INCREASING THE UNDERSTANDING OF CHEMICAL CONCEPTS: THE EFFECTIVENESS OF MULTIPLE EXPOSURES

By

Janet H. Bius

Submitted to the graduate degree program in Chemistry and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

_____

Chairperson Joseph A. Heppert

_____

Mikhail V. Barybin

_____

Heather Desaire

_____

Meagan M. Patterson

_____

David Weis

Date Defended: April 15, 2016

The Dissertation Committee for Janet H. Bius

certifies that this is the approved version of the following dissertation:

INCREASING THE UNDERSTANDING OF CHEMICAL CONCEPTS: THE EFFECTIVENESS OF MULTIPLE EXPOSURES

_____

Chairperson Joseph A. Heppert

Date approved: April 15, 2016

Abstract

Chemistry is difficult because it has multilevels of knowledge with each level presenting challenges in vocabulary, abstract thinking, and symbolic language. Students have to be able to transfer between levels to understand the concepts and the theoretical models of chemistry. The cognitive theories of constructivism and cognitive-load theory are used to explain the difficulties novice learners have with the subject of chemistry and methods to increase success for students.

The relationship between external representations, misconceptions and topics on the success of students are addressed. If students do not know the formalisms associated with chemical diagrams and graphs, the representations will decrease student success. Misconceptions can be formed when new information is interpreted based on pre-existing knowledge that is faulty. Topics with large amount of interacting elements that must be processed simultaneously are considered difficult to understand.

New variables were created to measure the number of times a student is exposed to a chemical concept. Each variable was coded according to topic and learning environment, which are the lecture and laboratory components of the course, homework assignments and textbook examples. The exposure variables are used to measure the success rate of students on similar exam questions.

Question difficulty scales were adapted for this project from those found in the chemical education literature. The exposure variables were tested on each level of the difficulty scales to determine their effect at decreasing the cognitive demand of these questions.

The subjects of this study were freshmen science majors at a large Midwest university. The effects of the difficulty scales and exposure variables were measured for those students whose exam scores were in the upper one-fourth percentile, for students whose test scores were

in the middle one-half percentile, and the lower one-fourth percentile are those students that scored the lowest on the exam. The most difficult for all three percentiles were the topics of acid/base equilibria and aqueous equilibria. The exposure variables of recall and algorithmic homework increased student success for all percentiles.

Students perform better on exam questions when they understand the terminology and symbolic representations of a topic.

Acknowledgements

To Joe Heppert whose support through the ups and downs of my career at The University of Kansas has been invaluable to my success. Thank you.

To Heather Desaire for providing a means for this research project.

To my committee members David Weis, Meagan Patterson, Misha Barybin, and Heather Desaire. Thank you for the sacrifice your time to serve on this committee.

To Dr. Mary Lynn DeShazo for her patience while teaching me the Krebs cycle, her mentorship, and her friendship.

To Mike for his encouragement and for keeping me focused on completing this dissertation.

To my son and daughter-in-law, Daniel and Tasha, and my grandsons Will and Noah, you will never have to hear the phrase “When I finish my dissertation, we will….” ever again!

To my brother-in-law, Brian, who oversees the Missouri branch of the family, thank you for all that you do for us and for your friendship throughout the years.

To my sisters, Diane and Mary, thank you for the help, support and love you have given to my family and me. I am blessed to have you as sisters and as friends.

I would like to dedicate this work to my mother, Mary Borgfeld, and to my son, William. I miss you both and wish you were here to help me celebrate achieving this life-long dream.

Table of Contents

Table of Tables

Table of Figures

Introduction

The study of chemistry is associated with the development of professionals in career fields ranging from healthcare to engineering. Subjects such as biology, geology, and agriculture require an understanding of chemistry. Yet students see chemistry as a difficult subject that is complicated and confusing, and students fear they will not be successful in the subject. Because they do not see a need for it in their future, they view it as worthless and boring. The combination of these factors leads to an unwillingness to invest time or effort in learning the chemistry subject matter (Bauer, 2008). This research project is designed to determine if multiple exposures in different learning environments will help students to become more successful in introductory chemistry courses at the collegiate level.

First, the learning theories of constructivism and cognitive load will be briefly reviewed, and the different learning theories will be used to explain why chemistry is so difficult for novice learners. Next, the pedagogies based on the learning theories will be examined. Finally, the importance of the learning environments and associated teaching methods will be discussed.

Learning Theories

Two learning philosophies that are the most relevant to this research project are the constructivist and cognitive load theories. Basic tenets of each theory are presented in the next section.

Cognitive load theory. This theory was developed from information attained through problem-solving experiments. The focus is to decrease the cognitive demand of instructional material to increase student understanding and learning. The theory is based on human cognitive structures which are working memory, long-term memory and schema acquisition (Bannert, 2002; Mostyn, 2012).

Information is processed in the working memory. A learner is cognizant of and can monitor only the contents of working memory. A limited amount of information can be processed at one time, and the amount depends on age and socioeconomic factors. For students in early adulthood, approximately five to seven pieces of information can be managed simultaneously (Bunce, 2005; Cowan, 2014; Sweller, van Merriënboer, & Paas, 1998).

After the information has been processed in the working memory, it is stored in the form of schemas in long-term memory. A schema is an organized group of information that is related by how the information will be used or how it connects to other stored schema (Cowan, 2014). As an example, when a child hears the word “dog,” the information he or she knows about dogs is recalled: furry, four legs, has a tail. As the child grows, more information is stored in the schema of “dog.” With this new information, dogs can be small or large and can have longhair or shorthair. As the schema grows, it incorporates the knowledge that dogs are mammals, were descended from the wolf (Carver), and its genus is canid. Complex schemas are built by combining lower-level schemes into higher-level ones. Schemas are treated as one element in working memory, and there is no limitation on size or complexity of the element. Schemas are stored in long-term memory and reduce working memory load (Sweller et al., 1998).

Automation in schema production occurs after extensive practice. When automation occurs, a process can be carried out with minimal conscious effort. Learners with automated schema have more working memory capacity and are able to solve more sophisticated problems (Sweller et al., 1998).

The focus of cognitive load theory is to design instructional material that reduces the cognitive load on working memory for novice learners. The principles used in creating this material are located in the pedagogy section of the introduction section.

Constructivism. The theory of constructivism states people construct their own knowledge through experience and reflection on those experiences. Piaget's Theory of Cognitive Development is based on this philosophy of learning (Bunce, 2001). First, Piaget's theory will be discussed, and then, the learning theory of constructivism will be reviewed.

Piaget argued knowledge is constructed when learners organize their experiences with their physical world to fit into pre-existing mental structures. Incorporation of new knowledge is done through the cognitive structures of assimilation and accommodation (Bodner, 1986).

When new knowledge can be integrated into a pre-existing conceptual framework, it is called assimilation, and there is conflict between the new ideas and the student's prior knowledge. As a result, the student's cognitive functions are at equilibrium. However, if the new information cannot be assimilated, the learner experiences disequilibrium, which is a stressful state. The learner will change their pre-existing conceptual structure to fit the new knowledge. This process is called accommodation (Bodner, 1986). Learning occurs because accommodation reduces the stress of disequilibrium.

Piaget described four stages of cognitive development. Each stage represents how children interpret this new knowledge. As children mature, they will pass through each stage of cognitive development. These levels are in Table 1 along with a brief description of the characteristics of the thinking processes for each stage (Bunce, 2001; Cracolice, 2005).

Table 1

Piaget's Four Levels of Cognitive Development

Level	Age	Developments
Sensorimotor	0 – 2 years	• Understands the permanence of objects • Understands the environment exists independently from self • Learns about physical world through touching and feeling
Preoperational	2 – 7 years	• Symbolical thinking - understands words represent people and things • Trouble seeing another person's point of view
Concrete operational	7 – 11 years	• Concrete thinking – focused on facts, the physical environment, and literal thinking • Can classify objects and put in increasing order • Conservation – changing the shape or appearance does not change the composition
Formal operational	11 – 15 years	• Understands abstract ideas and hypothetical situations • Uses deductive reasoning

Piaget stated most children pass through the first three levels of the hierarchy at approximately the same age. The last stage of formal operational depends on a child's educational background, and it is theorized that not all people reach this final level (Bird, 2010; Gould & Howson, 2011; Kuhn, 1979).

Bird (2010) tested 446 freshmen chemistry students and found 19% of the students were at concrete operational stage, 40% were in a transitional stage, and 41% were at the formal operational stage. Out of 131 freshmen students at an Oklahoma university, 50% were at the

concrete operational state, 25% were at post-concrete (a transition stage), and 25% were at the formal operational stage (McKinnon & Renner, 1971).

Bird (2010) described a relationship between a students' cognitive stage and their educational success in chemistry courses. The students in her study who earned an "A" in a freshman chemistry course were at the formal operational stage. Students with a course grade of a "B" were in the transitional stage of cognitive development, and students at the concrete operational stage earned a course grade of a "C". In addition, students at the formal operational stage in this study scored the highest on an ACS General Chemistry Examination, and the concrete operational stage scored the lowest. Transition stage students scored between the two extremes. Endler and Bond (2007) reported a relationship between educational achievement and increases in formal operational skills.

Personal constructivism. Knowledge is constructed in the mind of the learner (Bodner, Klobuchar, & Geelan, 2001). Conceptions about natural occurrences are developed through interactions with the physical environment. Meaning is constructed by students interpreting the new knowledge through their pre-existing frameworks (Garnett, Garnett, & Hackling, 1995). Conceptual change occurs when a student's existing conceptions are challenged by new information leading to accommodation (Garnett et al., 1995).

The principles of constructivism are that teaching should be student centered, students learn by interacting with the physical world, teaching should emphasize the process of learning, and teaching should recognize and accept the individual differences in learning (Bunce, 2001).

Social constructivism. This type of constructivism emphasizes the social aspects of learning. Higher cognitive functions develop with social interactions between individuals (Bunce, 2001), and the social context in which learning takes place is critically important

(Garnett et al., 1995). Students build their knowledge through personal experiences and interactions with others in the scientific community (Garnett et al., 1995).

Cognitive acceleration. Children pass through first three stages of Piaget's developmental levels at approximately the same age. Transitioning between concrete operational and formal operational is not as dependent on a specific age. Students in concrete operational stage cannot fully comprehend scientific thinking. If the cognitive demands of a task exceeds the cognitive level of a student, that student will not be able to complete the task (Endler & Bond, 2007).

Michael Shayer at King's College, London recognized the mismatch between the cognitive developmental stage of freshmen students and the cognitive demand of freshman chemistry courses. Michael Shayer and Philip Adey in 1981 created a teaching intervention based on the theories of Jean Piaget and Lev Vygotsky and on constructivist methodology (Endler & Bond, 2007). The object of the new pedagogy is to increase a pupil's developmental level from concrete operational to formal operational through increased exposure to the higher-order thinking skills. The intervention is called cognitive acceleration through science education (CASE).

CASE lessons usually have four parts. The first part is to introduce the definitions and variables of the lesson. Next, students explore or experiment with concrete models representing the concepts of the lesson. Then, a challenge to their perceptions of the experimental results of their experiments is given. This challenges their prior knowledge and personal theories.

Students work in small groups to explain the conflict between their observed results and the challenge to their observations through working problems and class discussions. These

processes expose students to the thinking of their peers. If students are on the verge of formal operational thinking, these activities should help them make the transition (Elkind, 1988).

Finally, the instructor leads them to apply this new knowledge to different but relevant contexts (Endler & Bond, 2007). Training for acceleration is more effective at the concrete to formal operational than it is at the lower levels of cognitive development (Elkind, 1988)

Detractors of the process do not believe the ability to think at the formal operational stage is increased, but instead they suggest the positive results are from training effects. These critics suggest the program is domain specific and does not transfer to other subjects (Haley & Good, 1976; Shayer, 1987).

Cognitive Demand of Chemistry

We have briefly discussed the constructivist and cognitive load theories that describe the cognitive skills and mental structures required for learning to occur. In this section, the cognitive demand of the subject of chemistry is discussed.

The first sub-section is a review of the chemical knowledge triplet introduced by Johnstone in 1982, which discusses the multi-level knowledge required to understand chemistry. Next we will discuss the difficulties novice learners have understanding external representations, and finally, we will investigate the part of misconceptions in student learning of chemistry.

Chemistry triplet. Johnstone designed the chemistry triplet to represent the relationships between the different levels of knowledge required to be proficient in chemistry. Later, in 1991, he arranged the levels in a triangle. A representation of the triangle is in Figure 1.

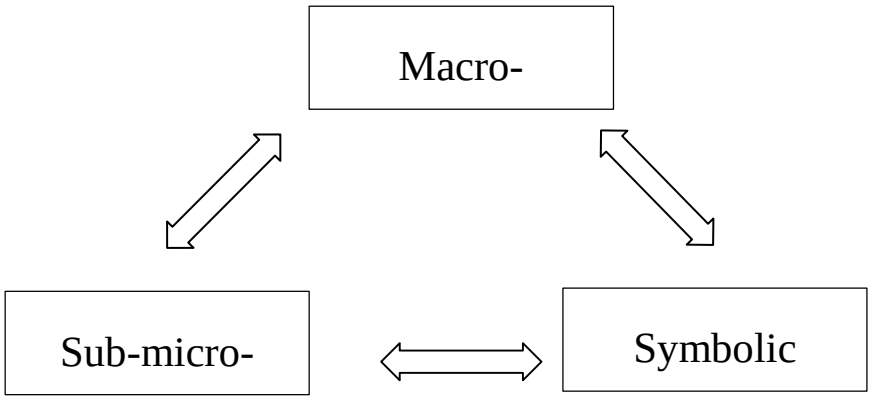


Figure 1. Johnstone's chemical triplet. The corners of the triangle represent the different levels or domains of knowledge between which students are required to be able to navigate (Tabor, 2013).

The macroscopic level represents the world that can be observed with our senses. At this level, students can manipulate matter and observe its chemical and physical properties. Students interact with chemistry at the macroscopic level every day when cooking, cleaning, and driving cars (Johnstone, 1982).

An obstacle to learning at this level is the vocabulary used to describe matter. Some terms and definitions have different meanings when used in day-to-day language when compared to scientific language (Garnett et al., 1995; Nakhleh, 1992; Tabor, 2013). As an example, in daily language, heat and temperature are used in similar context, but in chemistry, they are not the same. Heat is defined as the flow of energy from one substance to another due to differences in temperature; the amount of heat depends upon the amount of matter involved. Temperature is the average kinetic energy in a substance relative to some reference material and is not dependent on the amount of material (Nakhleh, 1992). These are subtle differences but can lead to student

misconceptions by creating different mental models of the concepts of energy and matter (Garnett et al., 1995).

The submicroscopic level describes the entities that constitute matter. The chemical nature of substances at this level explains why matter behaves the way it does at the macroscopic level (Johnstone, 1982). When atoms gain or lose electrons, they become ions and form ionic compounds. At the macroscopic level, ionic compounds are solids at room temperature, have high melting points, and conduct electricity when in a molten state or in an aqueous solution. Ionic compounds have these properties because of the ionic bonds formed at the submicroscopic level.

Matter and energy have different properties at the macroscopic and sub-microscopic levels. At the macroscopic level, matter has properties of mass, volume, and shape. Electromagnetic radiation (radiant energy) travels in waves that are described by the frequency, wavelength, and amplitude. At the sub-microscopic level, matter and energy can have both wave-like and particle-like properties; matter can have the properties of waves, and energy can have the properties of particulate matter. This dichotomy is called the wave-particle duality of matter and energy. The concept is a major departure from the observations students make at the macroscopic level and can exert a large demand on working memory, thus leading to cognitive overload (Garnett et al., 1995; Nakhleh, 1992; Tabor, 2013).

At the sub-microscope level, the representations of particles and their interactions are abstract. Students must create their own mental images from the theoretical models presented to them in textbooks, lectures, and diagrams. These abstract ideas have to be incorporated into the student’s preexisting knowledge base.

The symbolic level includes representations used to communicate with others in the chemical sciences (Nakhleh, 1992; Tabor, 2013). These representations include chemical formulas and equations that describe how matter reacts chemically, and the diagrams and graphs that represent theoretical models.

The impediment to learning at this level is the multiple symbolic representations of substances. Molecular structures can be represented in the form of condensed structural formulas, structural formulas, and ball and stick models. Examples of these different formulas are in Figure 2.

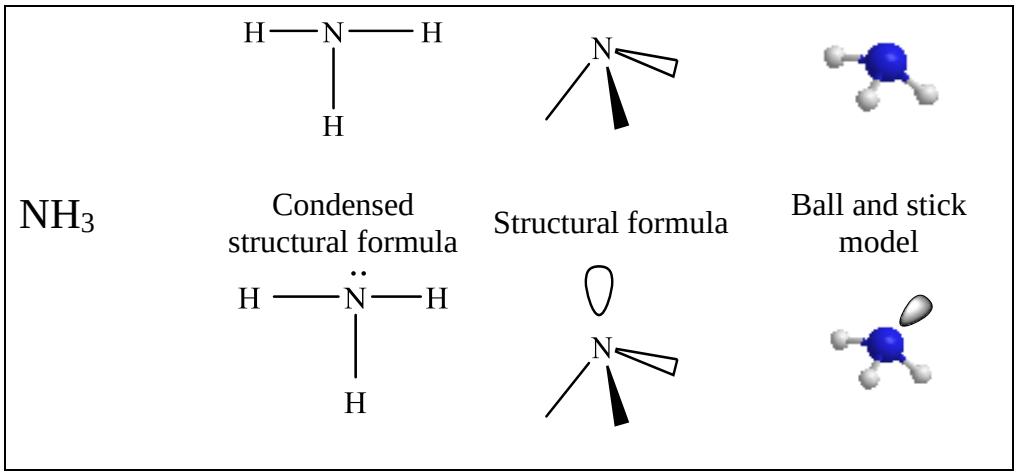


Figure 2. Representations of the molecule ammonia. The top row provides models of ammonia without showing the lone pair of electrons on the nitrogen. The second row includes the lone pair of electrons.

Each structure represents an ammonia molecule and depicts how the atoms are connected to each other, but each emphasizes a different aspect of the structure. The condensed structure model gives only the arrangement of atoms. The structural formula shows what atoms are connected and how they are orientated to each other. The ball and stick model is a three-

dimensional representation. Students organize their knowledge by course, topic, professor, etc., and using different representations can be difficult for them (Garnett et al., 1995).

In addition to the different molecular structure diagrams, each subatomic particle has several symbolic representations. A proton can be represented by ${}^{1}_{1}p$ or p^{+} or ${}^{1}_{1}H$. For a neutron, the symbols are ${}^{1}_{0}n$ or n^{0}. The electron can be represented by ${}^{0}_{-1}e$ or e^{-} and in nuclear chemistry as the beta particle ${}^{0}_{-1}\beta$, which should not be confused with the positron ${}^{0}_{+1}\beta$. Each symbolic representation for a particle is valid and can be used interchangeably. The choice of how to present information depends on the educational motive for the representation (Nakhleh, 1992).

Figure 3 is a visual representation of the shift between levels.

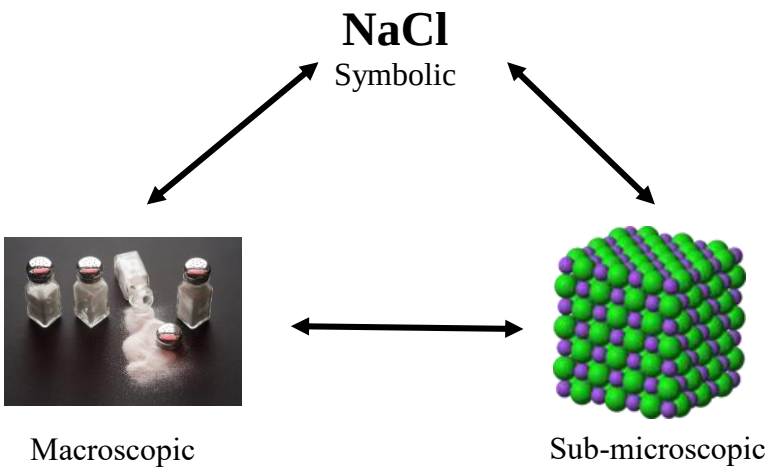


Figure 3. Example of the multilevels of chemistry knowledge. Sub-microscopic particles explain the properties of matter at the macroscopic level. Each level has its own symbolic language depicting the interactions of matter. Images from WikiMedia Commons (Benjah-bmm27 (Own work), 2012; Royan, 2006).

