planning to calculate the moles of octane and then convert them to the moles of product formed using the coefficients from the balanced chemical equation. Cullen follows through on this plan by converting the mass of octane to moles.

He then calculates the number of moles of octane, water, carbon dioxide and carbon monoxide. *"So then we add that many kg for 2650 grams. Which gives us 23.2 moles divided by 3, and so there's 208.88 moles of water. 22.2 divided by 3 times 16, gives us 123.73 moles of carbon dioxide. 23.2 times 8 gives us 61.9 moles of CO."* It appears that Cullen felt that it was important to have the amounts of all three products, perhaps because the total mass of the three products was given in the question.

At this point, the interviewer asks Cullen what concepts he's using to solve the task and he talks about not being sure if he can combine the two equations for complete and incomplete combustion into one. "*I've never seen a problem like this so I wasn't sure how to approach it, but I really don't know. I wasn't completely sure you could add equations like that."* Cullen's uncertainty with how to combine the equations together, and even if he should do that, led him to give up on this task and move on.

Overall, Cullen was on the right track for wanting to calculate the masses of the three products, but he didn't need to combine the two equations together. And he mentions that he's never seen a task like this before leading credence to the idea that this task was a problem for him. It appears that having the total mass given for the three products is leading many students to a least think about combining the two equations together. Overall however, Cullen tried to use the dimensional analysis approach to solving this task, instead of a more algebraic approach.

Seth – Gave Up Early with Two Equations

Seth is an Engineering student who got one quarter of the stoichiometry questions on the Toledo Test correct at the beginning of the semester. Seth's plan also appears to be based on the dimensional analysis procedure. His plan is to convert the gallons into a unit he's familiar with. He also talks right away about being unsure of how many equations to write. *"I'm not exactly sure what equation I should be writing cause there's the complete combustion, there's also the incomplete combustion."* It seems that when students' start off in this dimensional analysis mindset, the fact that two reactions are going on at once is confusing. The students, and Seth in particular, seem to be unsure of how to proceed with their usual problem solving procedure of writing a balanced chemical equation. Seth decides to write two equations. The equations Seth writes can be seen in Figure 40.

$$C_8H_{18} + O_2 \text{ -> } H_2O + CO_2$$

$$C_8H_{18} + O_2 \text{ -> } H_2O + CO$$

Figure 40 Seth's chemical equations

He seems unsure of what to do next. He mentions that he'd have to find the mass of carbon dioxide produced, but he can't think of an equation to get him there. *"Oh man,* *(pause) yeah, I, I mean I know that at some point I know I have to convert err find out which mass of the* CO_2 *would actually, would come through the 11.53 kg but I don't know how to set up the equation that would get me into that."* Seth seems to have gotten stuck on how to continue. It seems that Seth was unable to find a way to calculate the amount of carbon dioxide produced.

Seth's behavior is indicative of several students who were unable to move on past the writing of an equation. Some of the students couldn't even balance the equation or equations they had written and were at a loss as to what to do next. Seth is unable to even come close to setting up equations relating the masses of the carbon dioxide and carbon monoxide to the total mass of the products formed.

Joseph – Gave Up used Equations as Relationship

Joseph is a Science Majors student who got half of the stoichiometry questions on the Toledo Test correct at the beginning of the semester. Overall, Joseph came the closest of the five students highlighted to solving this task correctly. While he started out using the same dimensional analysis approach shown by the other students, he ended up writing an equation relating the percentage of complete and incomplete combustion reactions to each other.

Joseph started out with the usual steps in the dimensional analysis process to answer this task, namely writing and balancing two equations. His chemical equations are shown in Figure 41. Joseph converts gallons of octane into kilograms using the density given in the question and converts the mass to moles using the molecular weight.

$2\ C_8H_{18} + 25\ O_2 \text{ -> } 16\ CO_2 + 18\ H_2O$

$2\ C_8H_{18} + 17\ O_2 \text{ -> } 16\ CO + 18\ H_2O$

Figure 41 Joseph's chemical equations

He plans on calculating the masses of the products based on the amount of octane available and comparing it to the total mass of the products given. *"So at this point I think I would take how many moles of octane I have and then plug it through each of my equations to see what the weight or what the final product weight would be from those two and then compare that to the weight that I am given and that is 11.53 kilograms."* Joseph's plan to solve for the amounts of all three products and compare it to the total mass of products is similar to some of the other students' plans. However, while he's working this through, he realizes that the needs to think about the amount of oxygen available. *"the weight of the, the reactants is equal to the weight of the products so I could just take my weight of octane and add it to the weight of the oxygen that I know, so the octane times 23.199 times 2 scratch that I just need to find the weight of the oxygen."* Joseph used the idea of the conservation of mass to decide he needed to calculate the amount of oxygen available. While this question does rely on the idea of the conservation of mass, the amount of oxygen is not a key to this task. Joseph continues to move down this path of focusing on the amount of oxygen needed, by calculating the mass of oxygen needed for the complete combustion reaction. This value is added to the mass of octane available and compared to the total mass of products from the question. When comparing these values, Joseph notes a difference between the amount he calculated and the mass of all three products.

"The uh it seems like the weight is not adding up exactly. ... Um it gives me the weight of the, combined weight it tells me the weight of combined as the carbon monoxide, carbon monoxide and the water produced by the combustion sorry um its 11.53 kilograms so I know that the weight of the initial amount of octane plus the oxygen added should be equal that depending on which reaction, like how much of each reaction was used."

Joseph was using the law of conservation of mass to justify his answers when he noticed that there were significant differences in the amount of oxygen needed in the complete and incomplete combustion equations.

Joseph suggests that the mass of products given in the question would be somewhere between the amounts produced by the two reactions. From there he suggests doing algebra to find the percentage of complete versus incomplete reaction.

"Logically it will be 11.53 kilograms it will have to fall somewhere in between those two complete reaction things so I suppose I would see something like a number line like a which end it was closer to and I suppose I would do some sort of algebra stuff and would check to get a percentage of complete reaction versus incomplete reaction."

Joseph makes a leap in thought that most of the other students did not, to thinking that there was a way to use algebra and a relationship between the complete and incomplete reaction to solve this task. Five other students also used this idea to try and solve this task; their work will be discussed below.

While Joseph is using an appropriate idea, he's applying it to the incorrect compound. He continues his work by calculating the mass of oxygen available and converts it to moles. He then calculates the amount of oxygen needed for the complete and incomplete reaction using the balanced chemical equations shown in Figure 41.

Joseph's next step is to set up an equation relating the percentage of the reaction that is complete versus incomplete. *"So ok well I know that 277.510 should equal 289.99 x plus 197.19 y, x being the percentage that of the reaction that was complete and y being the percentage of the reaction that was incomplete so now I have this kind of 2 variable equation."* This equation is important because it shows the relationship between the complete and incomplete reactions. However, Joseph wrote his equation based on the amount of oxygen instead of octane. He may have done this because he thought of this task as a limiting reactant problem, where oxygen was the limited reactant. His knowledge of combustion reactions and how one gets carbon monoxide as a product may have led to the decision to base this equation on oxygen. Joseph rearranges the equation to solve for y, but is unable to solve the equation. *"Yeah, I changed the equation so I just did the algebra to make it just so y=-1.4706x+1.4073 and I am, I usually get stuck at this point because I am not quite sure where to go from that equation or what use to have that equation."* Rearranging the equation does not help Joseph solve this task. He doesn't see that this equation would be solvable if he related the amounts of carbon dioxide and carbon monoxide produced to each other. In fact he gives up on the question because he feels that this was an unproductive path to take. *"um yeah the more I think about it I think I might have done a unconstructive way to go about solving a problem um well no because I know how much, if I were to graph that line I would know that."* Because Joseph can't solve the task, he feels that he must have made a major error in thinking about the question in this way. In fact, he was very far along the path to successfully solving this task. It appears that when faced with an unfamiliar task, Joseph feels that if he can't get an answer, than the method being used must be incorrect. In fact, that may not be the case.

While Joseph correctly decided to write a mathematical equation relating the amount of complete and incomplete combustion that occurred when the octane was burned, the exact method he used was flawed. The question asked for the fraction of octane converted to carbon dioxide, indicating that calculating the amount of carbon dioxide and carbon monoxide produced should be based on the amount of octane available. In fact, only six students, three from each class, were able to recognize that writing a mathematical equation relating the amount of carbon dioxide and carbon monoxide produced was the key to solving this task. However, within this group of six students, only one successfully solved the equation and answered the question correctly. A description of what these six students did while solving this task will be addressed in the next section.