Chemistry is difficult because it has multilevels of knowledge with each level presenting challenges in vocabulary, abstract thinking, and symbolic language. Students have to be able to transfer between levels to understand the concepts and the theoretical models of chemistry that can result in working memory overload (Gabel, 1993; Tabor, 2013). The cognitive demand of the three levels of knowledge overloads the working memory of students and makes it difficult for students to learn chemistry.

External representations. For the discussions in this paper, external representations are the graphs, images, or diagrams used to explain a chemical phenomenon. They are created to help students understand abstract ideas and to decrease the cognitive demand of new information (Orgill & Crippen).

Chemistry uses certain formalisms when representing sub-micro particles and their interactions. If the students are not familiar with the format of these representations, the images meant to help students will instead increase the cognitive demand (Madden, Jones, & Rahm, 2010; Orgill & Crippen).

Costu (2010) compared test scores of algorithmic, conceptual, and graphical questions on the topics of solubility, chemical calculations, equilibrium, and radioactivity. Algorithmic questions require learners to apply formulae to answer a question, and conceptual questions require students to understand the ideas and relationships between the parts of a theoretical model. Graphical questions call for an understanding of the relationships between variables presented in a two-dimensional drawing. Costu found there was a significant difference in student scores for each type of question, with students scoring best on algorithmic, then conceptual questions, and the worst on graphical ones. Competency in one area of knowledge was necessarily repeated in the other knowledge areas.

When illustrations are used along with text, the student must be able to relate the information in the text and the information in the illustrations to each other and must incorporate the pieces of information into one coherent whole (Cheng & Gilbert, 2009; McTigue & Flowers, 2011). If text and diagrams are not integrated, it takes too much working memory to hold the information from the text while searching for the related information in the diagram. This is true when there is too much extraneous material in either the text or the diagram (Sweller et al., 1998).

Misconceptions. According to constructivism, knowledge is created in the mind of the learner through incorporating new knowledge into pre-existing, conceptual frameworks. The students interpret the new information based on their prior knowledge. If this knowldege is faulty, the process can lead to misconceptions. Students will selectively pay attention and pick and choose information that works best with their pre-existing knowledge. The result is that their construction of the abstract concepts are not what the instructor intended (Nakhleh, 1992).

Exposures

An exposure occurs when a student is presented a concept in the lecture portion of the course, the textbook, laboratory experiments, or homework problems. Concepts are the topics and subtopics of the Chem170/Chem 175 courses. In the chemistry lab, an exposure occurs when a student conducts an acid-base titration. In this case, the student is exposed to the concept of acid-base equilibrium at the macroscopic level. In the lecture class, an exposure could be a worked example of an algorithmic problem on calculating the pH of the acid. On a homework assignment, the exposure could be a conceptual question on the behavior of the sub-micro particles in acid-base neutralization reaction. These are all exposures in different locations and presented in different formats involving acid-base equilibrium.

The number of exposures and the location of the exposures are the predictor variables in this project. The research questions are answered by examining the effects of the predictor variables on the success of the students in the course.

Question Difficulty

The cognitive demand of an exam or homework question depends on the topic and the format of the question (Holme, Knaus, Murphy, & Blecking, 2011). The difficulty of a question can be classified by the complexity of the question or by the type of knowledge required to answer the question. Other procedures to classify question difficulty are based on type of learning needed and the cognitive skill level required to answer the question (Hartman & Lin, 2011; Holme & Murphy, 2011; Nakhleh, 1993b; Zoller & Tsaparlis, 1997).

Question type. The first method to be discussed is based on the learning process needed to find a solution (Hartman & Lin, 2011). The learning processes are recall, algorithmic, and conceptual.

Conceptual level questions require a student to evaluate what ideas are pertinent to a problem, understand the underlying concepts associated with the problem, and make the connections between the three levels of chemistry knowledge. Conceptual questions can be in the format of graphs or diagrams (Costu, 2010; Zoller, Lubezky, Nakhleh, Tessier, & Dori, 1995). Conceptual learning is sometimes associated with scientific reasoning (Deming, O'Donnell, & Malone, 2012).

The algorithmic level includes questions requiring the ability to use memorized procedures to find a numeric answer. Algorithmic questions usually have only one correct answer that is found by using a set of rules or formulas to answer the question. They are most likely questions that contain numerical quantities that have to be manipulated mathematically to

find the answer. This type of question can be solved successfully without any understanding of the concepts underlying the question (Nurrenbern & Pickering, 1987).

Recall questions are those that require students to remember memorized definitions or formulas, which are also called rote learning. This is considered the lowest level of learning.

In the following two figures, there is a conceptual question about the topic of limiting reagents (Figure 4). An algorithmic form of the question follows in Figure 5.

The equation for the combination reaction of nitrogen (●●) and oxygen (○○) is

$$N_2(g) + 2O_2(g) \rightarrow 2NO_2(g)$$

If the following molecular diagram represents the starting mixture,

which of the following is a representation of your product mixture?

a) b) c) d) e)

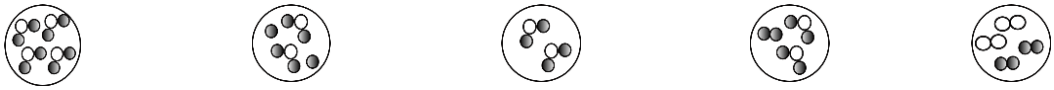


Figure 4. A limiting reagent conceptual question. From the balanced chemical equation and a cartoon of the starting material, the student must use the concept of the limiting reagent to choose the correct answer.

To solve the problem, the student must understand the ratio of reactants needed to produce products. In this case, the reactant ratio is one nitrogen molecule (●●) to two molecules of oxygen (○○).

The student has to evaluate the ratio to the amount of reactants to decide the limiting reagent. There are two nitrogen molecules and two oxygen molecules; therefore, there will be two molecules of nitrogen dioxide () formed and one molecule of nitrogen leftover. The answer to Figure 4 is letter "d." There is no process or algorithm to answer this question. The student must understand what a limiting reagent is and how to use this theory to answer the question.

In contrast, the following algorithmic question has a set process to find a numerical answer.

For the following reaction,

$$N_2(g) + 2O_2(g) \rightarrow 2NO_2(g)$$

A mixture contains 0.56 g of nitrogen (N_2) and 0.64 g of oxygen (O_2); which starting product is the limiting reagent?

Figure 5. Example of an algorithmic question. The student can convert the starting materials to moles and determine which is the limiting reactant.

The student uses an algorithm to convert mass to moles for each reactant, and then the stoichiometrically equivalent molar ratios of the reactants and products to decide which starting material is the limiting reagent. Each step uses a mathematical process, and the student can solve the problem without knowing the concept of limiting reagents (Nakhleh, 1993a; Nurrenbern & Pickering, 1987).

Cognitive skills. Another measure of question difficulty is the idea of cognitive skill levels. Some types of learning require higher cognitive skills to master than others (Zoller & Tsaparlis, 1997).

Lower-order cognitive skills (LOCS) requires the application of previously learned information in a systematic manner to familiar problems (Lewis & Smith, 1993). In this research project, the classification of lower-order cognitive skills includes definitions, recall of facts, and algorithmic problems where there is one correct way to find an answer.

Higher-order cognitive skills (HOCS) require students to relate previously learned knowledge to new, unfamiliar situations. It could be in the form of manipulating formulas or multiple substitutions to derive a method to calculate the answer to a question (Lewis & Smith, 1993; Thompson, 2011). Another type of higher-order thinking questions is the application of theories in unfamiliar questions.

The classification of higher order or lower-order thinking is not uniform. Problems that might seem difficult to one person and require higher-order cognitive skills might not be difficult for another and require only lower-order cognitive skills (Lewis & Smith, 1993). The classification relies on the previous knowledge a student has. Thompson (2011) included the issue of familiarity when deciding the cognitive skill level of a chemistry problem. He states the more a student practices a problem, the more it becomes lower-order versus higher-order thinking skills.

Bloom's Taxonomy. Bloom's Taxonomy of Learning Domains was developed by a committee of college and university examiners attending a conference of the American Psychological Association in Boston in 1948. The goal was to produce categories of learning that could be used to write and assess educational goals. It organized cognitive processes by

creating six levels of increasing cognitive complexity (Forehand, 2005). The first edition of the taxonomy appeared in 1956 (Choppin & Postlethwaite, 2014; Santa Barbara City College Student Learning Outcomes Project, 2015). Figure 6 gives the hierarchy of the six levels of the taxonomy.

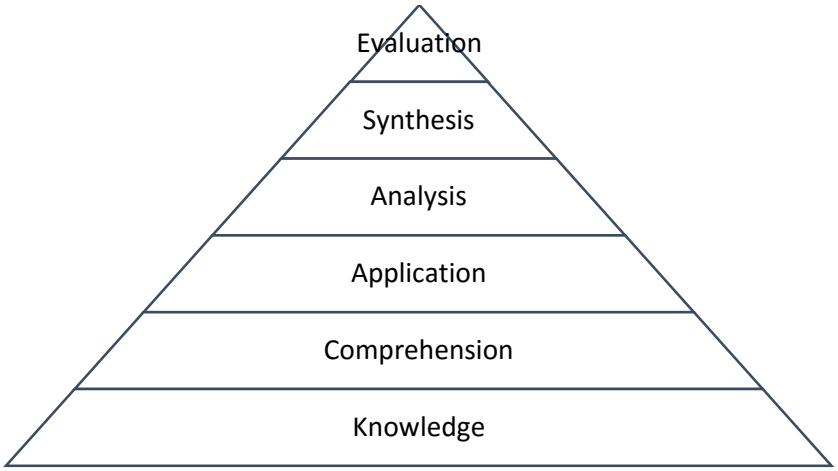


Figure 6. Bloom's Taxonomy Hierarchical Levels of Knowledge (Choppin & Postlethwaite, 2014).

Bloom’s Taxonomy has been revised and adapted to fit different needs of academic subjects and business and to include new theories from the field of cognitive development (Jones, Harland, Reid, & Bartlett, 2009; Thompson, 2011). Bloom’s Taxonomy is a familiar theory, and it is easy to use because of its simple structure that can be applied in varied areas. Bloom’s Taxonomy is used to design course content and to create assessments that test for the course objectives (Santa Barbara City College Student Learning Outcomes Project, 2015).

As the levels of Bloom’s Taxonomy progress from the lower to upper levels of the taxonomy, the cognitive processes become more difficult to master. Research performed by Tiemeier, Stacy, and Burke (2011) tested this idea by creating a modified Bloom’s Taxonomy

based on three levels: recall, application and analysis. Multiple-choice questions were created for each level and given to pharmacy students. The researchers found a statistically significant difference in percentage of correct answers in each domain (Tiemeier, Stacy, & Burke, 2011). Students generally must master a lower level of knowledge before they can begin to attempt to master a higher level of learning objectives.

One goal of this research is to investigate how the level of cognitive processing required by lab and lecture presentations and homework problems affects the success on exam questions of different levels of complexity. In an address to students and faculty, Saundra Yancy McGuire, Director Emerita of the Center for Academic Success and Retired Assistant Vice Chancellor, and Professor of Chemistry at Louisiana State University suggested introductory chemistry courses should not teach nor test above the analysis level of the taxonomy (S.Y. McGuire, private presentation, November 13, 2013).

A modified taxonomy was created to measure question difficulty. The original and the modified taxonomy are shown in Figure 7.

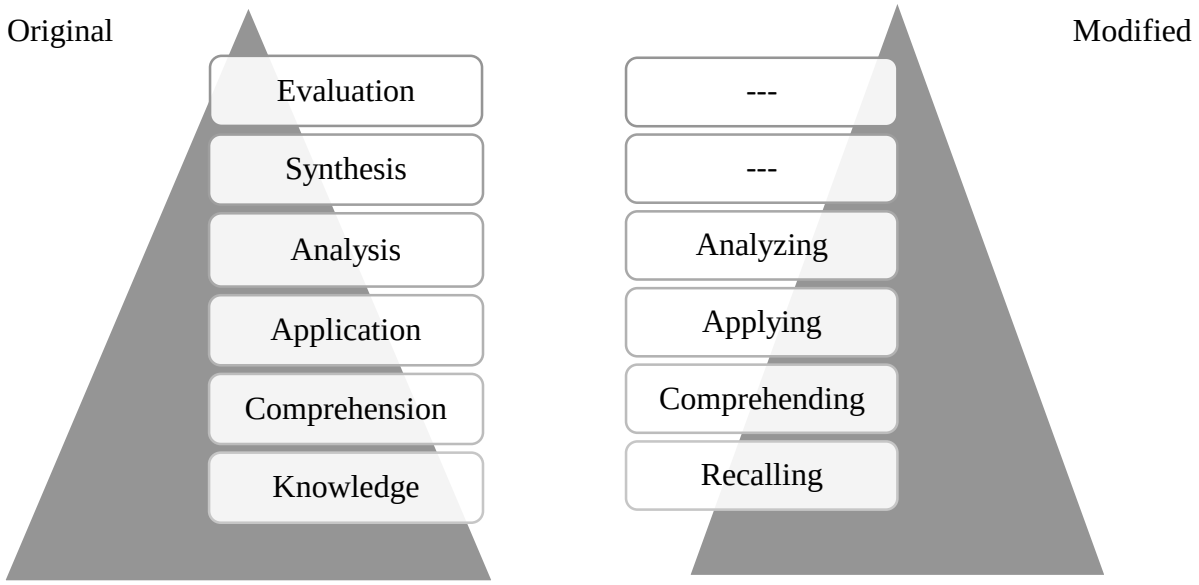


Figure 7. The 2011-revised version of the Bloom's Taxonomy and the taxonomy created for this research project. The modified taxonomy uses verbs to describe learning objectives. The two top-tiered levels of synthesis and evaluation are not included in the modified version (Dávila & Talanquer, 2010).

Recalling questions ask students to remember memorized material or facts. The ability to use simple equations would fall under this category. Recalling questions requires basic knowledge and must be mastered before a learner can achieve higher level (Jones et al., 2009). For example, the question, "What is the symbol for copper?" requires only the student to remember symbols for the elements.

Comprehending questions requires a student to demonstrate a basic understanding of the relationships between variables or the elements of a concept. A comprehension question also requires that a student understand how to do simple calculations and substitutions. These questions require mathematical answer.

Applying questions require the understanding of concepts and their interconnectedness to answer a question. The student is able apply equations to new or unfamiliar situations and to be able to derive the correct algorithm to answer the question.

Analyzing questions requires a student to evaluate the information and identify what is needed to find the solution to a problem. The student understands the concepts and the relationships between them.

Number of steps. The cognitive load theory describes the relationship between success rates on chemistry problems with the cognitive demand of a question and the number of pieces of information that must be processed in working memory. Sweller (1994) labeled these pieces of information as chunks or elements. He found that as the complexity and interactions between elements increases, the cognitive load on the short-term memory of a student's ability to solve a problem decreases.

In 2011, Hartman and Lin used a similar construct but designated the term "steps" as a way to measure the cognitive load placed on short-term memory. In their research, a step is an independent process needed to solve the question. They measured the difficulty level of exam questions by the number of steps needed to solve the problem. As the number of steps increased, there was a corresponding decrease in the exam success of their students. The number of steps was determined by the shortest route needed to answer the question (Hartman & Lin, 2011). Figure 8 presents the different steps required to name compounds.

CaO	**$Mg(OH)_2$**	**$FeCl_3$**	**NO_2**	**CH_3OH**
What type of compound is this?	What type of compound is this?	What type of compound is this?	What type of compound is this?	What type of compound is this?
Ionic	Ionic	Ionic	Molecular	Molecular
Is the metal a transition metal?	Is the metal a transition metal?	Is the metal a transition metal?	Name the first element using a prefix if there is more than one atom.	What is the functional group?
No	No	Yes	One atom of nitrogen: N	Alcohol: ending is –ol.
Write the name of the anion and cation.	Write the name of the cation and the polyatomic ion.	Determine the ionic charge of the metal from the cation.	Name the second element using a prefix if there is more than one atom.	How many carbon atoms?
calcium oxide	magnesium hydroxide	Calculate cation charge from anion.	Two atoms of oxygen: dioxide	One: first part is methyl.
		Write the name of the compound.	Write the name of the compound.	Write the name of the compound.
		iron(III) chloride	nitrogen dioxide	methanol

Figure 8. The number of steps to name organic and inorganic compounds. Each step adds to the cognitive load of a student.

The first decision students must make is what type of compound they are naming. Alkaline and alkaline earth (representative) metals have one system of naming compounds. Transition metals have a similar process, but the key difference is students must put the ionic charge of the metal in the name. Polyatomic ions can be found in ionic compounds with both transition and representative metals. Molecular compounds use prefixes to indicate the number of atoms in the molecule. Organic compounds use the functional groups and the number of carbon atoms to name them.

Topics. Students believe chemistry is difficult (Bauer, 2008; Sirhan, 2007). In addition to the three levels of knowledge represented by Johnstone's triplet and the problem of misconceptions, there are certain topics that students have more difficulty with than others.

Topics with large amount of interacting elements that must be processed simultaneously are considered difficult to understand. Learning a subtopic without learning its connection to the other subtopics of the concept prevents the student from understanding. Cognitive load is determined by an interaction between the nature of the material being learned and the expertise of the learner (Sweller et al., 1998).

Sirhan (2007) published a list of topics considered difficult based on chemical education research; these topics include thermodynamics, electrochemistry, and solution chemistry. Due to the complexity of the topic, the cognitive demand exceeds the working memory of students (Gulacar, Overton, Bowman, & Fynewever, 2013; Johnstone, 1983; Sirhan, 2007).

Cognitive scale relationships. Measuring the success of a teaching pedagogy requires a method to determine the cognitive demand of exam questions (Smith, Nakhleh, & Bretz, 2010). The different scales evaluate difficulty based on cognitive processes.

Figure 9 illustrates the relationships between the different levels of the classification methods.

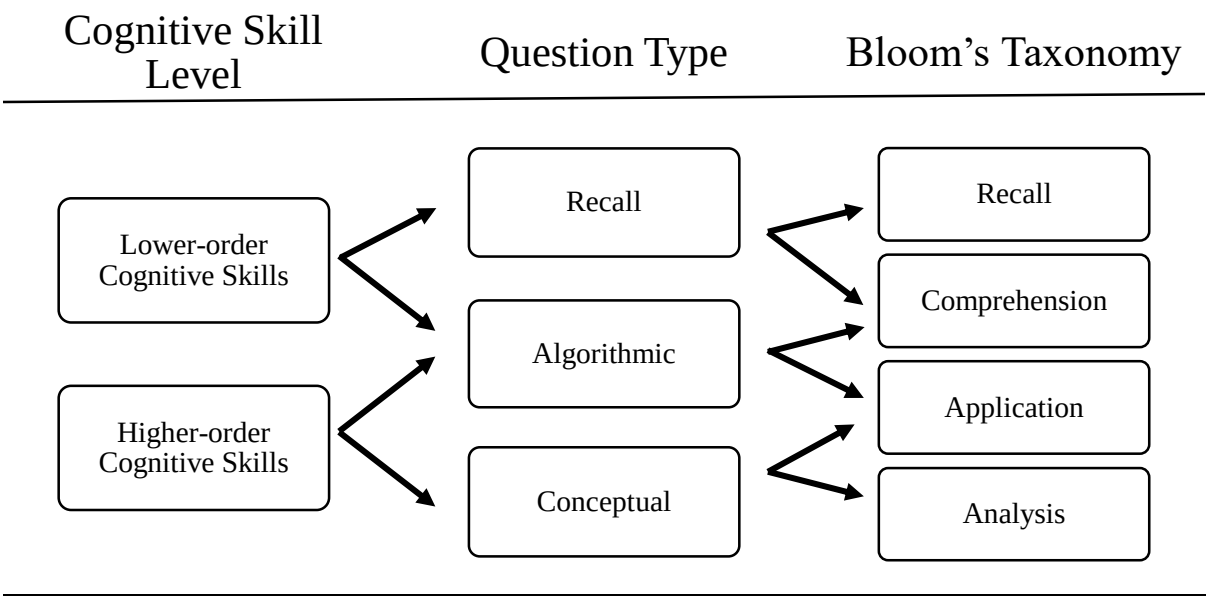


Figure 9. Relationships between the different levels of the difficulty scales.