Relationships using Balanced Equations

The same number of students in each course wrote a mathematical relationship between the amount of carbon dioxide and carbon monoxide produced. There were however, students who did not write mathematical equations, who still talked about ratios while working on the unfamiliar task. The number of times the word ratio was used while solving task two, was measured for both courses in order to get a measure of how students viewed relationships even if they didn't get to the point of using a relationship between the carbon dioxide and carbon monoxide. These uses of the word ratio were categorized into being used as part of an algorithm, for example as part of the dimensional analysis process, or as part of a relationship, for example as part of a mathematical equation relating amounts of CO and CO_2. For the most part, the word ratio was used as part of the phrase mole ratio. The number of students using the word ratio in each category is presented in Table 46 along with the total number of uses of ratio in each course during the task two interviews.

Table 46 Total Counts of use of ratio as algorithm or relationship

Course	Ratio as algorithm People (total uses)	Ratio as relationship People (total uses)
Engineering	4 (6)	2 (10)
Science Majors	3 (8)	2 (5)

Overall, the more interesting differences in the use of the word ratio during the task two interviews occur when the word is used in terms of relationships. While there are only two students who actively mention the word ratio as part of a relationship during their interviews in both courses, the number of instances of the word ratio related to relationships was twice as high in the Engineering course. This is due to one of the students in particular talking about the same topic using that word multiple times. When looking at the use of the word ratio as part of an algorithm, there was almost the same number of students in both courses using the word and almost the same number of instances in both courses. Overall, this suggests that despite the fact that the Engineering students have been in a lecture where the relationships between compounds were emphasized, they do not seem to be as aware of how to use that information while working on a novel task. The Engineering students tend to use the word ratio in the same way as the Science Majors students who have not had their problem solving discussions emphasize the relationship between compounds.

To further investigate the influence of how balanced equations are presented during problem solving in the two lectures, the six students who wrote mathematical equations relating the amount of carbon dioxide and carbon monoxide interviews were analyzed to determine how they developed their equations and whether or not they could successfully solve the equations they had written.

Joseph tried to calculate the percentage of the combustion that was complete and incomplete based on the amount of available oxygen to solve task two. He was not able to solve the equation he had written and so gave up on task two. The quotes to support this analysis are presented above. Jake, another student in the Science Majors course, used one equation to try and answer the question. *"Hmm,* C_8H_{18} *plus* O_2 *gives me* CO_2 *plus* H_2O *plus*

CO." He then accounted for the amount of water that would be produced in this reaction. *"that's how much octane we have, so 12 times 8, plus 18, which would be 2.650 divided by, … 114, (writing) K, so there's that many moles. (writing) Gives me, ok, then* H_2O*, that means how many moles of* H_2O *can I have? One of these gives me 9 of those, I'm gonna try dividing that by 9, I have 209 (writing) moles of* H_2O*."* Jake calculates the amount of oxygen available in moles. *" Given that, given 2.65 kg, that's wrong. (calculator noises) 5000 divided by 32, that's where I got the 276."* He actually had 277.5 moles oxygen on the page as a note. Jake then subtracts the moles of water produced from the moles of oxygen available to get moles of oxygen left over. *"so that's how many moles of oxygen we have, so that means we have 277 (calculator noises) 209.5 which gives me 67.75 moles of oxygen left."* He then writes two equations, one relating the amount of octane to the moles of carbon in carbon dioxide and carbon monoxide, and the amount of oxygen left to the moles of oxygen in carbon dioxide and carbon monoxide. "*We have 23.25 equals X plus Y we have 67.75 equals 2X plus Y."* However, when he goes to solve the equations, he is unable to get to a satisfactory answer. "*That should work, and then so then X equals 2,3, and Y 67.75, equals, (writing) 10.25 minus Y plus 1, so then we have 67.75, times 3.25, 6.5 minus 5, one the other side. Y has to equal a negative number."*

When Lin, the last Science Major student, starts to work, he writes two equations. *"Ok! So we got couple of equations I think I complete will be the first one and the second one will be because it is complete combustion yeah! So if there is not enough oxygen there would be the…(writes, balancing equation on worksheet)."* He sets up equations that relate the amount of CO and CO_2 produced to the amount of octane available converts everything to masses and includes the mas of water produced. This whole equation is set equal to the total mass of products mentioned in the question.

“um we can find the mole of the octane, I don’t know how to pronounce it so I will say it that way and Uh (calculates) it will be the 18.43 uh each mole and well so we can assume that about the each mole of the C_8H_{18} *reacts with the enough oxygen, so complete react so, now solve this as X, so this was 8X, so this will be 9X and the total will be 18.01, so the sum of them will be larger so we can no that it’s a 9. 18.43-X and 8(18.43-X) and then I think we should do is we can add them up and at least we know the mass and calculate and so..ok 8X multiplied by the* CO_2 *, that is 44, so put them together it will give 9, 18, 18 plus uh plus 8 ok its equal to the 11530 grams.”*

Lin’s equation relates the amount of carbon dioxide to X and the amount of carbon monoxide to 23.25 minus X, where 23.25 is the moles of octane. By relating the mass of carbon dioxide to the amount of octane, Lin is able to solve this equation for X and calculate the amount of carbon dioxide produced. *“Now its right, solve the X 19.9 8 Hm ok so its 19.98 mole.”* The last step is to divide the amount of carbon dioxide produced by the amount of octane available and multiplying by 100. “*That’s 19.9 divided by total it’s the answer. (worksheet 86%).”* Lin was in fact the only person who successfully used the mathematical equation to solve the task correctly.

Feng’s data will come from the written artifact from his interview because he didn’t talk all that much as he was working. After writing two balanced equations, he converting the gallons of octane into moles. *“*$m_{C8H18} = 1.000\ gal * (2.650 kg/gal) = 2650\ kg = 2.65 \times 10^6 g$ $n_{C8H18} = [m_{C8H18} / (12 g/mol + 1 g/mol * 18)] = 2.65 \times 10^6 g / 114\ g/mol = 23245.6\ mol$*”*

Feng then wrote three factors relating the masses of carbon dioxide and carbon monoxide to the mass of water produced. *“X mol* H_2O *16/18 X mol* CO_2 *16/18 X mol CO”* Where 16 is the coefficient in front of CO or CO_2 and 18 is the coefficient in front of H_2O. Feng sets up an equation relating these factors with X’s in them to the molecular weights of the various compounds and sets them equal to the total mass of the products. *“X*(1*2*16) + 16/18 X *(12 + 16*2) + 16/18 X * (12 + 16) = 11530”* He then solves for X and calculates the number of moles of CO_2. “*X = 140* n_{CO2} *= 124.44 mol”* Finally, Feng calculates the fraction of octane converted to carbon dioxide. *“(124.44 mole* CO_2 */ 23245.64 mol) * 100% = 0.5%”*. Feng’s issues stem from a mathematical error related to converting from gallons to kilograms of octane as well as relating the masses of carbon dioxide and monoxide to the amount of water produced instead of to the amount of octane available.

Randy, another Engineering student, works in a similar fashion to prior students. He knows he needs to relate everything in his equation to X, which is the amount of octane needed to produce carbon dioxide, but he blanks out on how to do that.

“We can set this up like a mathematical equation. Where, where X would be the amount of, would be the amount of octane that reacted to form CO_2 *and Y could be the amount of octane that reacted to form carbon monoxide. Let’s try that. So, hmm, ahh. 2 times X plus, I’ve got to relate everything to X so. Goodness I’m drawing a plan here.”*

Randy, ultimately gives up on working on this task.

Finally, Matt comes up with an incorrect answer to this task, despite setting up a mathematical expression. His main problem is that he picks what appear to be arbitrary values for the ratio of carbon dioxide to carbon monoxide produced instead of relating the amount produced to the amount of octane available. “*Yeah there’s two parts of CO plus 3* *parts of* CO_2*, so that'll give you 5 parts of everything. So if 3 parts of that is converted to* CO_2 *that's about 60%, about 60% of it is converted."*

Overall, when faced with a novel task, the problem solving training the students in the Engineering course were exposed to regarding relationships between compounds fails them. The students in the Engineering course tend to use relationships between compounds in the same way students who haven't been exposed to that type of training do, by using them as part of dimensional analysis algorithms or in ways that don't allow them to successfully solve the task. This means that the way relationships are presented in the Engineering course is not helping these students in novel situations, and that new methods may be necessary.

Common Errors and Overall Findings for Unfamiliar Task

The most common error for students while working on the unfamiliar task was that they applied what they knew about familiar stoichiometry tasks and how to solve them, to this unfamiliar task. Many of the students, even the ones who eventually use algebra and relationships between carbon monoxide and carbon dioxide to answer the question, started off writing balanced equations and converting the amount of octane given into a mass. This was the only work related to standard process of using dimensional analysis that was needed to solve this task. However, many other students continued to use the balanced equation or equations they had written to calculate the amount of carbon dioxide produced and divide that value by the total mass of products given. Other students gave up on the task when they couldn't decide whether or not to combine the complete and incomplete combustion equations together because the total mass of the three products were given. Finally, a few students were concerned about the amount of oxygen used in the reaction, and one student went so far as to calculate the amount of carbon dioxide produced based on the amount of oxygen available instead of octane, as implied in the question. Students seeing the question as a limiting reactant task may cause this concern about the amount of oxygen used. Since carbon monoxide is usually produced under low oxygen conditions, students may assume that oxygen is a limiting reactant in this question and adjust their problem solving strategies accordingly. All of these errors lead to students either giving up while working on the task, or getting the incorrect answer to the task.

Students could not conceive of a stoichiometry task that did not require dimensional analysis to solve. When faced with such an algorithmic task, students chose to use what they knew how to do instead of thinking critically about how to answer the task. Only six students of the 23 interviewed realized that the question could be answered using a balanced chemical equation and a mathematical equation relating the amount of carbon dioxide and carbon monoxide produced to the amount of octane available. Overall, quite a few students gave up while working on this task because their ideas of how to solve the task were not sufficient. A few students calculated the percentage of carbon dioxide produced using the dimensional analysis approach, most without taking into account any carbon monoxide produced. Overall, the students' problem solving skills were not sufficient to successfully answer this question for the majority of the people interviewed.

OVERALL FINDINGS FOR THE CHAPTER

In both the Engineering and Science Majors course the use of multiple stoichiometry questions that require dimensional analysis to solve presented in lecture trains students to use dimensional analysis to solve any stoichiometry task they are presented with. The way problem solving for these tasks is presented in lecture encourages students to identify stoichiometry tasks and use dimensional analysis to solve them. When working on the familiar task this was the path taken by most of the students and even then 61% of the students were unable to successfully solve the task. This level of error may indicate that the students know the process for the solving these types of tasks and that they were prone to lapses in attention when they were working through the problem. It is possible that they are so well trained at how to solve these tasks, that they go onto autopilot and don't check the details of their work.

When presented with an unfamiliar stoichiometry task, students often start working through the dimensional analysis procedure they know, even when it is not needed to successfully solve the task. Some students used this procedure to get an answer to the question while completely ignoring the fact that two reactions are going on. Other students give up when the dimensional analysis process is interrupted, by requiring the use of more than one chemical equation for example. There were also students who may have identified the task as a limiting reactant problem based on their knowledge of complete and incomplete combustion. Overall, the idea that students did not take the time to think through the unfamiliar task and decide what they needed to do was the main problem students had with this task. The fact that only six students wrote equations using a relationship between the products of the two reactions, and only one was able to successfully solve the equation he wrote is a major finding of this study.

Implications for Teaching

This study has implications for chemistry professors and the way that stoichiometry is taught. It is clear that the way stoichiometry is currently being taught, at least as far as gram-to-gram conversion, is not helping students develop their problem solving skills, only their skills in recognizing different types of tasks. If one of our goals in teaching chemistry is to improve students' problem solving skills, then we need to think about changing the types of tasks we give students while teaching the topic of stoichiometry. For one thing, students need to be given stoichiometry tasks that are problems for them, not just exercises. While exercises are valid ways to practice skills and the use of algorithms, they do not expand students' problem solving skills and critical thinking.

There of course risks with testing for problem solving skills and critical thinking. Students are generally concerned with test scores and if tests were more focused on problem solving, there is a possibility that very few students would get those questions correct. This would lead to low exam scores and students feeling demoralized about the exams and therefore not studying for them. This could lead to a decrease in students' understanding of material. Top improve the testing environment, perhaps problem solving skills and critical thinking need to be evaluated on homework assignments or quizzes, that generally carry less weight in final course scores as compared to exams. This would allow professors to test such skills without leading to as negative of testing environment.

Suggestions for designing questions that are problems for students are challenging to define, because what is a problem for one student may be an exercise for another. Potter and Overton (1984) wrote tasks that were designed to be problems for students. These problems "were designed which required an unfamiliar approach to obtaining a solution, used a real life context and which had insufficient data" (Overton & Potter, 2008). These suggestions tie in well with the data presented in this study. The IMMEX questions students worked on used real world contexts and insufficient data in the prompt to allow students to have to search for more information to solve the task, in a relevant context. The unfamiliar task that students were asked to solve in the interview required to students to use an algebraic approach instead of the dimensional analysis procedure they were used to. From the interview analysis in this study, using tasks that require two or more reactions and a relationship between the reactions is a way to develop a question that may be a problem for students.

Assigning stoichiometry tasks that have more characteristics of problems, such as a real world context, that require the application of knowledge in a new way, and leaving out key information, are ways that chemistry professors can teach students about problem solving skills as well as chemistry content.

Limitations of Study

The order of questions asked, and whether a simulation question was asked between the two-stoichiometry questions may have lead to fatigue in students. This is a limitation to the study, as all students should have been asked to answer the same questions in the same order to ensure answers were consistent across interviews. The fact that students volunteered to be interviewed, instead of being selected in a stratified way is a limitation of this study. One of the research questions to be answered in this study was what factors lead to a task becoming a problem for students instead of an exercise. Only two questions, one familiar and one unfamiliar were used in this study. It is possible that another question that was identified as having characteristics of both an exercise and a problem could have better answered this question. Finally, only students in an Engineering and Science Majors course at Iowa State University were compared and so the results are not be necessarily generalizable to other populations.

Future Work

It can be argued that once a task has been broken down into the behaviors that students use to solve it, as shown in Table 38, the task's content become irrelevant. One could compare problem solving behaviors for the two stoichiometry problems with the second thermochemistry question for example, which may be somewhere between a familiar and novel task for students. The thermochemistry questions could also be analyzed in the same way the stoichiometry questions were, however the thermochemistry content tends to be compartmentalized within the general chemistry curriculum as compared to stoichiometry.

CHAPTER 6: CONCLUSIONS

Summary of Findings

The study discussed in this work used a mixed method approach to study what influence the course structure and amount of content covered had on students' exam performance and problem solving skills. The study compared a one-semester general chemistry course for Engineering students to the first semester of two sequence for Science Majors. Lecture capture data, exam scores, interviews, the Attitude about the Subject of Chemistry Inventory version two and on-line homework assignments were collected. The data in the IMMEX chapter showed that students exhibit increasing effective and efficient problem solving skills over the course of increasing complex problems, when they know the chemistry context for the task. However, when faced with an unfamiliar context, they revert to using skills that are less effective and efficient.

The data in the Quantitative chapter showed that using resurrection points on a comprehensive final exam, led students to perform better than students using a replacement exam as a way to demonstrate learning of missed material, particularly the Middle Bottom quartile of students. This may be due to the Engineering students studying more of the breadth of material covered on the final exam as compared to the Science Majors students. Another finding from the Quantitative chapter is that taking laboratory leads students to perform better on exams where the content covered in the laboratory is closely related to the exam content coverage. Finally, there were some changes in students' attitudes about chemistry over the course of the semester. The top two quartiles of students in the Science Majors course made greater strides in their satisfaction with chemistry compared to the

Engineering students, but for the Middle Bottom and Bottom quartiles, it was the Engineering course that leads to greater gains in satisfaction over the course of the semester. While the amount of content covered in the Engineering course may have led to more students in each quartile feeling that chemistry was more challenging at the end of the semester, the Engineering course was also the one with the most students that held a more positive view of the difficulty of chemistry for each quartile at the end of the semester. The laboratory had no influence on changing the students' perceptions of the satisfaction they received from studying chemistry, but it did lead them to believe that chemistry was more challenging at the end of the semester, compared to the Engineering students who did not take the laboratory.

From the interview analysis in the Qualitative chapter the data suggests that when students are presented with a familiar problem they use algorithms they know to solve the problem and if they make an error it's usually because they are being as careful about the details of the process as they should be. As the problem solving related to the first interview question was similar in both courses, there is no influence on how students solved the familiar task during the interview. For the novel task, the students chose a variety of methods to solve the problem and almost all of them were unsuccessful. It appears that adding one extra layer of complexity, by requiring the students to deal with two equations while working on the task, is enough to change an exercise into a problem. One of the goals in the Engineering course was to make the relationships between the various compounds in a reaction explicit. In spite of this, the Engineering students were no more likely than the Science Majors students to notice the relationship between the two carbon containing products in task two. Even if they could write equations to show a relationship between the compounds, they were unable to successfully solve the equation. The type of relationship needed to answer the second stoichiometry task is one that's usually seen during the discussion of alloys or isotopes. From the interview analysis it is clear that students are not seeing the connection between the types of relationships discussed in other areas of chemistry and their applicability to stoichiometry tasks. These findings lead to implications for teaching involving how we measure problem solving, how we teach relationships in chemistry, and whether we require laboratory as a co-requisite to chemistry lecture.