Lower-order cognitive skills encompass recall questions and simple algorithmic problems. The higher-order cognitive skills contain more demanding algorithmic problems and conceptual knowledge. The rankings of question type are divided into the levels of Bloom's Taxonomy, giving further detail of question difficulty. The levels become more specific in their classification methods.

Pedagogy

Instruction methods based on constructivist theories focus on students' interactions with their environment. Cognitive load teaching pedagogies emphasize methods to decrease the cognitive load on the working memory of students.

Cognitive load theory. The cognitive demand of a task affects the success of a student at solving problems. Teaching pedagogies based on the cognitive load theory focus on decreasing the cognitive demand placed on students.

Modeling. Modeling is a process where the instructor describes the strategy and the motive for the choice of variables and the processes used to solve a problem. The instructor thinks out loud as he or she works the problem and reduces the each sub-goal into smaller steps that reduces the cognitive load on the student (Intel Teach Program, 2012). When the instructor explains the choice in variables, it reduces the cognitive load. Students do not have to search for and decide on the appropriate variables to use; this allows students to focus attention on the process (Kirschner, Sweller, & Clark, 2010).

Goal-free strategy. A means-ends problem solving strategy is a search strategy that creates a high-cognitive demand on a student's working memory (Ayres, 1993). When a student has been assigned a problem with a specific goal, the student looks at the beginning state of the problem and subtracts that from the goal state, and then finds a process to reduce the difference. The student must use working memory to keep all these elements in working memory at the same time. The student is less likely to focus on building a schema and will not be able to solve similar problems (Sweller et al., 1998).

A research project by Sweller and Levine (1982) found that the more students focused on the end goal, the less they learned of the problem structure. One method to prevent this outcome is to assign goal-free problems. In this type of problem, a non-specific goal is given and the students must discover the relationships between the variables associated with the concept.

A student using a means-end strategy focuses on the final goal and works backwards to find a method to achieve that goal. The problems with this strategy is students do not focus on the meaning of the variables and their interactions.

Using a goal-free strategy, a student may be asked to find all the relationships between the variables of a topic. The student concentrates on the different relationships, which leads to

the formation of schemas. These schemas can be used to solve other problems (Sweller & Levine, 1982).

Worked examples. When a student is involved in solving a problem, most of the working memory resources are allocated to the process of solving the problem, which means less space is available for learning. The learner searches for the correct formulas and relationships between the variables. Students have difficulty in determining what information is important and what is not (Lee & Anderson, 2013).

A worked example is a step-by-step process that provides a blueprint to solve a problem. A problem statement is given as are the variables are important and the steps to solve the problem (Crippen & Brooks, 2009). A worked example reduces the cognitive demand and reduces the drain on working memory (Lee & Anderson, 2013).

Worked examples are especially helpful for novice learners. Working memory space is released from ineffective searching for information and can be used to learn the reasons for the problem-solving steps (Lee & Anderson, 2013).

Constructivist-based pedagogy. The constructivist-based pedagogies center on student-centered learning. The pedagogies associated with constructivism center on learners' construction of knowledge by interaction with their physical environment. Social constructivism emphasizes the significance of social interactions in the learning of chemistry (Garnett et al., 1995).

Collaborative and cooperative learning. Collaborative and cooperative learning are small group techniques. The difference is that in collaborative learning students receive a group grade and in cooperative learning each student gets an individual grade (Gasiewski, Eagan,

Garcia, Hurtado, & Change, 2012). Peer-led team learning, inquiry labs, and problem-based learning are all small group methods.

Inquiry learning. Inquiry learning is a student centered teaching method based on personal constructivism (Garnett et al., 1995). The student learners are responsible for directing their own investigation of a problem. The problem could be designing a lab experiment, investigating solutions to an environmental problem, or any activity where the students have the responsibility of designing an investigation and evaluating the results (Keselman, 2003).

Inquiry learning mimics the scientific method. Students complete each stage of the scientific method by formulating hypotheses, designing experiments to test them, collecting information, and drawing conclusions (Kirschner et al., 2010). Inquiry learning increases critical thinking and problem-solving skills.

Problem-based learning. In problem-based learning (PBL), students are presented with an open-ended problem; together with a group of fellow students they plan and enact that plan to find a solution to the problem (Contributors; Sandi-Urena, Cooper, Gatlin, & Bhattacharyya).

The benefit of PBL is to engage students at a higher level of learning by emphasizing the comprehension of concepts and not the memorizing of facts. It is a form of inquiry learning, but the students are given a specific problem to work on; it includes a social constructivist component because of the small group learning.

Peer-led learning. Students meet in together in groups of eight to ten students outside of class once a week to work together on homework or worksheet problems. The peer leader, a more experienced student in the subject, is there for guidence if students have difficulties finding solutions (Cracolice & Deming, 2001; Gasiewski et al., 2012; Hockings, DeAngelis, & Frey, 2008).

Peer-led learning is based on the theories of social constructivism. The students work in small groups, which allows them to discuss scientific theories and how they can be applied to the problems they are facing. The role of the instructor is to help students connect their current conceptions with those accepted by the scientific community (Garnett et al., 1995).

Learning Environments

The location of an exposure to a concept or a fact is important. Different learning environments provide opportunities for teaching methods that fit the various learning preferences of students. The Visual-Aural-Read/Write-Kinesthetic model of learning styles (VARK) is a method to define the learning styles of a learner (Bretz, 2005; Fleming, 2001). There are four learning environments in this research project, and each environment accentuates at least one of the learning styles. For example, the classroom environment emphasizes auditory and visual learning styles. The laboratory environment emphasizes the kinesthetic learning, and the textbook and homework environment accentuates the reading and writing learning style. Each educational environment will be described and the importance of the setting to learning will be discussed.

Laboratory environment. According constructivism, learners construct knowledge by interacting with their environment. (Cracolice, 2005; Crippen & Earl, 2004). The chemistry teaching lab has many opportunities for students to incorporate new knowledge through observations and interpretation of their observations (Nakhleh, 1994). Students are able to have hands-on experiences, operate in small groups, and develop their cognitive skills through writing lab reports (Tsui, 2002).

Students have difficulty transferring between the different levels of the chemistry triplet (Tabor, 2013). The chemistry laboratory provides an environment where students can make the

connections between these levels of chemical knowledge. Students use symbolic language to represent chemical reactions, manipulate the chemicals and lab apparatus, observe, and relate these observations to the interactions of the sub-microscopic particles.

Writing is one way to develop students' critical thinking skills (Moore & Rubbo, 2012). Writing assignments consists of lab notebooks and lab reports. In their notebooks, students have to decide what data is important to record for later use. For lab reports, students must organize their thoughts in a coherent manner, report their data in the correct format, and use their conceptual understanding to produce a meaningful and rational discussion. Writing helps students understand or make sense of the lab experiment (Johnstone, 1983).

Appropriate teaching pedagogies to use in the lab environment are cooperative and collaborative learning, problem-based learning, and inquiry learning. The learning style of kinesthetic is present through the manipulation of chemicals. Writing lab reports supports the reading and writing style of learning. The visual learning style is supported by the observations of reactions, and the discussions between peers support the aural learning style.

Homework. Cognitive theories suggest the familiarity with instructional materials decreases the processing space and increases the number of elements that can be processed simultaneously (Cowan, 2014). When a student becomes more familiar with a concept, the process becomes more automatic and does not place as much demand on working memory. Homework assignments are one method to increase schema complexity and reduce the number of steps. The time spent doing homework outside of class is positively correlated with achievement (Richards-Babb, Drelick, Henry, & Robertson-Honecker, 2011).

Online homework. Graded homework assignments have shown to have a positive influence on student success across all subjects and grade levels (Fynewever, 2008). Immediate

feedback is beneficial regardless of learning situation (Buzzetto-More & Ukoha, 2009). Online homework assignments provide both immediate feedback and graded assignments.

A survey created to measure students' attitudes toward online homework indicates that students felt online homework promotes more consistent and beneficial study habits by increasing the amount and frequency of studying and reducing cramming for exams. Students felt online homework was worth the effort, relevant to the material presented in class, and challenging (Richards-Babb et al., 2011).

Textbook. Worked examples provide students with sample exercises to help them understand what is needed to solve a particular problem (Lee & Anderson, 2013). Sample exercises in chemistry textbooks provide step-by-step instructions on how to solve a problem and help the student focus on the problem structure, which they then can apply to new problems (Lee & Anderson, 2013). Textbooks are a source of worked examples.

In addition, theoretical models in chemistry are abstract, and for students at the concrete operational or transitional stage, abstract concepts are difficult to understand. Textbooks can help these students by using representations of abstract models to increase conceptual understanding (Mayer, 2005).

Electrochemistry is a topic with many abstract components such as the flow of electrons between the parts of a voltaic cell to produce an electrical current. Diagrams and other representations of this process can help build abstract models to increase student understanding (Orgill & Crippen). Figure 10 is a representation of the parts and processes of a voltaic cell.

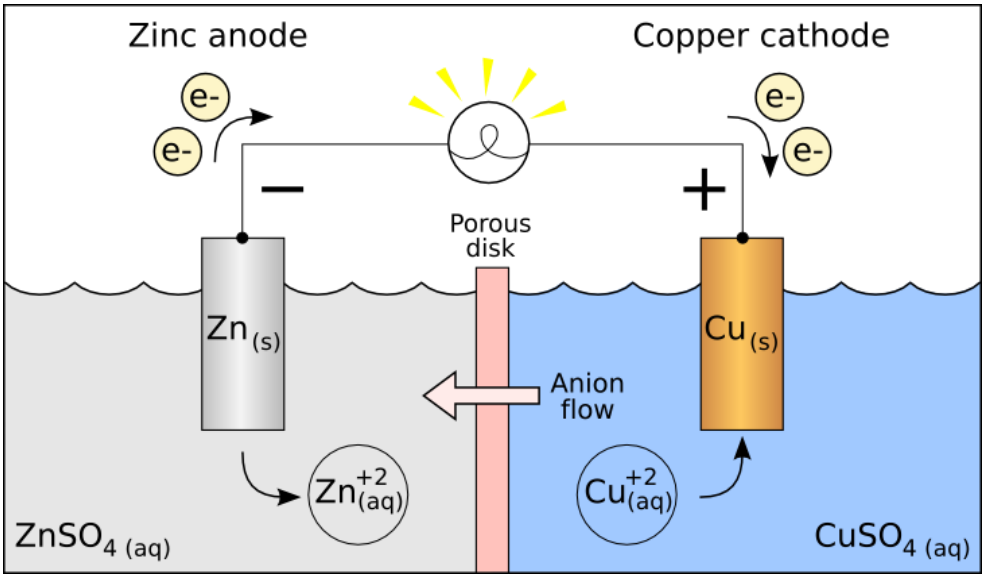


Figure 10. Voltaic cell. This representation of a voltaic cell is similar to ones found in textbooks It shows how the components interact to produce an electrical current, and this could increase the conceptual understanding of electrochemistry (Ohiostandard at en.wikipedia, 2016, February 11).

In this voltaic cell, there are four processes occurring at the same time: the reduction of the copper(II) ion at the cathode, the oxidation of zinc at the anode, the flow of electrons from the anode to the cathode, and the movement of ions to keep the solutions electrically neutral. To create an abstract model of a voltaic cell, the student needs to understand each process, and how they interact with each other. Illustrations and diagrams can aid a student's conceptual understanding of the different processes (Mayer, Bove, Bryman, Mars, & Tapangco, 1996; Sweller et al., 1998).

Lecture. Professor Heather Desaire taught the lecture courses for all four semesters of this research project. In addition, she consulted on the laboratory component of the course. The following descriptions of lecture as a teaching method are based on the courses she taught.

The lecture professor used different teaching strategies to produce both professor-oriented and student-oriented teaching pedagogies. The professor used the expository lecture as the main teaching method, but she included active learning and student-centered methods to increase student involvement and interest.

The benefits of lecture as a teaching method includes its ability to dispense large amounts of information in an efficient manner, help organize thinking about a subject, and promote problem solving skills (Di Leonardi, 2007). Professor Desaire used teaching to increase student-centered learning during the lecture through small group problem solving with the groups presenting their solution to the class. In-class, graded quizzes gave the students immediate feedback on their understanding. Oral questions presented in the lecture promote thinking even when not asking for a response. Classroom demonstrations of chemical phenomena were performed to help students see the relationship between the macroscopic and sub-microscopic worlds and to heighten their interest. The professors used small breaks to enable students to refocus (Smith, 2006). These are all methods to increase student focus.

The lack of prior knowledge can be detrimental to a student's ability to solve problems and understand concepts (Rittle-Johnson, Star, & Durkin, 2009). Lectures can help provide missing information from a student's conceptual framework by modeling or by worked examples. These methods help students to detect any missing knowledge they might need.

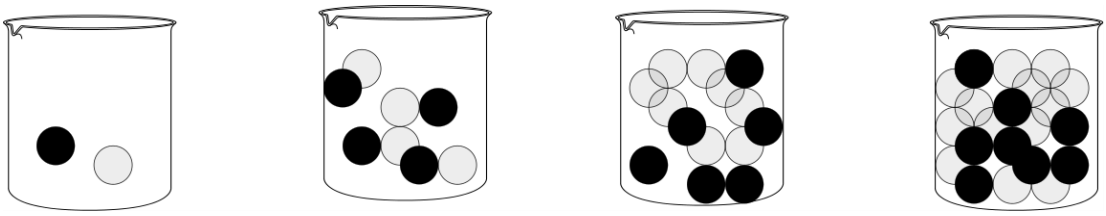
External representations can be used in the lecture portion of a class. Using diagrams to represent the behavior of sub-microscopic particles and graphs to demonstrate the relationships

between variables allows students to have both visual and auditory presentations to reinforce each other.

Figure 11 is an example that Professor Desaire presented to the students in the Spring 2014 Chem 175 course. In this example, she uses external representations to illustrate the differences in first and second reaction orders. She integrates the visual and numerical quantities so the students do not have to search for the relationships between the different representations.

For the general reaction of aA + bB → cD + dD, the rate equation is Rate = k $[A]^x[B]^y$

Visual representation of the change in concentration of products in first and second overall reaction orders:



Numerical quantities showing the change in concentration of the products:

● 1 M	4 M	6 M	8 M	1st Order
○ 1 M	4 M	9 M	16 M	2nd Order

Figure 11. External representation of the theory of kinetics. The two illustrations are integrated into one presentation to reduce the cognitive demand on working memory.

Teaching pedagogies suitable to the lecture are modeling and worked examples. If the configuration of the room allows it, small group learning is appropriate. Learning styles are aural and visual.

Research Hypothesis

The research hypothesis states the number of times a student is exposed to a chemical concept, the pedagogy used to present the concept, and the learning environment of the presentation will increase a student's success rate on exam questions.

- Does the number of times a student is exposed to a chemical concept affect the success rate of that student on exam questions covering the same topic?
- Does the location where the student is shown a concept affect the success rate of that student on exam questions testing the same topic?
- Does the method in which a concept is presented affect the success rate of students on exam questions?
- Do topics have different levels of difficulty, and if so, does the number of exposures decrease the difficulty of the more challenging topics?
- Do the learning environment and the number and type of exposures affect students differently at different cognitive levels?

Methods

The following sections describe the methods used to develop and quantify both the independent variables and dependent variables. Exposures, the independent variables, were coded by topic and by type of presentation. This process will be described and the reasoning for each coding decision will be given. The process used to measure the success of each treatment will be described.

An “exposure” is classified as any time a concept is presented to a student. It can take the form of a homework problem, a lecture presentation, or a lab experiment. The number of exposures, the type of exposure, and the location of an exposure are all treatment variables. Measuring instruments and research methods were developed to count, classify, and determine the location for each exposure.

Student Sample

The sample population consisted of students enrolled in a two-semester sequence of an introductory chemistry course developed for students majoring in a chemical science. The majority of the students were freshmen students with a declared major in chemical engineering.

The 2012-2013 school year was the first time this course was offered, and there was one course section offered each semester. The first course sequence (Fall 2012-Spring 2013) had a limit of 100 students for the sequence of courses. In the second year of the research (Fall 2013-Spring 2014), the enrollment cap was removed, and the number of students in course was determined by the number of lab spaces available, which was approximately 135 students for both semesters. The lecture portion of the course met one hour three times a week, and the lab portion met once a week for five hours.

Number of Exposures

The type of exposure, the location of the exposure, and the number of exposures are the independent variables. A method of counting each independent variable was developed and was applied to all of the independent variables.

Coding by concept. One unifying feature of this study was the subject matter presented in the courses. Introductory chemistry courses include a broad range of topics, so the first step was to organize the exposures by topic. The main topics were given numbers to distinguish them from the other topics. The main topics corresponded to the topics presented in the course textbook (Brown, LeMay Jr., Bursten, Murphy, & Woodward, 2012). The primary topics were further divided into subtopics, and each subtopic was labeled with a lower-case letter to denote it in the coding system.

Main topic is atomic structure. → 6(c) ← The subtopic is electron configuration.

As an example, the main topic of atoms and their structure (designated as topic 6) was divided into the subtopics of (a) the wave-particle duality of matter, (b) quantum mechanics, and (c) electron configurations. A homework problem about the electron configuration of a halogen would be designated as 6(c).

Coding by type of exposure. A secondary research question asks whether exam questions with different levels of difficulty are affected by the number and type of exposures in dissimilar ways. A classification method based on the strategies employed by Zoller et al. (1995) and Nurrenbern and Nakhleh (1987) was created to classify the types of presentations. The three types are recall, algorithmic, and conceptual presentations.

Recall presentations are definitions and material that require memorization. Algorithmic presentations are those that show how to solve mathematical problems. Conceptual presentations are those that explain the theories and concepts of the subject matter.

When an exposure ended and another began was defined by the topic and type of presentation. If the topic or subtopic of a presentation changed, it was coded as a new exposure. If the type of exposure changed (e.g., from recall to algorithmic type), a new exposure was started. This process is shown in Figure 12.

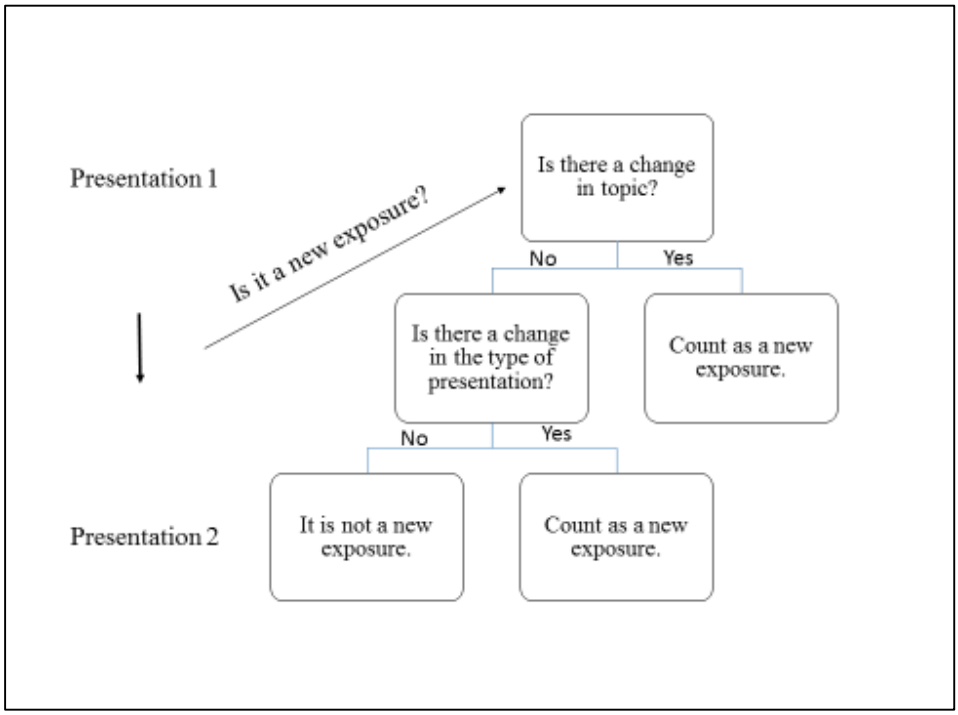


Figure 12. Process of classifying a presentation as a new exposure. The method uses the change in topic, subtopic, and presentation type to decide when there is a change from one exposure to another.