Implications for Teaching

One of the implications of this study is that if a professor wants to know more about students' problem solving skills, they need to ensure that the students understand the context of the question. If not, they may display less effective and efficient problem solving than they are capable based on the data from the IMMEX chapter. This study suggests, that in addition to using real-world contexts, applying and extending knowledge gained, and leaving out information, using examples where there are two reactions occurring and the relationship between the reactions is important to solving the task, are ways to turn exercises into problems for students (Overton & Potter, 2008). Another way chemistry professors could use the information from this study is when they are designing their courses, particularly for Engineering students. Using resurrection points on a final exam as a motivation to learn missed material seems to have more benefit for students, particularly lower performing students, than replacement exams. This assumes that one is going to use a comprehensive final exam as part of the course. Another consideration to make is whether or not to make laboratory mandatory for students or not. It has been shown in this study that when the content covered in the laboratory and the exam content are aligned; the laboratory students out perform non-laboratory students. If laboratory were not a co-requisite for the course, one would need to write exams with an awareness of the laboratory coverage so as to write exams that are fair to all students. These data could also be used to argue that all students need to take the laboratory, as it will improve their understanding of chemistry as measured by exams that are aligned with the laboratory content.

Future Work

There are questions that arise from this study that deserve further investigation. A survey asking about how long and what content the students are studying before the final exam, and perhaps throughout the semester, may shed light on differences between the two courses, if any exist. In particular, one could focus on the time between the fourth exam and final and see if there is a difference in between the study habits in particular between students who are and are not taking the replacement exam. There could also be a difference between groups of students who studies throughout the semester versus right before the final exam.

Another area of further research is in following these Engineering students into their Engineering course work. An analysis of the level of application of the chemistry content covered in the survey course would be important, along with the ability to answer questions about how they are using any problem solving skills they developed during their chemistry course. The Engineering professor feels that the chemistry content is important, but the problem solving skills that the students develop in his course are probably the most important things they will take from his course. It would be interesting to know if these skills are being transferred to the Engineering courses these students are taking in the future.

In the qualitative portion of this project, only two tasks were analyzed which represent extremes of the familiar-novel spectrum of task descriptions. It would be interesting to observe what happens when students work on a task that falls somewhere between the two extremes. The second written thermochemistry question may represent a middle ground. It could be analyzed in the say way that the two stoichiometry problems were and could be used as a point of comparison for the types of problem solving behaviors students exhibit when working on a moderately unfamiliar task. Another area of possible research with data that is already available, is to use the lecture capture data from the Engineering and Science Majors courses to identify differences in how material and/or problem solving skills were presented and then compare that to how students solved long answer questions related to that content. It is possible that how problem solving was presented for other topics may have lead to differences in students' problem solving behaviors as evidenced by their long answers to exam questions. These are all areas worthy of more analysis, some of which may already have data available.

ACKNOWLEDGEMENTS

This dissertation would not have come together without the help and guidance of many people. I would like to thank my committee, Dr. Greenbowe, Dr. Andre, Dr. Jeffries-EL, Dr. Winter, and of course my advisor, Dr. Holme. I would especially like to thank Dr. Holme for suggesting that I attend the first ever conference for Chemical Education graduate students at Miami University in Ohio. It was there that I met so many helpful colleagues and friends. I appreciate the time and effort that Stacey Lowry-Bretz and her graduate students put into designing and running that conference.

I would also like to thank the post-doc colleagues I have had the privilege of working with over the last three years. Without the help of Drs. Mary Emineke, Kim Linenberger, Megan Grunert, Jeffrey Raker, and Jacob Schroeder, I would not have been able to complete this work. I thank them for their guidance not only about work, but also about life. I wish them all the best of luck in their lives. Without Dr. Tanya Gupta's help, the data collection and analysis for the qualitative chapter would never have gotten done. Thank you for meeting me at all hours of the day and night to scan and code exams. To my group members, past and present, thank you for keeping my spirits up and offering your suggestions.

To the graduate student group members, thank you for your support. To the girls from group, thank you for the support and for being impressed and proud of my accomplishments. It means a lot to me. To the Friday night gaming group, thanks for giving me something to look forward to.

To my family, both new and old, thank you. Mom, Dad, and Laura, thank you for listening to the highs and lows. Laura, thanks for reading everything I sent to you, even if

you didn't understand. Mom and Dad, thanks for reminding me that I can do anything I put my mind to, and that this was no different.

And finally to my husband, Sean, I want to thank you for choosing to spend your life with me, no matter what. And thanks for supporting me and taking care of me through this process. You made my days in the Kraus group enjoyable, the days in the Holme group worthwhile. I look forward to being able to spend more time together now.

Thank you to everyone who made the last six and half years possible and even fun along the way.

BIBLIOGRAPHY

Bauer, C. F. (2008). Attitude towards Chemistry: A Semantic Differential Instrument for Assessing Curriculum Impacts. *Journal of Chemical Education , 85*, 1440-1445.

Bodner, G. M. (2000). Mental models: The role of representations in problem solving in chemistry. *University Chemistry Education , 4*, 24-30.

Bodner, G. M. (2003). Problem Solving: the difference between what we do and what we tell students to do. *University Chemical Education , 7*, 37-45.

Bowen, C. (1994). Think-aloud Methods in Chemistry Education: Understanding Student Thinking. *Journal of Chemical Education , 71*, 184-190.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How People Learn: Brain, Mind, Experience, and School.* Washington, D. C.: National Acadamy Press.

Burdge, J. (2009). *Chemistry 2nd Edition.* New York, New York: McGraw Hill.

Calimsiz, S. (2003). Problem solving cannot be taught: We can teach gap reducing techniques. West Lafayette, Indiana: Purdue University.

Chandrasegaran, A. L. (2009). Students' dilemmas in reaction stoichiometry problem solving: Deducing the limiting reagent in chemical reactions. . *Chemistry Education Research and Practice , 10*, 14-23.

Chen, X., & Weko, T. (2009). *Students Who Study Science, Technology, Engineering, and Mathematics (STEM) in Postsecondary Education.* National Center for Educational Statistics, Institute of Educational Sciences. National Center for Educational Statistics I.

Cooper, M., Stevens, R., & Holme, T. (2006). Assessing Problem-Solving Strategies in Chemistry using the IMMEX system. *Proceedings of the National STEM Assessment Conference* , 118-129.

Froyd, J. E. (2008). *White Paper on Promising Practices in Undergraduate STEM Education*. Retrieved 2012, May 31. from National Academies : http://www7.nationalacademies.org/bose/Froyd_Promising_Practices_CommissionedPaper.pdf

Hayes, J. (1981). *The Complete Problem Solver.* Philadephia, Pennsylvania: The Franklin Institute.

Herschbach, D. (1997). "Resurrection" Points. In S. Tobias, & J. Raphael, *The Hidden Curriculum Faculty-Made Tests in Science Part 1: Lower-Division Courses* (pp. 100-102). New York , New York: Plenum Press.

Holme, T. &. (2011). Assessing Conceptual and Algorithmic Knowledge in General Chemistry with ACS Exams. *Journal of Chemical Education , 88*, 1217-1222.

Holme, T. (1992). Using the Socratic Method in Large Lecture Courses. *Journal of Chemical Education , 69* (12), 974-977.

Jonassen, D. H. (2000). Toward a Design Theory of Problem Solving. *Educational Technology Research and Development , 48* (4), 63-85.

King, D. B. (2011). Using Clickers to Identify the Muddiest Points in Large Chemistry Classes. *Journal of Chemical Education , 88*, 1485-1488.

Kline, R. B. (2011). *Principles and Practices of Structural Equation Modeling* (3rd ed.). New York, New York, United States: The Guilford Press.

Kuenzi, J. J. (2008). *Science, Technology, Engineering, and Mathematics (STEM) Education: Background, Federal Policy, and Legislative Action.* Congressional Research Service. Congressional Research Service Reports.

Labov, J. B., Singer, S. R., George, M. D., Schweingruber, H. A., & Hilton, M. L. (2009). Effective Practices in Undergraduate STEM Education Part 1: Examining the Evidence. *CBE-Life Sciences Education , 8*, 157-161.

Lewis, S. E., & Lewis, J. E. (2005). The Same or Not the Same: Equivalence as an Issue in Educational Research. (J. W. Moore, Ed.) *Journal of Chemical Education , 82* (9), 1408-1412.

Lyon, D. C., & Lagowski, J. J. (2008). Effectiveness of Facilitating Small-Group Learning in Large Lecture Classes. *Journal of Chemical Education , 85*, 1571-1576.

Majerich, D. M., & Schmuckler, J. S. (2007). Improving Students' Perceptions of Benefits of Science Demonstrations and Content Mastery in a Large-Enrollment Chemistry Lecture Demonstration Course for Nonscience Majors. *Journa of College Science Teaching , 36*, 60-67.

Matz, R. L., Rothman, E. D., Krajcik, J. S., & Banaszak Holl, M. M. (2012). Concurrent Enrollment in Lecture and Laboratory Enhances Student Performance and Retention. *Journal of Research in Science Teaching , 49* (5), 659-682.

Mehrota, K., Mohan, C. K., & Ranka, S. (1997). *Elements of Artifical Neural Networks.* Cambridge, MA: MIT Press.

Murphy, K., Holme, T., Zelinsky, A., Caruthers, H., & Knaus, K. (2012). Building the ACS Exams Anchoring Concept Content Map for Undergraduate Chemistry. *Journal of Chemical Education , 89*, 715-720.

Nakhleh, M. B. (1993). Are our students conceptual thinkeres or algorithmic problem solvers? *Journal of Chemical Education , 70*, 52-55.

Nordstrom, B. H. (1990 April). Predicting Perfomance in Freshman Chemistry. 1-18. Boston, MA, USA.

Nurrenbern, S. &. (1987). Conceptual learning versus problem solving: Is there a difference? *Journal of Chemical Education , 64*, 508-510.

O'Connell, E. M. (2010, March 21). Rapid Knowledge Assessment: Correlating Student Reported Immediate First Steps and Problem Solving Efficiency. *239* .

Overton, T. & Potter, N. (2008). Solving open-ended problems, and the influence of cognitive factors on student success. *Chemical Education Research and Practice , 9*, 65-69.

Paas, F. G. (1992). Training Strategies for Attaining Transfer of Problem-Solving Skill in Statistics: A Cognitive-Load Approach. *Journal of Educational Psychology , 84* (4), 429-434.

Pallant, J. (2010). *SPSS Survival Manual.* Maidenhead, Berkshire, England: Open University Press McGraw Hill.

Raker, J. R., & Towns, M. H. (2010). Benchmarking problems used in second year level organic chemistry instruction. *Chemical Education Research and Practice , 11*, 25-32.

Reed, S. K., Ernst, G. W., & Banerji, R. (1974). The Role of Analogy in Transfer between Similar Problem States. *Cognitive Psychology , 6*, 436-450.

Rose, D., & Sullivan, O. (1996). *Introducing Data Analysis for Social Scientists* (2nd Edition ed.). Buckingham, England: Open University Press.

Sharma, M. D., Johnston, I. D., Johnston, H., Varvell, K., Robertson, G., Hopkins, A., et al. (2010). Use of interactive lecture demonstrations: A ten year study. *Physical Review Special Topics- Physics Education Research , 6*, 020119-1-020119-9.

Shell, D. F., Brooks, D. W., Trainin, G., Wilson, K. M., Kauffman, D. F., & Herr, L. M. (2010). *The Unified Learning Model: How Motivational, Cognitive, and Neuroboiological Sciences Inform Best Teaching Practices.* New York: Springer.

Stevens, R. (2008). A Value-Based Approach for Quantifying Student's Scientific Problem Solving Efficiency and Effectiveness Within and Across Educational Systems. In R. W. Lissitz, *Assessing and Modeling Cognitive Development in Schools: Intellectual Growth and Standard Setting.* Maple Grove, MN: JAM Press.

Stevens, R., & Thadani, V. (2007). Quantifying Student's Scientific Problem Solving Effeciency and Effectiveness. *Tech. Inst. Cognition and Learning , 5*, 325-337.

Stevens, R., Ikeda, J., Casillas, A., Palacio-Cayento, J., & Cylman, S. (1999). Artificial neural network-based perfomance assessments. . *Computers in Human Behavior , 15*, 295-313.

Stevens, R., Johnson, D. F., & Soller, A. (2005). Probabilities and Predictions: Modeling the Development of Scientific Problem-Solving Skills. *Cell Biology Education , 4*, 42-57.

Stevens, R., Soller, A., Cooper, M., & Sprang, M. (2004). Modeling the Development of Problem Solving Skills in Chemistry with a Web-Based Tutor. *Intelligent Tutoring Systems, 7th International Conference Proceedings*, (pp. 580-591).

Taasoobshirazi, G., & Glynn, S. M. (2009). College Students Solving Chemistry Problems: A Theoretical Model of Expertise. *Journal of Research in Science Teaching , 46* (10), 1070-1089.

Vendlinski, T., & Stevens, R. (2002). Assessing Student Problem-Solving Skills with Complex Computer-Based Tasks. *JTLA , 1*, 1-20.

Vygotsky, L. S. (1978). *Mind in Society.* Cambridge, MA: Harvard University Press.

Walker, J. D., Cotner, S. H., Baepler, P. M., & Decker, M. D. (2008). A Delicate Balance: Integrating Active Learning into a Large Lecture Course. *CBE-Life Sciences Education , 7*, 361-367.

Wheatley, G. H. (1984). *Problem solving in school mathematics.* Purdue University, School Mathematics and Science Center. West Lafayette : Purdue University.

Xu, X., & Lewis, J. E. (2011). Refinement of a Chemistry Attitude Measure for College Students. *Journal of Chemical Education , 88*, 561-568.

APPENDIX

QUANTITATIVE CHAPTER

EXAM AVERAGES – COMPARING QUARTILES WITHIN COURSES

Table 47 Kruskal-Wallis equality-of-populations rank test results of analysis of median differences within the Engineering and Science Majors Classes
*** = < 0.10, ** = < 0.05, *** = < 0.01**

χ^2 statistic (p-value)	Exam 1	Exam 2	Exam3	Exam 4	Final Exam
Engineering Class (df = 3)	251.2 (0.0001)***	259.7 (0.0001)***	254.8 (0.0001)***	307.1 (0.0001)***	320.5 (0.0001)***
Science Majors Class (df = 3)	204.6 (0.0001)***	247.0 (0.0001)***	221.4 (0.0001)***	187.6 (0.0001)***	262.9 (0.0001)***

Table 48 Ranksum results for Exam 1 for Engineering Class
*** = < 0.02, ** = < 0.008, *** = < 0.002**

Group Compared	z-statistic (p-value)	Group Compared	z-statistic (p-value)
Top / Middle Top	6.590 (0.0000)***	Middle Top/ Middle Bottom	5.930 (0.0000)***
Top/ Middle Bottom	10.567 (0.0000)***	Middle Top/ Bottom	10.714 (0.0000)***
Top/ Bottom	12.504 (0.0000)***	Middle Bottom/ Bottom	7.375 (0.0000)***

Table 49 Ranksum results for Exam 2 for Engineering Class
*** = < 0.02, ** = < 0.008, *** = < 0.002**

Group Compared	z-statistic (p-value)	Group Compared	z-statistic (p-value)
Top / Middle Top	6.407 (0.0000)***	Middle Top/ Middle Bottom	6.380 (0.0000)***
Top/ Middle Bottom	11.065 (0.0000)***	Middle Top/ Bottom	12.894 (0.0000)***
Top/ Bottom	14.017 (0.0000)***	Middle Bottom/ Bottom	9.039 (0.0000)***

Table 50 Ranksum results for Exam 3 for Engineering Class
*** = < 0.02, ** = < 0.008, *** = < 0.002**

Group Compared	z-statistic (p-value)	Group Compared	z-statistic (p-value)
Top / Middle Top	6.914 (0.000)***	Middle Top/ Middle Bottom	4.154 (0.0000)***
Top/ Middle Bottom	11.400 (0.0000)***	Middle Top/ Bottom	10.035 (0.0000)***
Top/ Bottom	13.770 (0.0000)***	Middle Bottom/ Bottom	4.632 (0.0000)***

Table 51 Ranksum results for Exam 4 for Engineering Class
*** = < 0.02, ** = < 0.008, *** = < 0.002**

Group Compared	z-statistic (p-value)	Group Compared	z-statistic (p-value)
Top / Middle Top	4.347 (0.0000)***	Middle Top/ Middle Bottom	0.948 (0.3429)
Top/ Middle Bottom	8.248 (0.000)***	Middle Top/ Bottom	7.447 (0.0000)***
Top/ Bottom	12.001 (0.0000)***	Middle Bottom/ Bottom	2.227 (0.0260)

Table 42 Ranksum Results for Final Exam for Engineering Class
*** = < 0.02, ** = < 0.008, *** = < 0.002**