As an example, in a lecture presented on March 14, 2014, Professor Desaire started the class with an example problem on how to make a buffer. The presentation was coded as an algorithmic presentation, and the topic was aqueous equilibria. The subtopic was the common ion effect. The next example problem in the class was how to calculate the pH of a buffer solution, which is also an algorithmic problem related to aqueous equilibria and the common ion effect. Because there was no change in topic, subtopic, or question type, these two presentations were counted as one exposure.

In another lecture, the instructor discussed the topic of chemical thermodynamics, and the subtopic was the first law of thermodynamics (Wedderburn, Bililign, Levy, & Gdanitz, 2006). It was coded as a presentation on the topic of chemical thermodynamics and the subtopic of the first law of thermodynamics. When the presentation changed from the subtopic of the first law to the subtopic of entropy, the switch in subtopic was coded as two different exposures.

Coding by location of exposure. The location of each exposure was an independent variable. Each exposure was coded using the process presented in Figure 12.

Lecture. Lectures were held in large lecture halls with a chalkboard or a white board, a document camera, and both an overhead and a mounted projector. The researcher attended all lectures. The lecture notes were dated according to each class session. Each lecture period was coded for the topics and types of presentation. Each lecture period was considered independent from each other when counting the number of exposures.

Homework. Homework problems were from the online homework program Mastering Chemistry published by Pearson Prentice Hall (12th edition, 2012). Most homework assignments focused on one topic. Assignments were aligned with the topics as they were presented in the lecture. The number of questions on each homework assignment ranged from ten to twenty-four,

and the assignments were due weekly. The average of the homework grades counted for 10% of the student's final grade. Each homework question was coded according to topic and subtopic in addition to type of presentation.

Laboratory. The lab for the Chem 170/175 course sequence consisted of weekly lab sections with approximately twenty students in each section. Students were organized into small groups of three and were responsible for how their group operated. Each week a different laboratory experiment was assigned, and all lab sections conducted the same experiment each week. Assessments for laboratories were quizzes, lab reports, worksheets, and lab notebooks graded by the teaching assistant. The researcher was a graduate teaching assistant for all four semesters of this study.

Each lab was analyzed to determine the topics and subtopics presented in the lab. If the topics in the lab were presented after the exam testing for that topic, it was not considered an exposure.

A laboratory experiment on molecular geometry was divided into sections A, B and C. In section A, students were asked to sketch pictures of different balloon models representing molecules, which would be considered a lower level of cognitive demand. For section B, they were asked to draw the Lewis structure of several different molecules and to give both the molecular geometry and shape for each. In this instance, the type of exposure would be algorithmic because it is a systematic process used to determine the Lewis structure, molecular geometry, and the molecular shape. In section C, the students were instructed to draw a three-dimensional model of several molecules. This is considered a conceptual question. The student has to bring together all the concepts of bonding and the parts of molecular geometry. Since each

section of this sample project was focused on a different presentation type, each section was counted as one exposure for a total of three exposures.

Textbook. The textbook used in this course was *Chemistry: The Central Science*, 12th edition by Pearson, Prentice Hall (2012). The professor urged students to use their textbooks as a resource to help with homework problems as source of information to help with concepts presented in class and those concepts they were to learn on their own. To determine the number of exposures, chapters in the textbook were grouped by topic and subtopic. Each sample exercise in a subtopic section was counted as an exposure.

Classification of Question Difficulty

The level of difficulty of exam questions depends on the topic of the question or the type of question being asked. The difficulty of each exam question was determined by developing methods based on the four different schemas discussed in the introduction. Each approach was modified and operationalized to match the abilities of freshmen chemistry students. Each exam was coded using the four schemas, and following that, each exam was coded twice at different time intervals. The consistency of the two different ratings was tested using an intra-rater reliability analysis. The analysis of test results are in the results section.

For each method of determining question difficulty, the categories were operationalized. The definitions and examples for the different methods are given in the following sections.

Number of steps. Both algorithmic and abstract methods were used to determine the number of steps required to answer an exam question. The algorithmic method used mathematical operations and numerical values to solve the problem. The abstract method used an outline of the number of steps needed to solve the problem. Each approach was applied

separately at different times to each exam question and compared to each other. Any discrepancies were examined and then resolved by doing each process again.

The algorithmic and abstract methods should produce the same number of steps. To assess the coding method, the researcher coded the multiple choice questions found in the Hartman and Lin (2011) paper. The validation of coding results were comparable to the results obtained by Hartman and Lin (2011).

An Intraclass Correlation Coefficient (ICC) was conducted on the number of steps determined by Hartman and Lin to the number of steps calculated by the mathematical method and to the number calculated by the abstract method. The reliability between the methods is evaluated by Cronbach's alpha. The results of the analysis show a very good reliability for the Hartman-Lin and mathematical method (Cronbach's $\alpha = .83$); for the Hartman-Lin and the abstract method Cronbach's $\alpha = .77$ is considered good reliability.

Question type. Type of presentation and the difficulty level of a question are based on the same schema. The difference between presentation and difficulty in this research project is that the former relates to the type of presentation whereas the latter measures how difficult the question is for the students.

Recall questions require the retrieval of information stored in short-term memory. This is rote learning. Questions that are answered by using memorized information are classified as recall questions.

Algorithmic questions require the use of memorized procedures to find an answer, and these procedures can be used to solve similar questions without modification. For instance, the process to find the percent composition of carbon in glucose is the same process used every time

the percent composition of a substance is requested. In addition, questions that use equations with few variables or limited substitutions are classified as algorithmic.

Conceptual questions require students to understand and combine concepts and theories of a subject to determine the answer to a question. This includes using two or more equations with associated variables and make substitutions to find the answer. In addition, questions that require students to apply known procedures in unfamiliar situations are considered conceptual questions. The decision on whether a situation is new or unfamiliar depends on the homework assignments and lecture presentations.

Each exam question and homework problem was classified as one of the above categories. In the following, a sample question from each category is given and the reason for its classification.

Recall questions:

- Question Example: Which of the following has the correct symbol for the element?

a) P, phosphorus

b) Po, potassium

c) Cu, copper

d) Mg, manganese

- Reason: The student must remember the factual or memorized information.

Algorithmic question:

- Question Example: How many moles of glucose ($C_6H_{12}O_6$) are in 22.6 gram sample?

a) 0.12 mol

b) 4.07×10^3 mol

c) 0.29 mol

d) 22.6 mole

- Reason: The student must use the correct process to determine the number of moles of a substance when given the mass. This same process is used in similar situations.

Conceptual question:

- Question Example: Which species would have the highest lattice energy?

a) KCl

b) CaO

c) NaCl

d) MgO

- Reason: The student must understand the relationships between ionic charge and radius and their effect on lattice energy.

Recall questions are generally considered the least cognitively demanding types of questions, and conceptual questions are the most demanding for students (Lord & Baviskar, 2007; Smith et al., 2010; Zoller & Tsaparlis, 1997). The difficulty of algorithmic questions falls between the two categories of recall and conceptual.

Cognitive skill level. Lower-order cognitive skills (LOCS) require students to recall memorized information and to use equations with few variables that express simple relationships and require few substitutions to find an answer. LOCS can be recall or algorithmic questions as classified by question type.

Higher-order cognitive skills (HOCS) are those that require a student to make connections between concepts and to apply these relationships in a new setting. For example, making substitutions using formulas representing complex relationships to determine the answer.

LOCS

- Question Example: For which of the following reactions are ΔS positive?

a) $2NO(g) + O_2(g) \rightarrow 2NO_2(g)$

b) $N_2(g) + 3H_2(g) \rightarrow 2NH_3(g)$

c) $CaCO_3(s) \rightarrow CaO(s) + CO_2(g)$

d) $3C_2H_2(g) \rightarrow C_6H_6(g)$

- Reason: The student must remember the general rules of when entropy increases or decreases.

HOCS

- Question Sample: For the reaction: $3O_2(g) \rightarrow 2O_3(g)$, ΔH is 146 kJ/mol and ΔS is -252 J/K•mol. This reaction is ___.

a) spontaneous at all temperatures

b) nonspontaneous at all temperatures

c) spontaneous only at low temperatures

d) spontaneous only at high temperatures

- Reason: The student must understand the relationships between change in enthalpy and entropy, temperature, and Gibbs free energy.

Each question is about entropy and neither requires a numerical answer. The difference between the two questions is the LOCS question requires the student to recall the circumstances where entropy increases, whereas the HOCS question requires the student to understand the relationships between variables of entropy, enthalpy, temperature, and Gibbs free energy. The student must link the variables correctly to answer the question.

Modified Bloom's taxonomy. Bloom's taxonomy categorizes question difficulty by the cognitive process needed to solve a problem. The taxonomy was modified to define the levels by scientific processes. In addition, the top two tiers of the taxonomy, evaluating and synthesis, have been removed.

The definition of each of the modified levels, examples of exam questions representing each level, and the reasoning for the classification follow.

Recall Question: The student must remember factual or memorized information.

- Question Sample: All atoms of an element have the same _____.

a) neutrons

b) electrons

c) electrons and neutrons

d) protons

- Reasoning: Student remembers that all atoms of the same element have the same number of protons, and that the number of protons determines the element.

Comprehension Question: The student demonstrations their understanding of equations that represent the simple relationships between variables by using them appropriately.

- Question Sample: The frequency of light emitted from a from a red brake light is 6.0 x 10^{14} s^{-1}. What is the wavelength (λ) of the light in m?

a) Determine equation to use.

b) Make substitutions.

c) Solve for λ.

- Reasoning: The student must recall the equation that relates the wavelength (λ) and the frequency (ν). A simple substitution is made and the equation solved for the unknown.

Application Question: The student is able to apply equations to new or unfamiliar situations and to be able to derive the correct algorithm to answer the question.

- Question Example: Which of the following substances will have the shortest wavelength when the substance is traveling at 100 cm/s?

a) marble

b) airplane

c) the planet Mars

d) uranium atom

- Reasoning: The student must make the connection that the wavelength is inversely proportional to the mass so the object with the greatest mass will have the shortest wavelength.

Analysis Question: The student is able to combine concepts to find the solution to a question and can evaluate what information is needed to answer the question.

- Example Question: Nitrogen oxides (NO_x) are a major component of air pollution. After analyses of an air sample, it was determined one oxide was 30% by mass nitrogen. Which of the following could be the compound?

a) N_2O

b) NO_2

c) N_2O_4

d) NO_2 or N_2O_4

- Reasoning: The student must understand the difference between molecular and empirical formulas, how to use subscripts to find the mass and then, how use moles of a compound to determine the final answer.

As the student progresses through each level, the relationships between concepts become more complex, which increases the difficulty level.

Measuring Success

To determine the influence of the number of exposures on student achievement, the percent correct answers on exam questions was used. The University of Kansas Testing provides information discrimination values, and item difficulty in addition to item analysis. This research project will use the item analysis information. The report lists the item number, which is the number of the exam question. The letters A, B, C, D, and E are the possible responses for the question, and the asterisks denote the correct answer. The 'Response Percentages' identifies the percentage of correct answers for that question. In addition to this information, the report supplies the percentage of correct answers for the students that had an overall exam score in the upper one-fourth of the class. Similar information is provided for students that scored in the middle one-half of overall exam grades and for the students scoring in the lower one-fourth of scores.

This information will be used in two aspects of this study. First, the number of responses is an indication of the number of exam questions and will be used to describe the distribution of exam questions. Secondly, the percent correct answers on the exam question will be reported as a measure of student success and will be used to measure the success of the types and number of exposures.

To designate which set of scores are being evaluated or discussed, the following abbreviations will be used. The percent of correct answers provided by all the students in this study will be labeled PCA-All percentiles. The percent correct answers produced by the students that scored in the upper one-fourth students will be designated as PCA-Upper percentile. The percent correct answers generated by the middle one-half percentile will be named the PCA-Middle percentile, and the lower one-fourth student percentile will be PCA-Lower percentile.

Results

The data obtained from the procedures developed in the methods section were analyzed using inferential statistics. The data analysis and distribution will be discussed, and the results of the varied analysis will be reported.

First, exploratory data analysis was performed to determine if the data is normally distributed and meets the requirements to generate applicable results from the inferential statistics used in this research. Second, each classification method of question difficulty was examined to determine the validity of each process. Finally, the effects of the number and types of exposure on student success was analyzed. All statistical results are produced by the Statistical Package for the Social Sciences (SPSS) version 22 (IBM, 2011).

Variables

. Each statistical test used in inferential statistics has certain assumptions associated with it that need to be met before the analysis can be used to produce valid results (Leech, Barrett, & Morgan, 2015; Morgan, Leech, Gloeckner, & Barrett, 2013; Rovai, Baker, & Ponton, 2014). The independent variables in this research project are the number of exposures. The dependent variable is the exam success of students.

First, the independent variables and the dependent variable will be described, and then, the appropriateness of its use in inferential statistics will be discussed. Thirdly, each scale of question difficulty will be investigated to determine its ability to distinguish between the cognitive demands of each question level. Lastly, the effects of exposure variables on student exam success will be explored.

For a variable to be suitable for use in inferential statistics, the data should be normally distributed, have equal variance and independent observations. A Wald-Wolfowitz Runs test was

used to determine the randomness of observations, and the Kruskal-Wallis H test is a nonparametric test used to evaluate the null hypothesis that there is no difference between the sums of ranks of a nominal or ordinal variable.

Nominal variables. Nominal variables are composed of groups or categories. There is no order of categories, and each level of the variable is mutually exclusive from each other. The data is described by frequency distributions, the mode of the data, and the number of categories.

The nominal variables in this research are the number of exam questions for each semester and for each exam. In addition, the number of exam questions for each topic is a nominal variable.

Semester variable. The distribution of exam questions for each semester is given in Figure 13.

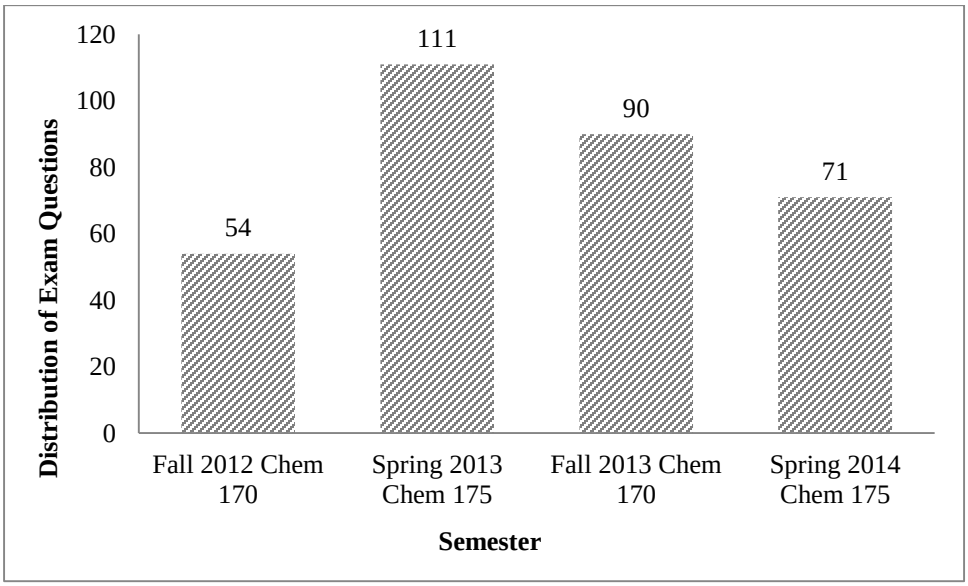


Figure 13. Distribution of the number of exam responses by semester.

The number of exam questions from the Fall 2012 Chem 170 semester is less than its companion course, which is the Spring 2013 Chem 175 semester. Data on the fourth exam from both the Fall 2012 semester and the Spring 2014 semester is missing. The researcher never received the statistical results for these two exams probably due to the increased activity at the end of a semester

The randomness of observations is when the observations of a variable are independent from all other observations of that variable (Price, 2000). The nonparametric Wald-Wolfowitz Runs test was used to determine if observations were random. The results of the runs test for the Fall 2012, Spring 2013, Fall 2013, and Spring 2014 semesters are in Table 2.

Table 2

Results of the Wald Wolfowitz Runs Test on the Semester Variable

	Semester			
	Fall 2012	Spring 2013	Fall 2013	Spring 2014
Test value[a]	80	93	92	70
Total cases	54	111	90	71
Number of runs	29	38	37	31
z	.45	.52	-.03	-.88
Sig.*	.65	.60	.98	.38

[a]Mode
*Asymp. Sig. (2-tailed)

Nonsignificant results for a Wald-Wolfowitz Runs test indicate randomness of observations. For each semester variable, the z score was not significant, and therefore, the observations are independent from each other for all semesters.

Both a parametric and nonparametric analyses were performed to determine if the distribution of percent correct answers are similar between semesters. The nonparametric Kruskal-Wallace H test results were not significant which indicates there is no difference between ranks of the semester grades, $\chi 2$ (3, N = 326) = 4.44, p = .22.

A one-way ANOVA is a parametric test to determine if there is a significant difference in means between levels of a nominal or ordinal variable. A significant Levene's test indicates the variance is not distributed equally, F(3, 322) = 3.12, p = .22; therefore, a Brown-Forsythe test was conducted instead of a one-way ANOVA. There was no statistically significant difference in the means between the semesters, F*(3, 277.58) = 2.29, p = .08.

The results support the randomness of observations and that there is no difference in the sums of ranks or means of the percent correct answers for each semester. Complete results of the statistical analyses of the semester variable are in Appendix B Section 1.

Exam variable. The distribution of exam questions is given in Figure 14.

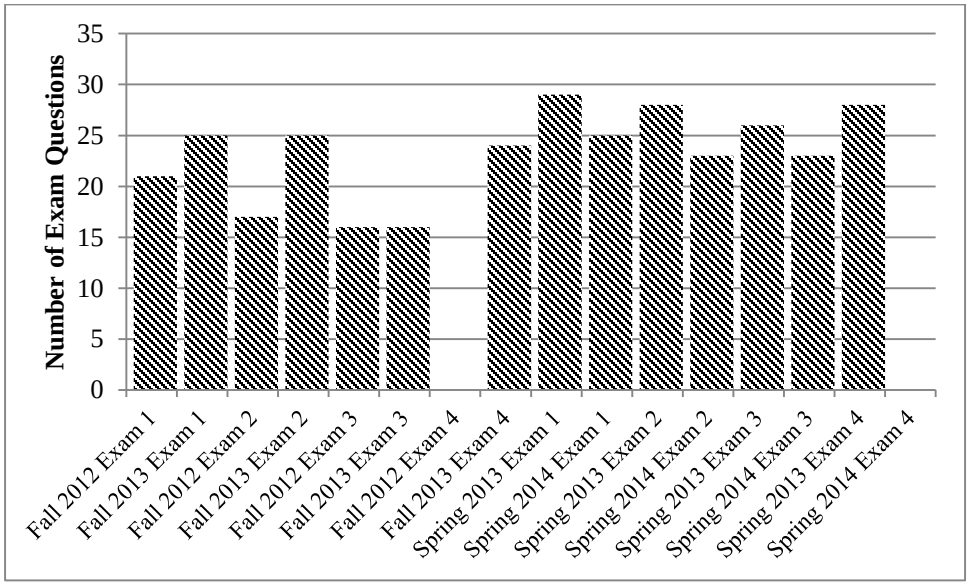


Figure 14. Number of exam questions for each exam in the study.

As was the case for the semester variable, data from the fourth exam is missing for both the Fall 2012 and Spring 2014 semesters. A Wald-Wolfowitz Runs test was performed to evaluate the randomness of the exam responses, and tests results provide evidence the observations are random and independent from each other.

A Kruskal-Wallis H test provided evidence that there is no difference in the sum of ranks of student success for each exam, $\chi^2(13, N = 312) = 21.49$, $p = .06$. In addition to the Kruskal-Wallis H test, a one-way ANOVA was performed on the PCA-All student group across all levels of the exam variable. The results were not significant, $F(3, 322) = 1.97$, $p = .12$, indicating there is no significant differences in grades between the levels of the exam variable. The statistical results for all analyses of the exam variable are in Appendix B Section 2.

There is no statistically significant difference in the percent correct answers between exams or between semesters. The observations are random for both the exam and semester variables. As a result, all response data was combined into one data set with $N = 326$.

Topics. The distribution of the number of exam questions for each topic is in Figure 15.