Group Compared	z-statistic (p-value)	Group Compared	z-statistic (p-value)
Top / Middle Top	12.681 (0.0000)***	Middle Top/ Middle Bottom	10.511 (0.0000)***
Top/ Middle Bottom	14.166 (0.0000)***	Middle Top/ Bottom	13.714 (0.0000)***
Top/ Bottom	14.527 (0.00000***	Middle Bottom/ Bottom	12.292 (0.0000)***

Table 53 Ranksum Results for Exam 1 for Science Majors Class
*** = < 0.02, ** = < 0.008, *** = < 0.002**

Group Compared	z-statistic (p-value)	Group Compared	z-statistic (p-value)
Top / Middle Top	5.986 (0.0000)***	Middle Top/ Middle Bottom	3.123 (0.0018)***
Top/ Middle Bottom	9.974 (0.0000)***	Middle Top/ Bottom	10.276 (0.0000)***

Table 53 Ranksum Results for Exam 1 for Science Majors Class
*** = < 0.02, ** = < 0.008, *** = < 0.002 ctd**

Group Compared	z-statistic (p-value)	Group Compared	z-statistic (p-value)
Top/ Bottom	12.874 (0.0000)***	Middle Bottom/ Bottom	7.403 (0.0000)***

Table 54 Ranksum results for Exam 2 for Science Majors Class
*** = < 0.02, ** = < 0.008, *** = < 0.002**

Group Compared	z-statistic (p-value)	Group Compared	z-statistic (p-value)
Top / Middle Top	10.285 (0.0000)***	Middle Top/ Middle Bottom	8.239 (0.0000)***
Top/ Middle Bottom	12.480 (0.0000)***	Middle Top/ Bottom	12.733 (0.0000)***
Top/ Bottom	13.458 (0.0000)***	Middle Bottom/ Bottom	9.786 (0.0000)***

Table 43 Ranksum results for Exam 3 for Science Majors Class
*** = < 0.02, ** = < 0.008, *** = < 0.002**

Group Compared	z-statistic (p-value)	Group Compared	z-statistic (p-value)
Top / Middle Top	9.873 (0.0000)***	Middle Top/ Middle Bottom	9.439 (0.0000)***
Top/ Middle Bottom	12.338 (0.0000)***	Middle Top/ Bottom	12.523 (0.0000)***
Top/ Bottom	13.362 (0.0000)***	Middle Bottom/ Bottom	9.204 (0.0000)***

Table 56 Ranksum results for Exam 4 for Science Majors Class
*** = < 0.02, ** = < 0.008, *** = < 0.002**

Group Compared	z-statistic (p-value)	Group Compared	z-statistic (p-value)
Top / Middle Top	11.078 (0.0000)***	Middle Top/ Middle Bottom	9.012 (0.0000)***
Top/ Middle Bottom	12.307 (0.0000)***	Middle Top/ Bottom	11.921 (0.0000)***
Top/ Bottom	13.320 (0.0000)***	Middle Bottom/ Bottom	10.573 (0.0000)***

Table 57 Ranksum Results for Final Exam for Science Majors Class
*** = < 0.02, ** = < 0.008, *** = < 0.002**

Group Compared	z-statistic (p-value)	Group Compared	z-statistic (p-value)
Top / Middle Top	5.925 (0.0000)***	Middle Top/ Middle Bottom	3.678 (0.0002)***
Top/ Middle Bottom	10.788 (0.0000)***	Middle Top/ Bottom	11.218 (0.0000)***
Top/ Bottom	13.158 (0.0000)***	Middle Bottom/ Bottom	5.103 (0.0000)***

Table 58 Variable Code Book for Stata Data

Variable Name	Label- short description	Long Description
studyid	study identification number	study identification number E### = Engineering ### = Science Majors
course	identifies which class student is part of 1 = Engineering 2 = Class B	identifies which class student is part of 1 = Engineering 2 = Science Majors
sex	student sex 0 = female 1 = male	student sex 0 = female 1 = male
section	recitation section number	recitation section number
courseperc	percentage earned by end of semester	percentage earned by end of semester assigned by course professor
placescore	score on departmental placement test given over the summer before semester	score on departmental placement test given over the summer before semester Max score = 55pts
toledoscore	score earned on chemistry items on Toledo Test given during first week of class	score earned on chemistry items on Toledo Test given during first week of class Max score = 45 pts
exam1	score on Exam 1	score on Exam 1 Max score = 100 pts
exam2	score on Exam 2	score on Exam 2 Max score = 100 pts
exam3	score on Exam3	score on Exam 3 Max score = 100 pts
exam4	score on Exam 4	score on Exam 4 Max score = 100 pts

Table 58 Variable Code Book for Stata Data ctd

Variable Name	Label- short description	Long Description
finalexam	score on final exam Max score Engineering = 200 pts Max score Science Majors= 150 pts	Engineering final was instructor written. Science Majors final was a combination of an ACS conceptual exam for a first semester general chemistry course and an instructor written final
bpre1-bpre8	rating of items one through eight on Bauer Semantic Differential given pre	rating of items on Bauer Semantic Differential given during the first week of the semester in recitation. All Items have been recoded such that 1 = positive end to 7 = negative end of scale. Items are as follows: easy-hard, simple-complicated, clear-confusing, comfortable-uncomfortable, satisfying-frustrating, not challenging-challenging, pleasant-unpleasant, and organized-chaotic.
bpost1-bpost8	rating of items one through eight on Bauer Semantic Differential	ratings were conducted in recitation the week after Thanksgiving break
prof	professor that taught lecture	1 = Course A, Professor A 2 = Course B, Professor B 3 = Course B, Professor C 4 = Course B, Professor D
eng	Engineering major or not 0 = Not Engineering Major 1 = Engineering Major	Majors were provided as part of demographic data from registrar, list of majors by College was obtained to determine Engineering Major status
group	Combination of course and eng status 1 = course A, Engineering Major 2 = course B, Non-engineering Major 3 = course B, Engineering Major	Combination of course and eng status 1 = course A, Engineering Major 2 = course B, Non- engineering Major 3 = course B, Engineering Major
hmwktotal	Total number of homework points earned	Total number of homework points earned, taken from excel spreadsheet provided at the end of the semester. Course A Max points = 80 Course B Max points = 75
b1ch- b8ch	Change in Bauer Semantic Differential pre-post	difference pre/post on each item (bpost-bpre)

Table 58 Variable Code Book for Stata Data ctd

Variable Name	Label- short description	Long Description
year	Year course taken	1 = 2010
quiz1-quiz16	recitation quizzes -> quizzes mostly algorithmic	Engineering had 11 quizzes (quiz1-quiz11) Science Majors had 16 quizzes (quiz1-quiz16
qe1	quizzes before exam1	Total points on quizzes before exam1 (Engineering = 33 points; Science Majors= 30 points)
qe2	quizzes between exam1 and exam2	Total points on quizzes between exam1 and exam2 (Engineering = 20 points; Science Majors= 30 points)
qe3	quizzes between exam2 and exam3	Total points on quizzes between exam2 and exam3 (Engineering = 20 points; Science Majors= 30 points)
qe4	quizzes between exam3 and exam4	Total points on quizzes between exam3 and exam4 (Engineering = 30 points; Science Majors= 31 points)
qfe	quizzes between exam4 and finalexam	Total points on quizzes between exam4 and finalexam (Engineering = 10 points; Science Majors= 30 points)
quiztotal	Total number of points on quizzes	Total points on quizzes overall (Engineering = 110 Max points; Science Majors= 161 Max Points)
c1-c36	Clicker quizzes in lecture -> usually testing material just covered in lecture	Engineering had 36 clicker quizzes (c1-c36) -> graded for attendance Science Majors had 28 clicker quizzes (c1-c28) -> graded for accuracy
ce1	Clicker quizzes between exam1 and exam2	Total points on clickers before exam1 (Engineering 16 Max points; Science Majors4 Max Points)
ce2	Clicker quizzes between exam1 and exam2	Total points on clickers between exam1 and exam2 (Engineering 18 Max points; Science Majors4 Max Points)
ce3	Clicker quizzes between exam2 and exam3	Total points on clickers between exam2 and exam3 (Engineering 23 Max points; Science Majors10 Max Points)
ce4	Clicker quizzes between exam3 and exam4	Total points on clickers between exam3 and exam4 (Engineering 34 Max points; Science Majors3 Max Points)
cfe	Clicker quizzes between exam4 and finalexam	Total points on clickers between exam4 and finalexam (Engineering 9 Max points; Science Majors8 Max Points)