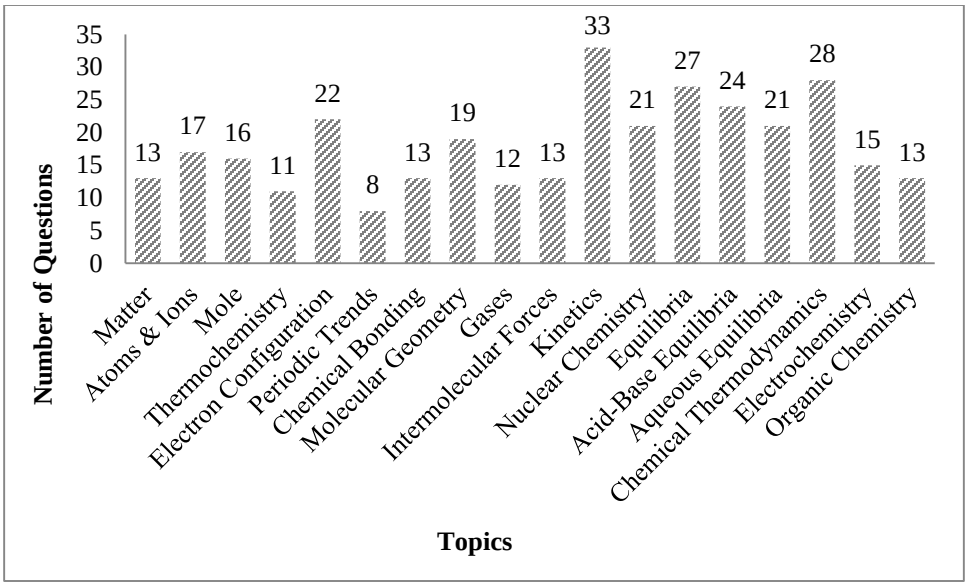


Figure 15. The distribution of exam questions per topic.

The topics studied in the Chem 170 courses are the first ten topics starting on the left in Figure 3, which are the subjects of matter through intermolecular forces. The topics in Chem 175 are the eight topics beginning with kinetics through organic chemistry.

The variable of topics does not have random observations, and there is a significant difference in means of each topic, $F(17, 325) = 3.20$, $p = .00$. The course instructor decides the number of exam questions for each topic and how much time is spent discussing each subject based on the instructor's experience and knowledge of chemistry. Therefore, there will not be randomness of observations. In addition, the means of the different topics vary because some topics are more complex and difficult.

Statistical results for the nominal variable of topics are in Appendix B Section 3.

Ordinal variables. Ordinal variables are similar to nominal variables in that each group or level is mutually exclusive from each other. Whereas nominal variables have random categories, ordinal variables have a sequence or an order to them. The ordinal variables in this

research are the modified Bloom's taxonomy and question type (recall, algorithmic, and conceptual). Both are procedures used to classify the difficulty of exam questions.

Modified Bloom's taxonomy. The taxonomy classifies the difficulty of both exam and homework problem by the type of knowledge required to find solutions for the questions.

A Wald-Wolfowitz Runs test was conducted on this ordinal variable (N = 326), and the results were not significant (runs = 118, p = .05) when using the median (1) as the cut point indicating there is randomness of observations.

The result of a one-sample Kruskal-Wallis H test was significant, $\chi^2(3, N = 326) = 37.94$, p = .00, indicating there is a difference between the PCA-All student group scores between the different levels of the taxonomy. Because Bloom's Taxonomy is a measure of question difficulty, there should be a difference in means between the different levels of the taxonomy. To determine if the rank means are significantly different within each level, a Kruskal-Wallis H test was used. The means of percent correct answers for each level of the taxonomy were compared by exam and by semester. Results are in Table 3.

Table 3

Kruskal-Wallis H Test for Rank Means within Each Level of Bloom's Taxonomy

Level	$\bar{X}$	χ^2	*df*	*N*	*p*	Grouping variable
Recall	82.84	1.23	3	91	.75	Semester
Comprehension	77.46	.84	3	134	.84	Semester
Application	67.83	7.36	3	70	.06	Semester
Analysis	63.97	4.38	3	31	.22	Semester

There is no significant difference in mean ranks within each level of the modified Bloom's taxonomy grouped by semester. Complete statistical analysis results for the modified Bloom's taxonomy are in Appendix B Section 4.

Recall, algorithmic and conceptual question type. Question type is another method to classify question difficulty. The method categorizes problems by the cognitive process needed to find a solution to a question.

The results of the Wald-Wolfowitz Runs test were significant (N = 326, runs =98, p = .00) using the median as the cut point (1) indicating the observations are not random. Results from the Kruskal-Wallis H test support the idea that there is a difference in sum of ranks for the levels of question type, $\chi^2(2, N = 326) = 26.56$, $p = .00$.

Question type is a measure of question difficulty so there should be a difference between the rank means of each level. To determine if the rank means were significantly different within a level, a Kruskal-Wallis H was used to test between semesters. Results are in Table 4.

Table 4

Kruskal-Wallis H Test for Rank Means within Each Level of Question Type

Level	$\bar{X}$	χ^2	*df*	*N*	*p*	Grouping variable
Recall	81.75	6.75	3	81	.08	Semester
Algorithmic	78.61	1.13	3	118	.77	Semester
Conceptual	68.90	3.17	3	127	.37	Semester

There is no significant difference in mean ranks within each level of question type whether grouped by exam or by semester. Statistical results on behalf of this variable are found in Appendix B Section 5.

Dichotomous variables. A dichotomous variable is a special case of a nominal variable where there are only two groups. The groups can be either ordered or random. Cognitive skill level is a dichotomous variable in this research project.

Cognitive skill level. The dichotomous variable of cognitive skill level organizes question difficulty by the type of critical thinking skills required to answer a question. The levels are divided into questions requiring lower-order cognitive skills (LOCS) and those requiring higher-order cognitive skills (HOCS).

Results from the Wald-Wolfowitz Runs test does support randomness of observation using the mean (.32) as the cut point (N= 326, runs = 133, p = .22). In addition, the results of the Kruskal-Wallis H test show there is a significant difference in the means of ranks, $\chi^2(1, N = 326) = 25.83$, $p = .00$.

A Kruskal-Wallis H test was conducted on both the LOCS and HOCS levels grouping by exams and by semesters. For the LOCS level, the results were insignificant and indicate there is no difference in mean ranks for exams, $\chi^2(3, N = 221) = 1.08$, $p = .78$, and for semesters, $\chi^2(3, N = 221) = 2.54$, $p = .47$. Likewise, the results were insignificant for the HOCS level when grouped by exams $\chi^2(3, N = 104) = 1.02$, $p = .80$ and by semesters, $\chi^2(3, N = 104) = 6.29$, $p = .10$.

The descriptive statistics, frequency distribution, and statistical tests for the dichotomous variable of cognitive skill level are in Appendix B Section 6.

Scale variables. Scale variables are described as having a true zero and continuous values. The scale variables in this research project are the number of exposures, the number of

steps (a measure of question difficulty), and percent correct answers for all student percentiles, which is the dependent variable.

Number of steps. This measure of question difficulty uses the number of steps required to answer a question as the means to decide the cognitive demand of questions. As the number of steps increases, the complexity level increases and the success rate of the students decreases (Hartman & Lin, 2011).

The Wald-Wolfowitz Runs Test was significant, (N = 326, runs = 125, p = .00), using the mean (2) as the cut point indicating there is no randomness in the observations. The results of a one sample Kolmogorov-Smirnov test were significant, $D(326) = 4.98$; $p = .00$, indicating there is not a normal distribution of the number of steps.

The number of steps is a measure question difficulty, and there should be no expectation of a normal distribution or randomness of observations. Statistical results for all analyses of this variable are in Appendix B Section 7.

Percent correct answers. The percent of correct answers is a measure of student achievement and is the percentage of students that answer an exam question correctly. This variable should have a normal distribution, and in the following section, the methods used to test scale data for normal distribution and equal variance will be described and the results given.

First, a Wald-Wolfowitz Runs test provided evidence that the sequence of values for the PCA-All student group variable is recorded in random order based on using a median of 80 as the cut point (N=326, runs = 152, p = .18).

Second, the assumption of normal distribution was tested. The distribution of scale variables is described by the skew value, the mode, the mean, and the median. A normally

distributed scale variable should have a skew value between -1 to +1 (Leech et al., 2015) and the mean, the mode, and the median should be equal.

The skew value for the PCA-All student group is -1.14 (std. error = .14) and is considered outside of the accepted range for a normal distribution (Rovai et al., 2014). The mean is 76 percent correct answers, the mode is 93 percent of correct answers, and the median 80 percent correct answers. These are all indications the dependent variable is not normally distributed.

A quantitative test used determine if a distribution is normal is the parametric Kolmogorov-Smirnov D test with the Lilliefors corrections. In this research project, it tests the null hypothesis that there is no difference in the distribution of PCA-All student scores (M = 75.61%, SD = 18.64, N = 326) and a normal distribution. Test results were significant, $D(326) = 1.92$, $p = .00$, indicating there is a difference between the two distributions, and gives evidence that the null hypothesis should be rejected (Rovai, Baker, & Ponton, 2013).

The results of the tests provide evidence the distribution of the percent correct answers scale variable does not have a normal distribution. The results of each statistical test are in Appendix B Section 8.

Transformation of the scale variable. Transformation of skewed data produces values comparable to a normal distribution. Recommended transformation methods are based on the severity of the skewness (Leech et al., 2015; Rovai et al., 2013). The possible transformation methods are presented in Appendix B Section 9.

When the results of the different transformations are compared, the moderate skew method produces the most normal distribution. The moderate-skew transformation, which is a square root conversion, produces a skew value of .34 and a standard error of .14. The mode is now 2.83%, the mean value is 4.70%, and the median is 4.58%.

Linearity of the dependent and independent variables. An important assumption for inferential statistics is that of a linear relationship between variables. To determine if this type of association exists, the transformed PCA-All student and the original PCA-All student data sets were regressed against the independent variables of lecture, homework, laboratory, and textbook exposures.

The strength of the relationship between the independent and dependent variables is determined by the coefficient of determination. It indicates how much variability in the dependent variable is explained by the independent variable. If the *R*-squared (R^2) value is between .7 and 1.0, there is a strong relationship between the two variables; an R^2 value between .4 and .7 indicates a moderate relationship between the variables; a value less than .4 indicates a very weak association (Pecher & Boot, 2011).

The following section gives the regression equation and the R^2 value for each of the exposure variables.

Lecture. In Figure 16, both the PCA-All student group and the transformed PCA-All student scores are regressed on the number of lecture exposures. There is a regression line, regression equation, and a R^2 value for each data set.

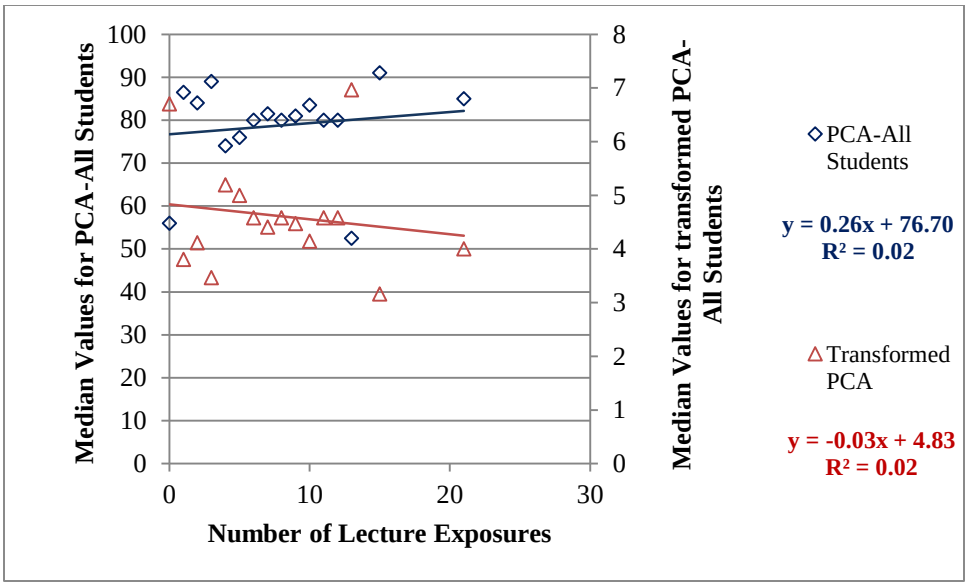


Figure 16. Regression lines for the both the original and transformed data sets regressed on the number of lecture exposures. The slope of the transformed data is negative due to the square root transformation.

The coefficient of determination (R^2 = .02) is very small and indicates a very weak relationship for both the transformed and original student data.

Homework. The regression lines for both the original and transformed student data sets regressed on the number of homework exposures are plotted in Figure 17.

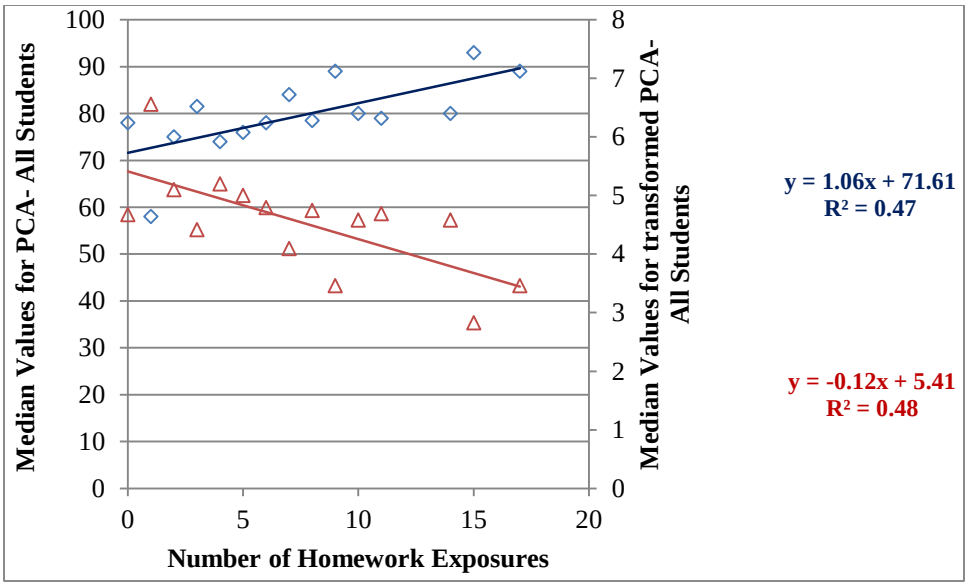


Figure 17. Regression lines for both the PCA-All student and the transformed student data sets regressed on the number of homework exposures.

The R^2 value for the PCA-All student group is .47, and the transformed data has a value of .48. A linear relationship exists between the number of homework exposures and the dependent variable of percent correct answers for all students. Likewise, there is a linear relationship between homework exposures and the transformed data. The R^2 value represents a moderate relationship.

Textbook. The PCA-All student and the transformed PCA-All student data sets were regressed on the number of textbook exposures. The results are in Figure 18.

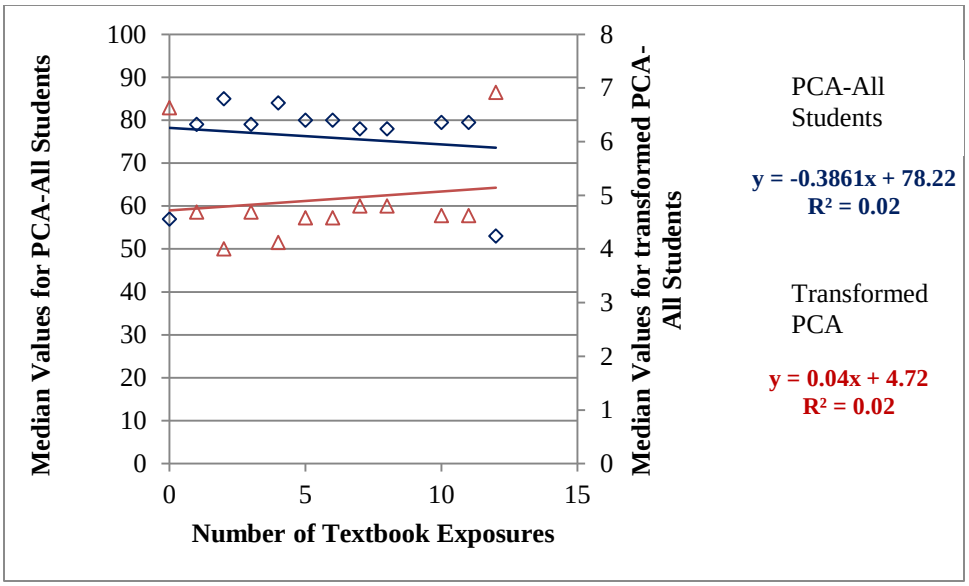


Figure 18. Regression lines of both the PCA-All student and the transformed student data regressed on textbook exposures. The slope for the original data is negative.

The coefficient of determination for both regression lines is very small (R^2 = .02), indicating a very weak relationship between textbook exposures and the percent correct answers.

Laboratory. The transformed and original PCA-All student data sets were regressed on the independent variable of laboratory exposures. The regression lines, regression equations, and R^2 values are in Figure 19.

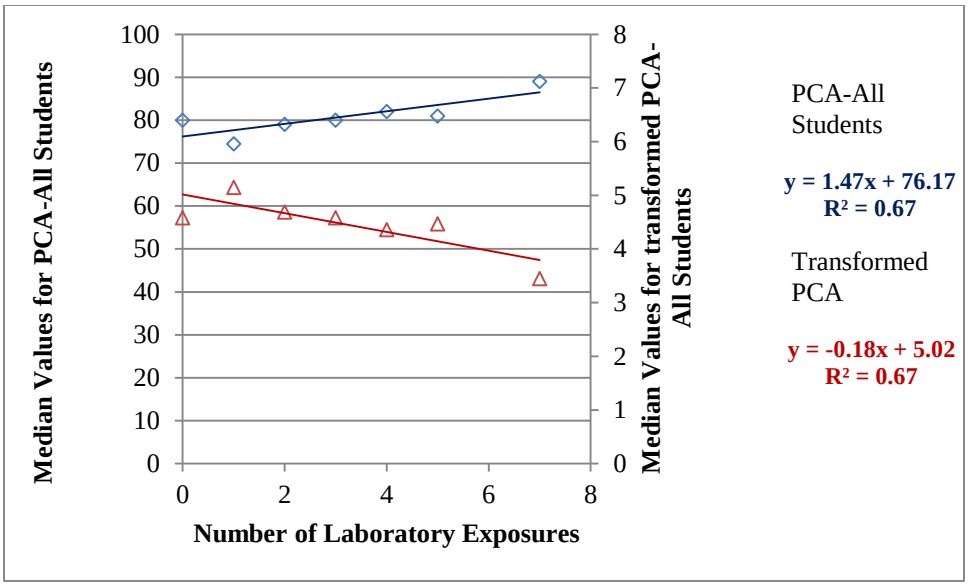


Figure 19. Regression lines for the original and transformed data regressed on the number of laboratory exposures.

For both the PCA-All student group and for the transformed data, there is a linear relationship between the dependent variable and the independent variable of laboratory exposures. The coefficient of determination (R^2 = .67) is high, implying a strong relationship between the independent and dependent variables.

Interpretation of the transformed data. The ability to understand the transformed data is not intuitive. The two regression equations from Figure 19 will be used to draw attention to the difficulty of interpreting results of the transformed data. Both regression equations describe the relationship of the percent correct answers for the PCA-All student group regressed on lab exposures.

Original Data:	$\hat{y}=1.47x+76.17$	$R^2 = 0.67$
Transformed Data:	$\hat{y}= -0.18x+5.02$	$R^2 = 0.67$

The coefficient for the variable produced from the original data is positive. This indicates that an increase in lab exposures will increase the success of students. However, the regression equation constructed from the transformed data shows a negative relationship between lab exposures and student achievement. In other words, negative coefficients represent positive relationships between variables when using the transformed date. In addition, the y-intercept for the transformed data has no real meaning when discussing student grades.

However, the transformed and original data produce the same R^2 value for each of the exposure variables. The R^2 value represents the strength of the relationship between the independent and dependent variables. Since the R^2 values of both data sets are the same and this research is interested in the strength of relationships between variables, the results will be reported for the original data to make interpretation of the results more understandable.

Question Difficulty

Each method of measuring question difficulty will be examined to determine if the classification methods are valid. The statistical tests used in this section are correlations, ANOVAs, and *t*-tests.