Table 58 Variable Code Book for Stata Data ctd

Variable Name	Label- short description	Long Description
clickertotal	Total number of points on clicker quizzes	Total points on clicker quizzes overall (Engineering = 100 Max Points; Science Majors= 29 Max Points)
grade	courseperc divided into five groups based on percentages 1 = 90-100% 2 = 80-90% 3 = 70-80% 4 = 60-70% 5 = > 60%	In letter grade fashion, 1 = A 2 = B 3 = C 4 = D 5 = F
z_neg	Negative end of z-score of toledoscore -> used to determine equivalency	negative z-score used in equivalencing around toledoscores
z_pos	Positive end of z-score of toledoscore -> used to determine equivalency	positive z-score used in equivalencing around toledoscores
z_toledoscore	z-scores of toledoscore	z-scores of toledoscore as determined using Stata code zscore; shows each students relative variation from the mean toledoscore for each course
tsgrp	Equivalent group determined as described in chapter **(?)**	tsgrp=1 -> in equivalent group tsgrp=0 -> Not in equivalent group
cpq	Course percent divided into quartiles within each Engineering recombined together into one variable 1 = bottom 25% 2 = 50% 3 = 75% 4 = top 25%	Determined quartile cutoffs using pctile command for each course. Divided each course by these cutoffs and label 1 = bottom 25% 2 = 50% 3 = 75% 4 = top 25% Recombine Engineering and Science Majors data into one variable

Table 58 Variable Code Book for Stata Data ctd

Variable Name	Label- short description	Long Description
cpd	Course percent divided into deciles within each Engineering recombined together into one variable 1 = bottom 10% 2 = 20% 3 = 30% 4 = 40% 5 = 50% 6 = 60% 7 = 70% 8 = 80% 9 = 90% 10 = top 10%	Determined decile cutoffs using pctile command for each course. Divided each course by these cutoffs and label 1 = bottom 10% 2 = 20% 3 = 30% 4 = 40% 5 = 50% 6 = 60% 7 = 70% 8 = 80% 9 = 90% 10 = top 10% Recombine Engineering and Science Majors data into one variable
ze1	z-score of exam 1 for each class recombined together	zscore exam1 for each Engineering recombined together
ze2	z-score of exam 2 for each class recombined together	zscore exam2 for each Engineering recombined together
ze3	z-score of exam 3 for each class recombined together	zscore exam3 for each Engineering recombined together
ze4	z-score of exam 4 for each class recombined together	zscore exam4 for each Engineering recombined together
zfe	z-score of finalexam for each class recombined together	zscore finalexam for each Engineering recombined together
deltaze1ze2	Difference between z-scores for exam1 and exam2	Subtracted zscore for exam1 (ze1) from zscore for exam2 (ze2) for each Engineering recombined
deltaze2ze3	Difference between z-scores for exam2 and exam3	Subtracted zscore for exam2 (ze2) from zscore for exam3 (ze3) for each Engineering recombined
deltaze3ze4	Difference between z-scores for exam3 and exam4	Subtracted zscore for exam3 (ze3) from zscore for exam4 (ze4) for each Engineering recombined
deltaze4zfe	Difference between z-scores for exam4 and finalexam	Subtracted zscore for exam4 (ze4) from zscore for finalexam (zfe) for each Engineeringnd recombined

Table 58 Variable Code Book for Stata Data ctd

Variable Name	Label- short description	Long Description
finalp	Final exam score converted to percent; out of 100 percent so it can be compared to other exam scores	Final exam for each class divided by maximum number possible points times 100. Engineering = (finalexam/200)*100 Science Majors= (finalexam/150)*100
biapre	Pre Intellectual Accessibility Factor on modified Bauer Semantic Differential; sum of items 1, 2, 3, and 6 pre of the modified Bauer	Sum of items 1, 2, 3, and 6 on Bauer pre columns for both classes. Range = 4-28
biapost	Post Intellectual Accessibility Factor on modified Bauer Semantic Differential; sum of items 1, 2, 3, and 6 post of the modified Bauer	Sum of items 1, 2, 3, and 6 on Bauer post columns for both classes. Range = 4-28
bespre	Pre Emotional Satisfaction Factor on modified Bauer Semantic Differential; sum of items 4,5,7 and 8 pre of the modified Bauer	Sum of items 4,5,7 and 8 on Bauer pre columns for both classes. Range = 4-28
bespost	Post Emotional Satisfaction Factor on modified Bauer Semantic Differential; sum of items 4,5,7 and 8 post of the modified Bauer	Sum of items 4,5,7 and 8 on Bauer post columns for both classes. Range = 4-28
ce1p	Points earned on clickers, when graded for accuracy before exam 1 as a percentage	Graded for accuracy; Engineering = (ce1/16)*100 Science Majors= (ce1/4)*100 Recombined into one variable
ce2p	Points earned on clickers, when graded for accuracy before exam 2 as a percentage	Graded for accuracy; Engineering = (ce2/18)*100 Science Majors= (ce2/4)*100 Recombined into one variable
ce3p	Points earned on clickers, when graded for accuracy before exam 3 as a percentage	Graded for accuracy; Engineering = (ce3/23)*100 Science Majors= (ce3/10)*100 Recombined into one variable
ce4p	Points earned on clickers, when graded for accuracy before exam 4 as a percentage	Graded for accuracy; Engineering = (ce4/34)*100 Science Majors= (ce4/3)*100 Recombined into one variable

Table 58 Variable Code Book for Stata Data ctd

Variable Name	Label- short description	Long Description
cfep	Points earned on clickers, when graded for accuracy before final exam as a percentage	Graded for accuracy; Engineering = (cfe/9)*100 Science Majors= (cfe/8)*100 Recombined into one variable
examtotal	Sum of exam1-4 and finalp	Sum of exams 1-4 and finalp for both classes
examtotalpercent	conversion of examtotal to a percentage	Engineering = (examtotal/600)*100 Science Majors= (examtotal/550)*100 Recombined into one variable
zcfep	z-score of cfep	z-score of cfep for each Engineeringnd then recombined into one variable
zce4p	z-score of ce4p	z-score of ce4p for each Engineeringnd then recombined into one variable
zce3p	z-score of ce3p	z-score of ce3p for each Engineeringnd then recombined into one variable
zce2p	z-score of ce2p	z-score of ce2p for each Engineeringnd then recombined into one variable
zce1p	z-score of ce1p	z-score of ce1p for each Engineeringnd then recombined into one variable
z_fp	z-score of finalp	z-score of finalp for each Engineeringnd then recombined into one variable
z_ep	z-score of examtotalpercent	z-score of examtotalpercent for each Engineeringnd then recombined into one variable
z_cp	z-score of course percent	z-score of course percent for each Engineeringnd then recombined into one variable
grade_e	examtotalpercent divided into five groups based on percentages 1 = 90-100% 2 = 80-90% 3 = 70-80% 4 = 60-70% 5 = > 60%	In letter grade fashion, 1 = A 2 = B 3 = C 4 = D 5 = F

Table 58 Variable Code Book for Stata Data ctd

Variable Name	Label- short description	Long Description
cpq_e	examtotalpercent divided into quartiles within each Engineeringnd recombined together into one variable 1 = bottom 25% 2 = 50% 3 = 75% 4 = top 25%	Determined quartile cutoffs using pctile command for each course. Divided each course by these cutoffs and label 1 = bottom 25% 2 = 50% 3 = 75% 4 = top 25% Recombine Engineering and Science Majorsdata into one variable
cpd_e	examtotalpercent divided into deciles within each Engineeringnd recombined together into one variable 1 = bottom 10% 2 = 20% 3 = 30% 4 = 40% 5 = 50% 6 = 60% 7 = 70% 8 = 80% 9 = 90% 10 = top 10%	Determined decile cutoffs using pctile command for each course. Divided each course by these cutoffs and label 1 = bottom 10% 2 = 20% 3 = 30% 4 = 40% 5 = 50% 6 = 60% 7 = 70% 8 = 80% 9 = 90% 10 = top 10% Recombine Engineering and Science Majorsdata into one variable
deltazrawsum	Sum of delta z-scores for all exam pairs	deltaze2e1 + deltaze3e2 + deltaze4ze3 + deltazfeze4
abs_deltazrawsum	Absolute value of deltazrawsum	abs(deltazrawsum)
deltazrawcuttofs_symm	Symmetric cutoffs for determine small, medium and large changes in delta z-scores	Symmetric cutoffs for determine small, medium and large changes in delta z-scores; units = standard deviations
deltazrawcutoffs_unsymm	Unsymmetric cutoffs for determine small, medium and large changes in delta z-scores	Unsymmetric cutoffs for determine small, medium and large changes in delta z-scores; units = standard deviations

Table 58 Variable Code Book for Stata Data ctd

Variable Name	Label- short description	Long Description
taexperience	number of general chemistry courses taught by the TA since Fall 2006 semester	Values range from 0-7
lab	1 = no lab 2 = lab	1 = no lab 2 = lab
no_labgrp	1 = lab, 0 = nolab	

QUALITATIVE CHAPTER

INTERVIEW GUIDE

Interviewer Questions:

Have you had chemistry before?
Did you enjoy chemistry in high school?
What do you remember most about chemistry in high school?
Are you enjoying chemistry now?
Are you taking laboratory?