Reliability of coding methods. To test the reliability of the coding methods used to categorize the different levels of question difficulty, an intraclass correlation coefficient (ICC) analysis was performed. The test measures how consistent an individual rater is over time (Rovai et al., 2014). The ICC was used to analyze the ratings from the Fall 2013 Chem 170 course (exams one and two) and from the Spring 2013 Chem 175 course (exams three and four). The results of the analyses are in Table 5.

Table 5

Intra-rater Reliability (ICC) for Question Difficulty Ratings

Difficulty measure	ICC[a]	Degree of reliability[b]	95% Confidence interval Lower-bound	Higher bound
Lin-Hartman	.16	poor	-.02	.33
Question type	.55	moderate	.41	.66
Bloom's taxonomy	.58	moderate	.45	.69
Cognitive skill level	.41	fair	.25	.55

[a]Single measures
[b] (StatsToDo)

The interrater reliability between the process used in this project and the Lin paper (Hartman & Lin, 2011) is poor. The three rating scales adapted from methods found in the chemical education literature have a fair to moderate degree of reliability. All results are in Appendix B Section 10.

Number of steps. The relationships between the percent correct answers for all student percentiles and the number of steps required to complete an exam question were analyzed using a Pearson correlation and linear regression. Table 6 has the results of the Pearson Correlation analysis for each student percentile.

Table 6

Pearson Correlation between All Student Percentiles and the Number of Steps

Student percentile	Number of steps	Significance	Coefficient of Determination (r^2)
PCA-All Students	$r(326) = -.18$**	.00	.03
PCA-Upper[a]	$r(326) = -.09$	.09	.01
PCA-Middle	$r(326) = -.18$**	.00	.03
PCA-Lower	$r(326) = -.17$*	.00	.03

[a]Results for the PCA-Upper student percentile are not significant.
* $p < .05$ (two-tailed).
** $p < .01$ (two-tailed).

There is a statistically significant, negative relationship between the number of steps and student success except for the PCA-Upper student percentile. The relationships are weak (Rovai et al., 2014) and explain 3% of the variance in the percent correct answers. The negative correlation indicates that as the number of steps increases, the success rate on exam questions decreases.

The results of a regression analysis of the percent correct answers for all student percentiles regressed on the number steps are in Table 7.

Table 7

Number of Steps Regressed on the All Student Group

Student percentile	Regression model	*p*	R^2
PCA-All Students	$F(1, 324) = 10.44$	.00	.03
PCA-Upper 1/4th percentile	$F(1, 324) = 2.82$	.09	.01
PCA-Middle ½ percentile	$F(1, 324) = 10.97$	.00	.03
PCA-Lower 1/4th percentile	$F(1, 324) = 9.43$	.00	.02

The number of steps predicted a significant amount of the variance in student success for all percentiles except the PCA-Upper 1/4th percentile. The number of steps is a significant predictor of student success for the percent correct answers for all students, $\beta = -.18$, $t(325) = -3.23$, $p = .00$, for the middle ½ percentile, $\beta = -.18$, $t(325) = -3.31$, $p = .00$, and for the lower 1/4th percentile, $\beta = -.17$, $t(325) = -3.07$, $p = .00$.

Cognitive skill level. The relationship between the variable of cognitive skill level (LOCS and HOCS) and percent correct answers for all students was tested using a point-biserial correlation. The results are in Table 8.

Table 8

Point-Biserial Correlation between All Student Percentiles and Cognitive Skill Level

Student percentile	Cognitive skill level	Significance	Coefficient of Determination (r^2)
PCA-All students	r_{pb} (326) = -.31*	.00	.10
PCA-Upper percentile	r_{pb} (326) = -.27*	.00	.07
PCA-Middle percentile	r_{pb} (326) = -.31*	.00	.10
PCA-Lower percentile	r_{pb} (326) = -.26*	.00	.07

* $p < .01$ (two-tailed).

A statistically significant correlation exists between the cognitive skill level required to answer an exam question and the success of students. Each student percentile has a negative correlation indicating that the change from the lower-cognitive skill level questions to the higher-order ones is difficult for all students.

The cognitive skill level explains 10% of the variance in the dependent variables of PCA-All student group and PCA-Middle student percentile. For the PCA-Upper and the PCA-Lower percentiles, the cognitive skill level accounts for 7% of the variance in student success.

In addition to the Pearson correlation analysis, an independent samples *t* test was performed on each student percentile. The results are in Table 9.

Table 9

Independent Sample t Test on Cognitive Skill Level and All Student Percentiles

Variable	*M*	*SD*	*t*	*df*	*p*
PCA-All[a]			5.22	151.51	.00
LOCS	79.59	15.33			
HOCS	67.11	22.01			
PCA-Upper[a]			4.04	128.11	.00
LOCS	93.57	10.17			
HOCS	85.12	20.19			
PCA-Middle[a]			5.13	150.09	.00
LOCS	82.30	16.50			
HOCS	68.95	24.03			
PCA-Lower			4.83	324	.00
LOCS	60.69	23.04			
HOCS	47.08	25.03			

[a]The *t* statistic and degrees of freedom were adjusted due to unequal variances.

In each case, there is a statistically significant difference between the different levels of cognitive skill level and the success of each student percentile. For the PCA-All student group, the effect size (d = .85) exceeds Cohen's definition of a large effect size of d = .80 (Rovai et al., 2014), and the 95% confidence level is CI [7.76, 17.21].

The PCA-Upper student percentile has an effect size (d = .71) that is between the values of a medium to large effect with a 95% confidence interval of CI [4.31, 12.60]. For the PCA-Middle percentile, the effect size (d = .84) exceeds the value for a large effect, and the 95% confidence level is CI[8.21, 18.50]. Finally, the PCA-Lower student percentile has an effects size

(d = .54) considered to have a medium effect according to the conventions of Cohen's scale (d = .50). The 95% confidence level is CI[8.07, 19.15].

Modified Bloom's taxonomy. Bloom's taxonomy classifies exams and homework questions by the kind of knowledge needed to solve problems. The least cognitively demanding level of the taxonomy is recall, and the difficulty of the questions increases through each level to analysis, which is the highest level of the taxonomy used in this research. A one-way ANOVA or a Brown-Forsythe test analyzed each student percentile across every level of Bloom's taxonomy to determine if there is a significant difference in student success between the levels. For the post hoc analysis, either the Games-Howell or the Tukey HSD were used.

Levene's test for the homogeneity of variance provides evidence that the variance in the percent correct answer scores of the PCA-Lower percentile is statistically equivalent, $F(3, 322) = 1.82$, $p = .14$, and the post analysis used was the Tukey HSD.

The Levene's test for the PAC-Upper percentile, $F(3,322) = 13.2$, $p = .00$, the PAC-Middle percentile, $F(3, 322) = 12.47$, $p = .00$, and the PCA-All student group, $F(3, 322) = 8.88$, $p = .00$, were significant, and the variance in these percentiles is not statistically equivalent. To correct for these violations of homogeneity, a Brown-Forsythe test was performed on these percentiles, and the post hoc analysis was the Games-Howell test. The results of all student percentiles are in Table 10.

Table 10

Modified Bloom's Taxonomy and Student Success

Student percentile	*F* Value[a]	Significance	Effect size (η^2)
PCA-All students	$F^*(3, 138.06) = 11.93$	.00	.12
PCA-Upper	$F^*(3, 116.64) = 6.05$	.00	.07
PCA-Middle	$F^*(3, 126.80) = 11.33$	.00	.12
PCA-Lower	$F(3, 322) = 11.67$	.00	.10

[a]*F* values are represented by *F** for the Brown-Forsythe test.

A statistically significant different relationship exists between the levels of the modified Bloom's taxonomy and each student percentile. The effect sizes (η^2) are large for the PCA-All students group and the PCA-Medium percentile. The PCA-Upper percentile has a medium effect, and the PCA-Lower student group is classified as having a medium to large effect. (Rovai et al., 2014).

A Games-Howell post hoc analysis was performed on the PCA-All student group to determine the source of the difference between the levels of the taxonomy. Students score significantly higher on recall questions ($M = 82.84$, $SD = 15.93$) compared to their scores on application ($M = 67.83$, $SD = 22.23$), $p = .00$ and analysis questions ($M = 63.97$, $SD = 22.36$), $p = .00$. This is also true for the comprehension level ($M = 77.76$, $SD = 14.33$) when compared to the application ($M = 67.83$, $SD = 22.23$), $p = .01$ and analysis levels ($M = 63.97$, $SD = 22.36$), $p = .01$.

A statistically significant difference does not exist between the levels of recall ($M = 82.84$, $SD = 15.93$) and comprehension ($M = 77.76$, $SD = 14.33$), $p = .05$ and between the levels of application ($M = 67.83$, $SD = 22.23$) and analysis ($M = 63.97$, $SD = 22.36$), $p = .85$.

For the PCA-Upper student percentile, a Games-Howell post hoc analysis was used to determine which levels of the taxonomy are significantly different from each other. As was the case for the PCA-All student group, the PCA-Upper student percentile score was significantly higher on recall questions (M = 93.99, SD = 10.03) compared to their scores on application (M = 85.77, SD = 21.55), p = .02 and analysis questions (M = 83.50, SD = 19.59), p = .03. Students score higher on the comprehension questions (M = 93.13, SD = 9.58) than on the application, (M = 85.77, SD = 21.55), p = .04.

A statistically significant difference in means does not exist between the levels of recall (M = 93.99, SD = 10.03) and comprehension questions (M = 93.13, SD = 9.58), p = .92, between comprehension (M = 93.13, SD = 9.58) and analysis questions (M = 83.50, SD = 19.59), p = .05, and between the levels of application (M = 85.77, SD = 21.55) and analysis (M = 83.50, SD = 19.59), p = .95.

The results provided by the PCA-Middle student percentile mimic the results for the all student group and upper student percentile. A Games-Howell post hoc analysis was used to determine the significant differences between the different levels. Students in the PCA-Middle percentile score significantly higher on recall questions (M = 85.64, SD = 16.46) compared to their scores on application (M = 69.57, SD = 24.02), p = .00 and analysis questions (M = 65.65, SD = 25.47), p = .00. This applies to the comprehension level as well with student scoring higher on the comprehension questions (M = 80.17, SD = 15.75) than on the application (M = 69.57, SD = 24.02), p = .01 and analysis questions (M = 65.65, SD = 25.47) p = .02.

On the other hand, a statistically significant relationship does not exist between the levels of recall (M = 85.64, SD = 16.46) and comprehension questions (M = 80.17, SD = 15.75), p =

.06, and between the levels of application (M = 69.57, SD = 24.02) and analysis (M = 65.65, SD = 25.47), p = .89.

For the PCA-Lower student percentile, a Tukey HSD test was used as the post hoc analysis. Students in the PCA-Lower percentile score significantly higher on recall questions (M = 66.49, SD = 24.34) compared to their scores on comprehension (M = 56.77, SD = 21.33), p = .01, application (M = 47.64, SD = 24.79), p = .00, and analysis questions (M = 44.42, SD = 25.75), p = .00. These students score higher on comprehension questions (M = 56.77, SD = 21.33) than on application (M = 47.64, SD = 24.79), p = .04 and analysis questions (M = 44.42, SD = 25.75), p = .04. A statistically significant difference in means does not exist between the levels of application (M = 47.64, SD = 24.79) and analysis (M = 44.42, SD = 25.75), p = .92.

There is a significant difference in student scores between the levels of the Bloom's taxonomy. Linear regression was used to determine the effect each level of the taxonomy had on the different student percentiles. The results of the regression are in Table 11.

Table 11

Student Success Regressed on Levels of Bloom's Taxonomy

Levels of Bloom's taxonomy	B	*SE*(B)	β	*t*	*p*
All student percentile					
Recall level	10.02	2.24	.24	4.48	.00
Comprehension	-	-	-	-	-
Application	-9.91	2.46	-.22	-4.03	.00
Analysis	-12.87	3.45	-.20	-3.73	.00
Upper 1/4th percentile					
Recall	4.32	1.80	.13	2.40	.02
Comprehension	3.84	1.64	.13	2.34	.02
Application	-6.50	1.95	-.18	-3.34	.00
Analysis	-8.15	2.74	-.16	-2.98	.00
Middle ½ percentile					
Recall	10.55	2.42	.24	4.35	.00
Comprehension	-	-	-	-	-
Application	-10.78	2.66	-.22	-4.06	.00
Analysis	-13.70	3.74	-.20	-3.66	.00
Lower 1/4th percentile					
Recall	14.08	2.93	.26	4.81	.00
Comprehension	-	-	-	-	-
Application	-11.10	3.25	-.19	-3.41	.00
Analysis	-13.18	4.57	-.16	-2.88	.00

For each student percentile the levels of application and analysis predicts a statistically significant decrease in student scores. Recall levels predict a statistically significant increase in student exam scores for all student percentiles. For the PCA-Upper percentile, comprehension questions predicted a significant amount of exam success. This is the only percentile where this occurred.

Recall, algorithmic, and conceptual question types. This method to classify the difficulty of exam questions is based on the cognitive process required to answer a question. The levels are recall, questions requiring memorized facts; algorithmic, using mathematical processes; and conceptual, understanding the underlying concepts. A one-way ANOVA or a Brown-Forsythe test was used to determine if a significant difference exists in the scores for each student percentile across the different levels of question difficulty. Post hoc tests were either the Games-Howell test or the Tukey HSD.

The F values and significance levels for each student group are in Table 12.

Table 12

Question Type and Student Success

Student percentile	*F* Value	Significance	Effect size (η^2)
PCA-All Students	$F^*(2, 295.68) = 15.40$	.00	.09
PCA-Upper	$F^*(2, 265.39) = 11.26$	.00	.06
PCA-Middle	$F^*(2, 295.79) = 18.93$	.00	.10
PCA-Lower	$F(2, 323) = 13.05$	.00	.07

Note. F^* designates the *F* statistic from a Brown-Forsythe analysis.

There is a statistically significant difference in the relationships between the levels of question type and each student percentile. The effect sizes are for the PCA-Upper and PCA-lower student percentiles are classed as medium, and the PCA-All and the PCA-Middle student percentiles have effect sizes ranked as medium to large.(Rovai et al., 2014)

To investigate the significant difference in scores between the levels of question difficulty, a Games-Howell post hoc test was used to analyze the Brown-Forsythe results for the PCA-All student group. Students scored higher on recall questions (M = 81.75, SD = 16.09) than on conceptual questions (M = 68.91, SD = 21.47), p = .00, and they scored significantly higher on algorithmic question (M = 78.61, SD = 14.40) than on conceptual questions (M = 68.91, SD = 21.47), p = .00. There is not a significant difference in the scores between recall (M = 81.75, SD = 16.09) and algorithmic scores (M = 78.61, SD = 14.40), p = .34.

A Games-Howell post hoc analysis was used to analyze the Brown-Forsythe results for the PCA-Upper percentile. Students scored higher on recall questions (M = 92.83, SD = 11.68) than on conceptual questions (M = 86.43, SD = 18.80), p = .01, and they scored higher on algorithmic question (M = 94.32, SD = 9.36) than on conceptual questions, p = .00. There is not a significant difference in the scores between recall (M = 92.83, SD = 11.68) and algorithmic scores (M = 94.32, SD = 9.36), p = .60.

For the PCA-Middle student percentile, there is a significant difference in the means of recall and conceptual questions with students scoring higher on recall questions (M = 84.41, SD = 16.56) than on conceptual questions (M = 70.17, SD = 23.51), p = .00. These students also scored higher on algorithmic questions (M = 82.14, SD = 15.35) than on conceptual questions, p = .00. There is not a significant difference in the scores recall M = 84.41, SD = 16.56) and algorithmic questions (M = 82.14, SD = 15.35), p = .59.

A Tukey HSD analysis was used to examine the results of a one-way ANOVA of the PCA-Lower student percentile. There is a significant difference in the means of recall and algorithmic questions with students scoring higher on recall questions (M = 66.11, SD = 24.18) than on algorithmic questions (M = 57.47, SD = 22.44), p = .03. Students also scored higher on recall questions (M = 66.11, SD = 24.18) than on conceptual ones (M = 49.08, SD = 24.36), p = .00. In addition, students in this group scored higher on algorithmic questions than on conceptual questions (M = 49.08, SD = 24.36), p = .02.

For each student percentile, the percent correct answers was regressed on each level of question type. Results are in Table 13.

Table 13

Student Success Regressed on Levels of Question Type

Question Type Levels	B	*SE*(B)	β	*t*	*p*
All student percentile					
Recall	8.17	2.35	.19	3.48	.00
Algorithmic	4.70	2.14	.12	2.20	.03
Conceptual	-10.98	2.03	-.29	-5.41	.00
Upper 1/4th percentile					
Recall	-	-	.-	-	-
Algorithmic	5.40	1.67	.18	3.24	.00
Conceptual	-7.28	1.62	-.24	-4.50	.00
Middle ½ percentile					
Recall	8.48	2.55	.18	3.33	.00
Algorithmic	6.42	2.30	.15	2.79	.01
Conceptual	-12.89	2.18	-.31	-5.91	.00
Lower 1/4th percentile					
Recall	12.99	3.06	.23	4.24	.00
Algorithmic	-	-	-	-	-
Conceptual	-11.91	2.71	-.24	-4.40	.00

For each student percentile, conceptual questions predict a significant decrease in student scores. Recall and algorithmic level questions predict an increase in student exam scores for all student percentiles except for the upper one-fourth and the lower one-fourth percentiles.

For the upper percentile, recall questions did not have a statistically significant relationship with student success. For the lower one-fourth percentile, algorithmic level questions did not predict a significant amount of variance in the exam success.

Topics. Another measure of question difficulty is the topic being studied. A one-way ANOVA or Brown-Forsythe test was used to determine if there is a significant difference in student scores across all levels of topics. A Games-Howell or Tukey HSD was used for post hoc analyses.

Test results show there is a statistically significant difference in percent correct answers for the PCA-All student group, $F(17, 308) = 3.20$, $p = .00$, $\eta^2 = .15$, for the PCA-Upper percentile, $F^*(17, 308) = 2.63$, $p = .00$, $\eta^2 = .13$, for the PCA-Middle percentile, $F(17, 308) = 3.37$, $p = .00$, $\eta^2 = .16$, and for the PCA-Lower percentile, $F(17, 308) = 2.80$, $p = .00$, $\eta^2 = .13$. In each case, the effect sizes are large.

A Tukey HSD post hoc analysis was performed on the results of the PCA-All student group one-way ANOVA, and the results show a significant difference between several topics. These students scored higher on the topic of matter ($M = 82.92$, $SD = 13.09$) compared to the topic of aqueous equilibria ($M = 60.71$, $SD = 20.60$), $p = .04$. This student group scored higher on electron configuration topics ($M = 83.83$, $SD = 10.26$) than questions on the topic of aqueous equilibria ($M = 60.71$, $SD = 20.60$), $p =.00$ and the topic of acid-base equilibria ($M = 62.46$, $SD = 23.11$), $p = .01$. In addition, scores were higher on questions over the topic of nuclear chemistry ($M = 82.57$, $SD = 14.91$) than those covering the topic of acid-base equilibria ($M = 62.46$, $SD = 23.11$), $p = .02$ and those covering aqueous equilibria ($M = 60.71$, $SD = 20.60$), $p = .01$. Last, students in this group scored higher on equilibria questions ($M = 79.30$, $SD = 16.31$) than those on the topic of aqueous equilibria ($M = 60.71$, $SD = 20.60$), $p = .04$.

For the PCA-Upper student percentile, a Games-Howell post hoc analysis was performed, and the results show students score significantly higher on topics of thermochemistry ($M = 97.09$, $SD = 3.36$) than on aqueous equilibria ($M = 81.76$, $SD = 16.29$), $p = .03$. This is true

for topic of electron configuration (*M* = 96.86, *SD* = 4.23) and aqueous equilibria (*M* = 81.76, *SD* = 16.29), with students scoring higher on the electron configuration questions, *p* = .03.

A post hoc Tukey HSD was used to analyze the differences between means of each topic for the PCA-Middle student percentile. The topics students scored higher on are presented in Table 14.

Table 14 :

Topics Student Scored Higher on Compared to Aqueous Equilibria

Topic	*p*
Matter	.02
Atoms	.01
Moles	.00
Electron configuration	.00
Nuclear chemistry	.04
Equilibria	.03
Energy	.04

PCA-Middle student percentile scored higher on these topics compared to the topic of aqueous equilibria. Matter, atoms, moles and electron configurations are topics presented in the first semester of the Chem 170/Chem 175 sequence.