Stoichiometry

1) What mass of oxygen is needed to completely combust 1.00 g of ethanol to produce carbon dioxide and water vapor?

Interviewer Question:

Can you select your first step off the list?
A using a mole ration
B writing the balanced equation
C calculating the moles of oxygen
D calculating moles of ethanol
E calculating the molar mass of ethanol
F writing the chemical formula for ethanol
G reading the exercise, however I am not sure how to start the exercise

2) Simulation:

a) Click on the Select Gas tab and pick a gas (one of the known compounds). Balance the equation using the lowest ratio of whole numbers and submit the equation.

b) Add some of the gas to the reaction container and start the reaction. The simulation will burn the gas and pass the products through filters that will absorb the product molecules so that they can be weighted. Click on the Product button. Record the data you collect in the following table:

	C H +	O_2 →	CO_2 +	H_2O
Inital Amount (moles) - I				
Change (moles) - C				
Ending Amount (moles) - E				
Inital Amount (grams) - I				
Change (grams) - C				
Ending Amount (grams) - E				

c) Click on the Select Gas tab and pick the **unknown hydrocarbon C_xH_y**. Add some of the gas to the reaction container and start the reaction. Click on the Product button. Record the data you collected in the table below.

	Cx Hy +	O_2 →	CO_2 +	H_2O
Inital Amount (moles) - I				
Change (moles) - C				
Ending Amount (moles) - E				
Inital Amount (grams) - I				
Change (grams) - C				
Ending Amount (grams) -E				

d) Determine the possible values for x and y. Balance the equation using the resulting hydrocarbon.

$$+ \quad O_2 \quad \rightarrow \quad CO_2 + \quad H_2O$$

Interviewer Questions:

What assumptions are you making when working on this reaction?
What do the I, C, and E stand for?

Have you seen this kind of question before? (particularly the unknown hydrocarbon)

3) Octane (C_8H_{18}) is a component of gasoline. Complete combustion of octane yields H_2O and CO_2. Incomplete combustion produces H_2O and CO, which not only reduces the efficiency of the engine using the fuel but is also toxic. In a certain test run, 1.000 gallon (gal) of octane is burned in an engine. The total mass of CO, CO_2, and H_2O produced is 11.53 kg. Calculate the efficiency of the process; that is, calculate the fraction of octane converted to CO_2. The density of octane is 2.650 kg/gal.

Interviewer Questions:

If a student gets stuck, ask if they could solve whatever issue they're having, what would be there next steps to solve the task.

Thermochemistry

1) What is the final temperature (in ^{o}C) when 1 gallon of water evolves 118.8 kJ of heat when it cools from 32.5^{0}C?

Interviewer question:

A	**converting kJ to J.**
B	**writing the equation for heat.**
C	**converting the volume into metric.**
D	**converting the mass of water to grams.**
E	**solving the heat equation for change in temperature.**
F	**writing the equation for heat with both initial and final temperature.**
G	**reading the exercise, however I am not sure how to start the exercise**

2) Simulation:

a) Use the button to pick LiCl. Leave the water volume at 20 mL and the amount of LiCl at 0.50 g. Record the beginning condition of the solution in the table below:

Choose one compound from each of the following lists:

A)
LiCl
$CaCl_2$
NaOH
$Mg(NO_3)_2$
$Ca(NO_3)_2$
Na_2CO_3
$ZnSO_4$

B)
NH_4NO_3
KCl
NaCl
NH_4Cl
$NaNO_3$

C)
Sucrose
Urea

Compound	Mass of Solution	Mass of Compound	Initial Temp	Final Temp	Change in Temp	q_{soln}	ΔH

b) Click on the Start button. Fill in table above.

c) Write a chemical equation representing the process of dissolving the compounds you have chosen from lists A and B above. Draw a picture diagram for the process for any compound.

Interviewer:

Please include a key.

d) Are the following statement True or False. Please provide a reason for your answer.

1) The number of dissolved particles (ions or moles of ions) is related to the temperature change.

2) Certain cations are associated with either exothermic or endothermic processes.

3) Certain anions are associated with either exothermic or endothermic processes.

4) The amount of heat gained or released by a compound is different for different compounds.

Interviewer:

You can write True or False, and talk about your reason out loud.

3) One of the most popular approaches to dieting in recent years has been to reduce dietary fat. One reason many people want to avoid eating fat is its high Calorie content. Compared to carbohydrates and proteins, each of which contains an average of 4 Calories per gram (17 kJ/g), fat contains 9 Calories per gram (38 kJ/g). Tristearin, a typical fat, is metabolized (or combusted) according to the following equation:

$C_{57}H_{110}O_6\,(s) + 81.5\,O_2\,(g) \rightarrow 57\,CO_2\,(g) + 55\,H_2O\,(l)\ \Delta H^o = -37{,}760$ kJ/mole

Although the food industry has succeeded in producing low-fat versions of nearly everything we eat, it has thus far failed to produce a palatable low-fat doughnut. The flavor, texture, and what the industry calls "mouth-feel" of a doughnut depends largely on the process of deep-fat frying. Fortunately for people in the doughnut business, though, high fat content has not diminished the popularity of doughnuts.

According to the information obtained from www.krispykreme.com, a Krispy Kreme original glazed doughnut weighs 52 g and contains 200 Cal and 12 g of fat.

a) Assuming that the fat in the doughnut is metabolized according to the given equation for tristearin, calculate the number of Calories in the reported 12 g of fat in each doughnut.

b) If all the energy contained in a Krispy Kreme doughnut (not just the fat) were transferred to 6.00 kg of water originally at 25.5 ^oC, what would be the final temperature of the water?

c) When a Krispy Kreme apple fritter weighing 101 g is burned in a bomb calorimeter with C_{cal}= 95.3 kJ/^oC, the measured temperature increase is 16.7 ^oC. Calculate the number of Calories in a Krispy Kreme apple fritter.

d) What would the ΔH^o value be for the metabolism of 1 mole of the fat tristearin if the water produced by the reaction was gaseous instead of liquid?

IMMEX CHAPTER

Table 59 IMMEX Problem Sets

Topic	Real World Application	Activities or Information needed to Solve Problem
Identification of an unknown compound	Determine if a brush fire was natural or arson	1) Use wet chemical tests 2) Interpret ^{1}H, ^{13}C, MS and IR spectra 3) Correctly identify compound 4) Identify which plant, compound came from
Identification of an unknown compound	Determine the compound used to poison a professor	1) Use wet chemical tests 2) Interpret ^{1}H, ^{13}C, MS and IR spectra 3) Identify which student poisoned the teacher
Buffers	Acid Neutralization Capacity (ANC) of a lake	1) Find the Acid Neutralization Capacity (ANC) equation 2) Collect the concentrations of different components of ANC equation, bicarbonate, carbonate, pOH of lake and pH of rain 3) Convert pH to pOH 4) Multiply concentrations by appropriate coefficient 5) Calculate the ANC value of the lake in mg/L 6) Find the type of bedrock the lake is made of, and find the ANC value they calculated
pH equilibrium	Crystallization of a protein	1) Identify pI of protein 2) Determine the amino acid sequence of protein 3) Determine the sites on the protein that can be protonated 4) Determine pKa of each amino acid 5) Determine which sites will be protonated at the pI 6) Determine the percent of sites protonated 7) Identify the buffer used to reach the pI 8) Select the correct buffer and percent protonation for list

Table 59 IMMEX Problem Sets ctd

pH titration	Determine the amount of aglime to spread on your field	1) Determine current pH of soil 2) Identify crop and pH needed to maximize yield 3) Unit conversions to convert cm^2 to acres 4) Determine the amount of H^+ to be neutralized in sample titrated 5) Use pH titration information to determine amount of acid neutralized by NaOH 6) Use balanced equation or conversion factor to determine amount of aglime needed to neutralize acid in sample 7) Do unit conversion from mol calcium carbonate to tons 9) Divide tons of calcium carbonate by acres 10) Select appropriate answer from drop down menu
Battery; Electrochemistry	Build three batteries out of scrap metals and citrus	1) Find equation needed to determine EMF of a cell/battery 2) Identify anode and cathode of battery they wish to build 3) Fill in table with EMF of all possible combinations of 6 metals that would give a positive EMF 4) Identify the required combined EMF of three batteries 5) Determine the unique combination of 3 EMFs that equals the required EMF 6) Determine the metal combinations that form the three batteries 7) Select the correct three batteries from a list