In addition to the above topics, the middle percentile of students scored significantly higher on the following topics when compared to the topic of acid-base equilibria. The topics are listed in Table 15.

Table 15

Topics Students Scored Higher on Compared to Acid-Base Equilibria

Topic	*p*
Moles	.03
Electron configuration	.01
Nuclear chemistry	.04

The PCA-Middle student percentile has a significant difference in means between nine topics with students scoring lower on the topics on equilibria.

The results from the Tukey HSD post hoc analysis for the PCA-Lower percentile show a significant difference exists between nuclear chemistry (M = 68.81, SD = 22.05) and acid-base equilibria (M = 40.96, SD = 23.88), p = .01 and between nuclear chemistry and aqueous equilibria (M = 40.38, SD = 23.05), p = .01. As was the case for the PCA-Middle percentile, these students score lower on the equilibria questions.

Exposures

Exposures are the independent variables in this project. An exposure is information presented to students in the form of lecture, homework assignment, textbook example, or a laboratory experiment. The effects of these exposure variables were investigated by using linear regression.

Each student percentile was regressed on every one of the exposure variables. In the second analysis, all exposure variables were entered simultaneously and a backward regression was used to investigate whether a combination of exposures predicted a statistically significant amount of variance in student exam success. Finally, each student percentile was regressed on

the main exposures variables of lecture and homework that were gathered into levels of recall, algorithmic, and conceptual groups.

Single exposures on student success. The exam success rate of each student percentile and the all student group were individually regressed on the independent variables of lecture, homework, textbook, and laboratory exposures. There is not a statistically significant relationship between student achievement and the number of exposures.

Number of exposures. Each student percentile was regressed on the exposure variables of lecture, homework problems, textbook examples, and laboratory experiments that were entered simultaneously in the regression model to determine if the number of exposures predicted a significant amount of variance in student exam success. The results were not significant.

Location of exposures. The exposure variables were entered simultaneously and then a backward elimination was used to investigate the amount of variation predicted by the location of an exposure in student exam achievement. For the PCA-Upper student percentile, homework exposures predicted 1% in the variation of student success, $F(1,308) = 4.78$, $R^2 = .01$, $p = .03$. The effect size is small.

Categorized exposure variables. The independent variables of lecture and homework were categorized into recall, algorithmic, and conceptual types of exposures. These variables along with the textbook and lab exposures were entered all together and a backward elimination was conducted. The order of entry was recall lecture, algorithmic lecture, conceptual lecture, recall homework, algorithmic homework, conceptual homework, textbook, and lab.

The choice of this method was to determine if any combination of the categorized variables predicted a significant amount of variation in student exam performance. Results for the all student group are in Table 16.

Table 16

All Students Group Regressed on Categorized Exposures

Variable	B	*SE*(B)	β	*t*	*p*
Model					
Constant	72.24	2.47		29.30	.00
Recall homework problems	2.20	.69	.18	3.17	.00
Algorithmic homework problems	1.54	.52	.21	2.98	.00

Note. VIF < 2.00.

Both recall and algorithmic homework problems explained a significant amount of the variance in the exam grades of this percentile. The two criterion variable explained 4% of the variance in exam scores, $F(3, 322) = 5.37$, $p = .00$.

For the PCA-Upper student percentile, the same process was used. The independent variables of recall lecture, algorithmic lecture, conceptual lecture, recall homework, algorithmic homework, conceptual homework, textbook, and lab exposures were entered simultaneously and then a backward elimination was performed. Results are in Table 17.

Table 17

PCA-Upper Student Percentile Regressed on Categorized Exposure Variables

Variable	B	SE(B)	β	*t*	*p*
Model					
Constant	86.47	1.62		53.55	.00
Recall homework problems	1.49	.55	.16	2.73	.01
Algorithmic homework problems	.87	.34	.15	2.60	.01

Note. VIF < 2.0.

Recall homework and algorithmic homework problems significantly predicted student performance on exam questions for this percentile. The regression model explained 3% of the variance in exam scores, $F(2, 323) = 5.36$, $p = .00$.

The categorical exposure variables were entered simultaneously and a backward regression was executed on the dependent variable of PCA-Middle student percentile. The results of the regression are in Table 18.

Table 18

PCA-Middle Student Percentile Regressed on Exposure Type

Variable	B	*SE*(B)	β	*t*	*p*
Model					
Constant	71.66	2.21		32.48	.00
Define homework problems	2.76	.75	.21	3.69	.00
Algorithmic homework problems	1.08	.46	.14	2.36	.02

Note. VIF < 2.00

For the PCA-middle student percentile, recall homework and algorithmic homework problems significantly predicted student performance on exam questions. The regression model explained 4% of the variance in student success, $F(2,323) = 7.58$, $p = .00$.

As for the other student percentiles, the categorical exposure variables were entered simultaneously, and a backward elimination was conducted to determine their effect on the success of the PCA-lower student percentile. The results of the regression are in Table 19.

Table 19

PCA-Lower Student Percentile Regressed on Categorized Exposure Variables

Variable	B	*SE*(B)	β	*t*	*p*
Model					
Constant	53.71	3.25		42.19	.00
Recall homework problems	2.05	.91	.13	2.25	.02
Algorithmic homework problems	2.15	.68	.22	3.15	.02
Textbook examples	-1.37	.63	-.15	-2.16	.03

Note. VIF < 2.00

For the PCA-Lower student percentile, recall and algorithmic homework problems in addition to textbook examples significantly predicted student performance on exam questions. The regression model explained a significant amount of the variance in exam scores, $F(3, 322) = 4.43$; $p = .00$; $R^2 = .03$.

Topic and the effects of categorized exposures. To investigate the effects of the different exposures on the topics studied in the Chem 170/175 courses, the categorical variable of lecture and homework exposures were entered simultaneously with the main exposures of textbook and laboratory, and a backward elimination was performed on each student percentile.

Results for the PCA-All student group are in Table 20 (Chem 170) and Table 21 (Chem 175).

Table 20

All Student Group Regressed on Topics for Chem 170

Variable	B	*SE*(B)	β	*t*	*p*
		Thermochemistry			
Model					
Constant	71.80	3.86		18.56	.00
Recall homework problems	8.02	2.62	.72	3.06	.01
		Electron Configuration			
Model					
Constant	91.17	3.18		28.71	.00
Conceptual lecture exposure	-1.91	.67	-.54	-2.86	.01
		Gases			
Model					
Conceptual lecture exposure	17.44	6.63	.85	2.63	.03
Laboratory exposure	-9.23	3.08	-.97	-2.99	.02

Note. VIF < 2.00.

For the topic of thermochemistry, recall homework problems predicted a significant amount of success of the PCA-All student group. The variable explained 46% of the variance in exam scores, $F(1, 9) = 9.40$, $p = .01$.

Conceptual lecture exposures explained a significant amount of the variance in exam scores on the topic of electron configurations, $F(1, 20) = 8.19$, $p = .01$, $R^2 = .26$.

Finally, for the topic of gases, conceptual lecture presentations and laboratory exposures had a significant role in predicting student success and explained a statistically significant portion of the variance in exam scores, $F(2, 9) = 4.77$, $p = .04$, $R^2 = .41$.

The results of the effects of exposure type on the topics in the Chem 175 course are in Table 21.

Table 21

All Student Group Regressed on Topics for Chem 175

Variable	B	*SE*(B)	β	*t*	*p*
		Nuclear Chemistry			
Model					
Constant	102.61	6.84		15.01	.00
Textbook examples	-3.91	1.77	-.42	-2.21	.04
		Equilibrium			
Model					
Constant	64.96	5.21		12.47	.00
Algorithmic lecture exposure	3.31	1.03	.54	3.20	.00
		Acid-Base Equilibria[a]			
Model					
Constant	97.90	18.651		5.25	.00
Algorithmic lecture exposure	-14.46	3.97	-1.57	-3.62	.00
Conceptual homework problems	-6.90	2.87	-.82	-2.40	.03
Textbook examples	5.20	2.31	.72	2.25	.04
		Aqueous Equilibria			
Model					
Constant	45.83	7.62		6.01	.00
Conceptual homework problems	20.83	9.02	.47	2.31	.03
		Electrochemistry			
Model					
Constant	65.48	6.3		10.40	.00
Textbook example	3.12	1.36	.54	2.30	.04

Note. VIF < 2.00.

Categorized predictor variables predicted the exam success of the PCA-All student group. For the topic of nuclear chemistry, textbook examples made a signification prediction of the success of all students and predicted a significant portion of the variance of exam scores, $F(2, 18) = 5.14$, $p = .02$, $R^2 = .29$.

Algorithmic lecture exposures, conceptual homework problems, and textbook examples predicted a significant amount of student success on questions about acid-base equilibria. These criterion variables explained 50% of the variance in student success, $F(4, 19) = 6.81$, $p = .00$.

Algorithmic lecture exposures predicted a statistically significant amount of variance in student success when testing the topic of equilibria, $F(1, 25) = 10.26$, $p = .00$, $R^2 = .26$. For the topic of additional aqueous equilibria, conceptual homework problems accounted for 18% of the variance in exam scores, $F(1, 19) = 5.33$, $p = .03$, $R^2 = 18$. Finally, textbook examples explained a significant amount of the variance in exam scores, $F(1, 13) = 5.29$, $p = .04$, $R^2 = .24$, for questions on the topic of electrochemistry.

For the PCA-Upper student percentile, the categorical variables of lecture and homework exposures in addition to the main exposure variables of textbook examples and lab experiments were entered simultaneously and backward elimination was performed on the exam scores. The results are in Table 22.

Table 22

PCA-Upper Student Percentile Regressed on Topics for Chem 170/175

Variable	B	*SE*(B)	β	*t*	*p*
		Molecular Geometry			
Model					
Constant	104.98	2.68		39.15	.00
Algorithmic lecture presentations	-2.84	.81	-.65	-3.49	.00
		Nuclear Chemistry			
Model					
Constant	103.70	5.20		19.95	.00
Textbook exposures	-3.42	1.48	-.47	-2.32	.03
		Equilibrium			
Model					
Constant	80.64	3.19		25.29	.00
Algorithmic lecture presentations	2.82	.63	.67	4.46	.00
		Acid-Base Equilibria			
Model					
Constant	44.33	14.18		3.13	.00
Lab exposures	21.31	7.58	.51	2.81	.01
		Aqueous Equilibria			
Model					
Constant	65.83	5.28		12.47	.00
Conceptual homework problems	22.30	6.25	.63	3.57	.02

Note. VIF < 2.00.

For the upper percentile of students, algorithmic lecture presentations predicted a meaningful amount of success in exam scores on questions about molecular geometry and explained 38% of the difference in student success, $F(1, 17) = 12.20$, $p = .00$, $R^2 = .38$. In regards

to the topic of nuclear chemistry, textbook examples accounted for 18% of the variance in exams for this student percentile, $F(1, 19) = 5.36$, $p = .03$.

Each of the three topics on equilibria has a different exposure variable explaining a significant amount of the variance in the exam scores. Algorithmic lecture presentations explained 42% of the variance in scores of equilibria questions, $F(1, 22) = 19.89$, $p = .00$. Laboratory experiments accounted for 23% of the difference in test scores for the subject of acid-base equilibria, and for aqueous equilibria, conceptual homework problems explained 37% of the variance in student scores, $F(1, 19) = 12.74$, $p = .00$.

The categorical variables of lecture and homework exposures in addition to the main exposure variables of textbook examples and lab experiments were entered simultaneously and backward elimination was performed on the exam scores of the PCA-Middle student percentile. The results are in Table 23.

Table 23

PCA-Middle Student Percentile Regressed on Topics for Chem 170/175

Variable	B	*SE*(B)	β	*t*	*p*
		Electron Configuration			
Model					
Constant	93.92	3.15		29.81	.00
Conceptual lecture presentations	-1.97	.66	-.55	-2.96	.01
		Gases			
Model					
Constant	63.67	6.36		10.01	.00
Algorithmic lecture exposures	7.33	3.18	.59	2.31	.04
		Equilibrium			
Model					
Constant	67.07	5.86		11.45	.00
Algorithmic lecture presentations	3.13	1.16	.48	2.70	.01
		Acid-Base Equilibria[a]			
Model					
Constant	103.58	16.18		6.40	.00
Algorithmic lecture exposures	-16.35	3.45	-1.79	-4.74	.00
Conceptual homework problems	-8.12	2.49	-.97	-3.27	.00
Textbook examples	5.62	2.00	.78	2.81	.01
		Electrochemistry			
Model					
Constant	63.01	6.65		9.47	.00
Textbook examples	4.03	1.43	.62	2.81	.02

Note. VIF < 2.0 except for Acid-Base Equilibria.
[a]VIF values for acid-base equilibria < 10.

For the PCA-Middle percentile, conceptual lecture presentations explained 27% of the variance in exam scores on problems about electron configuration, $F(1, 20) = 8.79$, $p = .01$. For the topic of gases, algorithmic lecture exposures accounted for 28% of the variance in exam scores, $F(1, 10) = 5.32$. Textbook examples explained 33% of the difference in exam scores concerning electrochemistry and resolved 33% of the variation in exam scores $F(1, 13) = 7.91$, $p = .02$. Last, algorithmic lecture presentations explained 19% of the variance in equilibria exam questions, $F(1, 25) = 7.26$, $p = .01$.

Success on the topic of acid/base equilibria questions can be predicted by algorithmic lecture exposures, conceptual homework problems, and textbook examples. These variables explained 62% of the variance in exam scores, $F(4, 19) = 10.29$, $p = .00$.

For the PCA-Lower student percentile, the categorical variables that form lecture and homework exposure along with the main exposures of textbook examples and laboratory experience were simultaneously entered and a backward elimination was performed on exam scores. Results for the Chem 170 course are in Table 24, and for the Chem 175 course, results are in Table 25.

Table 24

PCA-Lower Student Percentile Regressed on Topics for Chem 170

Variable	B	*SE*(B)	β	*t*	*p*
		Thermochemistry			
Model					
Constant	48.20	7.12		6.76	.00
Define homework problems	12.36	4.82	.65	2.56	.03
		Electron Configuration			
Model					
Constant	72.58	5.55		13.07	.00
Algorithmic lecture exposure	2.15	1.24	.44	1.73	.1
Conceptual lecture exposure	-4.83	1.56	-.78	-3.10	.01

Note. VIF < 2.00.

For the PCA-Lower student percentile, define homework problems explained 36% of the variance in thermochemistry exam questions, $F(1, 9) = 6.56$, $p = .03$. Algorithmic lecture exposures explained 18% of the variations in exam scores on the topic of equilibria, $F(1, 25) = 6.69$, $p = .02$.

When the exam questions tested for electron configuration topics, 27% of the variance in scores was accounted for by conceptual lecture exposures, $F(2, 19) = 4.89$, $p = .02$.

In Table 25, the results are given for the regression analysis of the PCA-Lower student percentile scores on the topics presented in Chem 175.

Table 25

PCA-Lower Student Percentile Regressed on Topics for Chem 175

Variable	B	*SE*(B)	β	*t*	*p*
		Nuclear Chemistry			
Model					
Constant	99.06	9.98		9.93	.00
Textbook examples	-5.75	2.58	-.42	-2.23	.04
		Equilibrium			
Model					
Constant	44.29	8.71		5.09	.00
Algorithmic lecture exposure	4.46	1.73	.46	2.59	.02
		Acid-Base Equilibria[a]			
Model					
Constant	111.24	21.32		5.22	.00
Conceptual lecture exposure	-8.55	1.69	-.74	-5.05	.00
Algorithmic homework problems	-21.67	5.97	-2.15	-3.63	.00
Textbook examples	11.28	3.58	1.50	3.15	.00

Note. For Acid/Base Equilibria VIF <19.00.

Textbook exposures significantly predicted scores on nuclear chemistry exam questions for the PCA-Lower student percentile and explained 31% of the variance in the scores, $F(2, 18) = 5.54$, $p = .01$. For the topic of equilibria, algorithmic lecture exposures explained 18% of the variance in student success, $F(1, 25) = 6.69$, $p = .02$.

The topic of acid/base equilibria had high VIF values for several of the predictor variables. Collinearity diagnostics results indicate that the conceptual homework problems variable (VIF = 18.04) and algorithmic homework problems (VIF = 4.47) are highly correlated.

Since these two variables are subcategories of homework exposures, they were combined into one variable. When this combination variable was entered into the model, textbook

examples and laboratory exposures became correlated so these two variables were also combined into one variable.

When these combination variables were added to the regression, conceptual lecture presentations significantly predicted the success on exam questions testing the topic of acid/base equilibria, $\beta = -.60$, $t(23) = -3.50$, $p = .00$ and accounted for 33% of the variance in exam scores, $F(1, 22) = 12.20$, $p = .00$.

Interaction Effects

In the previous section, the effect of exposures on student success has been measured by regressing student success on a single exposure or a group of exposures entered simultaneously. In this section, the effects of the exposures on each other will be investigated using interaction terms. Interaction terms are those that demonstrate how two or more independent variables interact to affect the dependent variable. The interaction relationships between these exposures are listed below.

- Interaction terms
    - lecture exposures interacting with homework exposures
    - lecture exposures interacting with textbook exposures
    - lecture exposures interacting with laboratory exposures
    - homework exposures interacting with textbook exposures
    - homework exposures interacting with laboratory exposures
    - laboratory exposures interacting with textbook exposures

Homework problems are assigned to students to increase their understanding of the topics presented in the lecture portion of the class, and for the same reason, students are urged to read their textbooks. These are customary methods teachers use in an attempt to increase student course success. The interaction terms of lecture exposures interacting with homework and

textbook exposures were created to test this practice. In addition, interaction terms were created to test the relationships between the laboratory variable and the other exposure variables.

The laboratory component of a chemistry class is considered to be essential in learning the subject (Nakhleh, 1994). Teaching labs have many methods considered important in educational research such as writing to increase scientific reasoning skills, peer-teaching, and hands-on activities. Learning is hampered when students have an inadequate understanding of the concepts of the experiment and are not able to connect the topics of the lab to the theoretical knowledge presented in the lecture (Nakhleh, 1994). Lectures, textbook examples and homework problems would be expected to improve a student's comprehension. Because of these reasons, interaction terms between the lecture, textbook and homework exposures with laboratory experiments were created and tested.

Interactions of main exposures. The four main exposures, lecture, homework, textbook, and laboratory, plus the interaction terms between each exposure were regressed on by each student percentile. For the PCA-All student group and the PCA-Upper and PCA-Middle percentiles, there were no significant models.

The interaction of homework exposures with laboratory experiences created a significant F change on the performance of the PCA-Lower student percentile, $\Delta F(1, 300) = 4.27$, $p = 0.04$, $R^2 = .01$. There was not a significant model for the main exposures.

Categorized predictor variables. The main effects of lecture and homework exposures were grouped into levels of recall, algorithmic, and conceptual presentations and were used to classify homework assignments and lecture presentations. The grouped predictor variables were entered simultaneously and then the interaction terms were added in a stepwise manner. Only the PCA-All student group was analyzed in this section of the research project.

An explanation of the figures in this section is needed here. Each graph in this section has a truncated y-axis. When examining student success, 6% of student grades fell below the 40% level of percent correct answers. To increase the clarity of the trends, the y-axis starts at this number.

In addition to the changes in the y-axis, both back casting and forecasting were used to aid in the understanding of the trends in the data. Assumptions made for the interpolation of the data include if there is a linear relationship between the number of exposures and student success and that this relationship is constant. For each visual representation in this section, the figures before modification can be found in Appendix C.

Interactions with recall lecture presentations. The effects of recall lecture presentations interacting with other exposure types on student success are investigated in this section. After the categorized, predictor variables were entered; their interaction terms were then entered in the following order: recall lecture interacting with algorithmic homework, recall lecture presentations with conceptual homework, and recall lectures interacting with labs.

The addition of the interaction between recall lecture presentations and laboratory experiments produced a significant F change, $\Delta F(1, 314) = 5.28$, $p = .02$, $\Delta R^2 = .02$. For the final model, significant predictors of student success include recall homework problems, $\beta = .22$, $t(325) = 3.40$, $p = .00$, algorithmic homework problems, $\beta = 1.75$, $t(325) = .24$, $p = .01$, and the interaction between recall lectures and laboratory experiments, $\beta = .16$, $t(325) = 2.30$, $p = .02$. These variables explained 5% of the variance in student success scores, $F(11, 314) = 2.57$, $p = .00$, VIF < 5.0.

A visual representation of the interaction is in Figure 20.

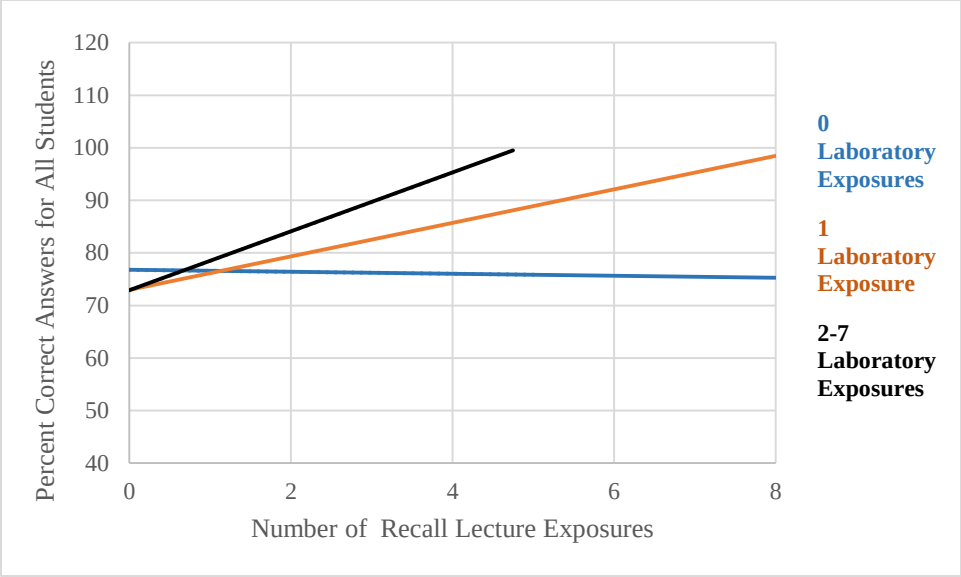


Figure 20. The interaction of recall lecture exposures with laboratory experiments. As the number of recall lectures increases, student success increases when the number of laboratory exposures is large. The original figure is in Appendix C Figure 1.

The interaction term of laboratory experiments with recall lectures shows that a comparatively large number of labs increase the effectiveness of recall lecture exposures. The lack of labs decreases the effect of this type of lecture.

Interactions with algorithmic lecture presentations. All categorized predictor variables were entered simultaneously. The following interactions were entered in a stepwise manner: algorithmic lecture interacting with conceptual homework problems, algorithmic lectures with algorithmic homework, and algorithmic lecture with laboratory exposures The addition of the interaction term of algorithmic lecture with conceptual homework produced a significant F change, $\Delta F(1, 316) = 10.37$, $p = .00$, $\Delta R^2 = .03$.

Significant predictors of student success are recall homework problems, B = 2.88, $t(325) = 3.64$, $p = .00$, algorithmic homework problems, B = 1.45, $t(325) = 2.20$, $p = .03$, and the

interaction between algorithmic lectures and conceptual homework, B = .58, $t(325) = 3.08$, $p = .00$. Together these variables explain 6% on the variance in percent correct answers, $F(11, 314) = 2.99$, $p = .00$, VIF < 3.00.

Figure 21 is a graphical view of this interaction.

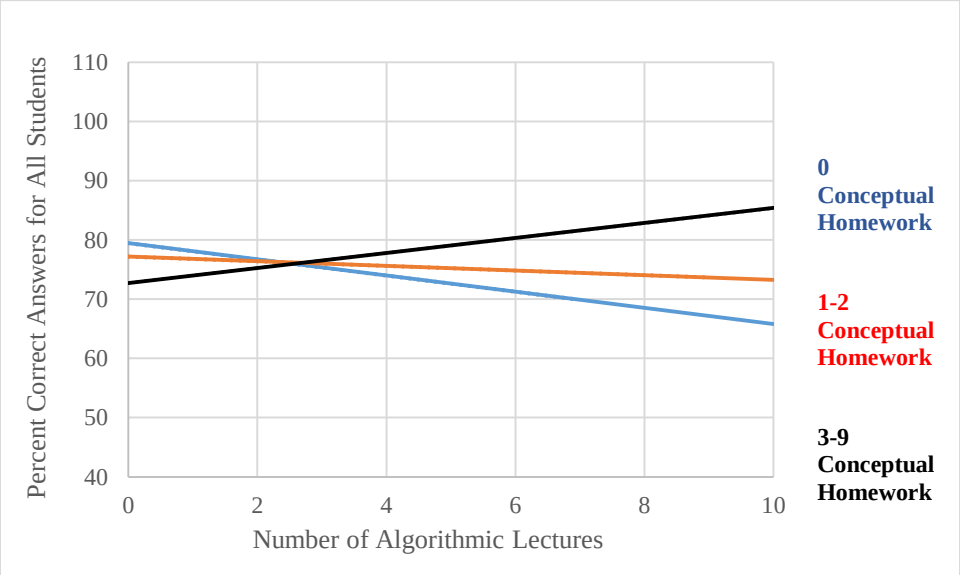


Figure 21. The interaction of algorithmic lectures exposures with conceptual homework problems. Student success increases as the number of algorithmic lectures increase when the number of conceptual homework problems is high. The original figure can be found in Appendix C Figure 2.

The interaction term of algorithmic lectures with conceptual homework problems shows the correlation between the quantities of conceptual homework problems on the effectiveness of algorithmic lectures. Student success increases as the number of algorithmic lecture presentations increase when there are a sizeable number of conceptual homework problems. The number of

lectures decreases student performance when the number of conceptual problems is low or nonexistent.

Interactions with conceptual lecture presentations. All categorized, predictor variables were entered simultaneously, and then the following interactions were entered in a stepwise manner: conceptual lecture interacting with algorithmic homework problems, conceptual lectures with conceptual homework, and conceptual lecture with laboratory exposures.

None of the interaction terms produced a significant F change. Variables that significantly predict student success are recall homework problems, $\beta = .20$, $t(325) = 2.88$, $p = .00$ and algorithmic homework exposures $\beta = .23$, $t(325) = 2.71$, $p = .01$. These variables predict 3% of the variance in student success $F(11, 314) = 1.90$, $p = .04$, VIF < 2.50.

Interactions with textbook examples. All categorized predictor variables were entered simultaneously. The following interactions were entered in a stepwise manner: textbook examples interacting with recall homework problems, textbook with recall lectures, textbook by algorithmic homework, and textbook with conceptual homework. The addition of the interaction term of textbook examples with algorithmic homework problems produced a significant F change, $\Delta F(1, 314) = 6.98$, $p = .01$, $\Delta R^2 = .02$, as did the term of textbook examples interacting with conceptual homework problems, $\Delta F(1, 313) = 8.00$, $p = .00$, $\Delta R2 = .0^2$.

For the final model, the interaction term of textbook examples by conceptual homework problems significantly predicted the student scores, B = .68, $t(325) = 2.83$, $p = .00$. Overall, the interaction term explained 8% of the variance in exam scores, $F(12, 313) = 3.26$, $p = .00$, VIF < 3.50.

The visual representation of the interaction between textbook examples and conceptual homework problems is in Figure 22.

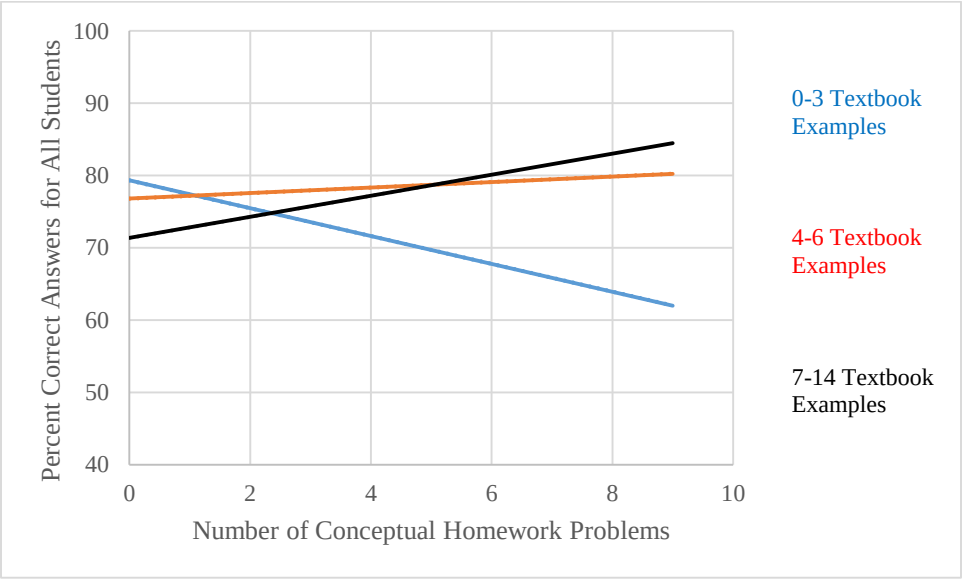


Figure 22. Interactions of textbook examples with the number of conceptual homework problems. With a high number of textbook exposures, an increase in the number of conceptual homework problems increases student success. Original figure is in Appendix C Figure 3.

The effect of conceptual homework problems on student performance increases as the number of textbook examples increases. When there are zero textbook examples, the impact on the number of conceptual homework problems is negative.

Interactions with laboratory exposures. The categorized, independent variables were entered simultaneously followed by the interactions terms of lab by conceptual homework, lab by algorithmic homework, and lab exposures interacting with recall homework. None of the interactions produced a significant F change.

Recall homework exposures significantly predicted student exam scores, β = .21, $t(325) = 3.18$, $p = .00$, as did algorithmic homework, β = .20, $t(325) = 2.18$, $p = .03$. These two predictors predicted 3% of the differences in exam grades, $F(10, 315) = 1.97$, $p = .04$, VIF < 3.00).

Interaction Effects of Main Exposures on Question Difficulty

Each measure of question difficulty has shown a significant, negative effect on all levels of student success. In this section, the effects of the interaction terms on the difficulty level of questions will be investigated. For each regression, the variable of question difficulty was entered first, then the main exposure variables were entered simultaneously, and finally, the interaction variables were entered in a stepwise manner. Using the stepwise entry method allows the detection of any variable that has a significant change in the *F* statistic.

Modified Bloom's taxonomy. The result of a Pearson correlation shows Bloom's taxonomy has a significant negative relationship with the success of each student percentile. The results of the Pearson correlation are in Table 26.

Table 26

Pearson Correlations between Student Success and the Modified Bloom's Taxonomy

Student percentile	Pearson *r*	Significance	Coefficient of Determination (r^2)
PCA-All Students	$r(325) = -.34$**	.00	.12
PCA-Upper	$r(325) = -.25$**	.00	.06
PCA-Middle	$r(325) = -.34$**	.00	.12
PCA-Lower	$r(325) = -.31$**	.00	.10

**Correlation is significant at the .01 level (two-tailed).

The taxonomy has a negative correlation with student success for each student percentile. Each correlation has a medium effect size.

Results from a the multiple regression showed the modified taxonomy explained 12% of variance in exam scores for the PCA-All student group, $F(1, 308) = 43.20$, $p = .00$. For the PCA-Upper percentile, the modified Bloom's taxonomy explained 6% of the variance in exam scores, $F(1, 308) = 19.94$, $p = .00$. Bloom's taxonomy explained 12% of the variance in exam scores for the PCA-Middle student percentile, $F(1, 308) = 40.17$, $p = .00$. The question difficulty scale of the modified Bloom's taxonomy explained a significant amount of variance in student exam success for the PCA-Lower student percentile, $F(1, 308) = 37.24$, $p = .00$, $R^2 = .11$.

Recall, algorithmic, and conceptual question types. The question type method classifies the difficulty of exam questions based on the type of thinking required to answer an exam or homework question. A Pearson correlation shows the relationship between question type and student success in Table 27.

Table 27

Pearson Correlations between Student Success and Question Type

Student percentile	Pearson *r*	Significance	Coefficient of Variation (r^2)
PCA-All students	$r(325) = -.28$**	.00	.08
PCA-Upper	$r(325) = -.19$**	.00	.04
PCA-Middle	$r(325) = -.29$**	.00	.08
PCA-Lower	$r(325) = -.27$**	.00	.07

**Correlation is significant at the .01 level (two-tailed).

For each student group, the difficulty measure of question type has a significant negative relationship with student achievement. Each correlation has a medium effect size.

Results from a multiple regression indicated the question type measure of question difficulty explained 8% of the variance in student exam scores for the PCA-All student group, $F(1, 308) = 26.92$, $p = .00$. For the PCA-upper student percentile, the question type measurement scale explained a significant amount of variance in exam scores, $F(1, 308) = 10.76$, $p = .00$, $R^2 = .03$. Question type explained 9% of the variance in exam scores for the PCA-Middle student percentile, $F(1, 308) = 29.06$, $p = .00$. The question difficulty scale of the question type explained a significant amount of variance in student exam success for the PCA-Lower student percentile, $F(1, 308) = 26.15$, $p = .00$, $\Delta R^2 = .08$.

Cognitive skill level. Cognitive skill level is another method to classify question difficulty. The cognitive skill difficulty scale has a negative correlation with each student percentile. Each correlation has a medium effect size. Results of a point bi-serial correlation for each student percentile are in Table 8.

Results from a multiple regression indicated the cognitive skill level measuring scale explained 10% of the variance in student exam scores for the PCA-All students group, $F(1, 308) = 32.97$, $p = .00$. For the PCA-Upper student percentile, the cognitive skill level explained a significant amount of variance in exam scores, $F(1, 308) = 22.09$, $p = .00$, $\Delta R^2 = .07$. For the PCA-Middle student percentile, the cognitive skill level explained 9% of the variance in exam scores, $F(1, 308) = 30.92$, $p = .00$. For the PCA-Lower student percentile, a significant amount of variance in student exam scores was explained by the difficulty of cognitive level questions, $F(1, 308) = 22.15$, $p = .00$, $R^2 = .07$.

Number of steps. The number of steps measures question difficulty by the complexity of the question. The number of steps has a significant negative relationship with all student percentiles except for the PCA-Upper percentile. The effect size for the significant correlations are small. The results of a Pearson correlation analysis can be found in Table 6.

The number of steps predicted a significant amount of variance in the all student group exam scores, $F(1, 308) = 15.03$, $p = .00$, $R^2 = .05$. For the PCA-Upper student, the number of steps explained 2% of the variance in exam scores, $F(1, 308) = 5.63$, $p = .02$. For the PCA-Middle student percentile, the number of steps explained a significant amount of variance in exam scores, $F(1, 308) = 15.67$; $p = .00$; $R^2 = .05$. Finally, The number of steps explained a significant amount of variation in exam scores for the PCA-Lower percentile, $F(1, 308) = 12.65$, $p = .00$, $R2 = .04$.

Exposures on Different Levels of Question Difficulty

The methods used to classify question difficulty have different levels associated with them. To determine if there are exposures that affect student success at different levels of question difficulty or have an effect on a different student percentile, the categorized main exposures are regressed across each level of question difficulty for each student percentile

For each student percentile, a regression analysis was performed selecting for a certain level of question difficulty. The categorized main exposures variables were entered simultaneously, and a backward elimination was used to determine if any exposure type had a significant effect on student achievement. This method was applied to each student percentile.

Modified Bloom's taxonomy. Bloom's taxonomy classifies question difficulty by the type of knowledge required to solve a problem. To determine the amount variance in student success explained by the predictor variables, all of the categorized main exposures were entered

simultaneously and then a backward elimination was used. Each regression selected for a level of the taxonomy.

Table 28 has the results for the percent corrects answers for the all student group.

Table 28

All Student Group Regressed on the Levels of the Modified Bloom's Taxonomy

Bloom's taxonomy level	B	*SE*(B)	β	*p*
Recall				
Constant	80.03	3.13		
Algorithmic homework problems	3.04	.80	.48	.00
Textbook examples	-1.81	.74	-.31	.02
Comprehension				
Constant	72.74	2.43		
Recall homework problems	2.50	.95	.24	.01
Application[a]	-	-	-	
Analysis[a]	-	-	-	

[a]No significant model.

Both algorithmic homework problems and textbook examples significantly predicted exam scores and explained 12% of the differences in exam grades for all students, $F(2, 88) = 7.30$, $p = .00$, VIF < 2.00.

For the comprehension level of question difficulty, recall homework problems explained 4% of the variance in student exam success, $F(2, 131) = 3.71$, $p = .03$, VIF < 2.00.

The predictor variables did not explain a significant amount of variance in exam scores for the levels of application and analysis.

The procedure used to analyze the PCA-All student group was also used to examine the PCA-Upper student percentile. Results are in Table 29.

Table 29

PCA-Upper Percentile Regressed on the Levels of the Modified Bloom's Taxonomy

Bloom's taxonomy level	B	*SE*(B)	β	*p*
Recall				
Constant	90.83	1.80		
Algorithmic homework problems	.88	.41	.22	.04
Comprehension				
Constant	89.39	1.62		
Recall homework problems	1.53	.64	.22	.02
Algorithmic homework problems	.81	.36	.21	.03
Application[a]	-	-	-	
Analysis[a]	-	-	-	

[a]No significant model.

Algorithmic homework problems explain a significant portion of the variance in student performance on recall questions, $F(1, 89) = 4.56$, $p = .04$, $R^2 = .04$, VIF < 2.00.

For the level of comprehension questions, both recall and algorithmic homework problems predicted a significant amount of student success of the PCA-Upper student percentile and explained 2% of the variance in the exam scores, $F(2, 131) = 3.97$, $p = .02$, VIF < 2.00.

The taxonomy levels of application and analysis have no exposure variables that account for a statistically significant amount of variance in student success.

The process of entering the predictor variables and then conducting a backward elimination was used for the PCA-Middle student percentile. The results are in Table 30.

Table 30

PCA-Middle Percentile Regressed on the Levels of the Modified Bloom's Taxonomy Levels

Bloom's taxonomy level	B	*SE*(B)	β	*p*
Recall				
Constant	78.10	3.93		
Recall homework problems	2.06	1.02	.22	.05
Algorithmic homework problems	2.23	.69	.34	.00
Comprehension				
Constant	73.21	3.08		
Recall homework problems	2.54	.96	.22	.01
Application[a]	-	-	-	
Analysis[a]	-	-	-	

[a]No significant model.

For the recall questions, both recall and algorithmic homework problems predicted a significant amount of the success for the PCA-Middle student percentile. These variables explained 11% of the differences in exam scores for the recall level of Bloom's taxonomy, $F(3, 87) = 4.84$, $p = .00$, VIF< 2.00.

Recall homework problems account for 6% of the variance in student achievement on comprehension questions, $F(2, 131) = 5.18$, $p = .01$, VIF < 2.00. There were no significant models for either the application or analysis level for this student percentile.

The same process used to analyze the PCA-All student group, PCA-Upper, and the PCA-Middle student percentiles was applied to the PCA-Lower student percentile scores. Results are in Table 31.

Table 31

PCA-Lower Percentile Regressed on the Levels of the Modified Bloom's Taxonomy

Bloom's taxonomy level	B	*SE*(B)	β	*p*
Recall				
Constant	64.23	4.66		
Textbook exposures	-3.69	1.10	-.41	.00
Algorithmic homework problems	5.25	1.19	.54	.00
Comprehension[a]	-	-	-	
Application[a]	-	-	-	
Analysis[a]	-	-	-	

[a]No significant models.

For the PCA-Lower student percentile, the exposure variables of algorithmic homework problems and textbook examples explained 17% of the variance in student performance, $F(2, 88) = 10.10$, $p = .00$, VIF < 2.00.

For the levels of comprehension, application, and analysis question difficulty, there are no regression models that explain a significant amount of difference in student success.

Recall, algorithmic and conceptual question types. The difficulty measure of question type classifies problems by the mode of learning and the cognitive processes needed to answer a question. Each level of question type was analyzed by selecting for a question level and then entering each categorized variable simultaneously. Then, applying a backward elimination. The same method was used on each student percentile. The results for the PCA-All student group are in Table 32.

Table 32

PCA-All Student Group Regressed on the Levels of Question Type

Level	B	*SE*(B)	β	*p*
Recall[a]	-	-	-	
Algorithmic[a]	-	-	-	
Conceptual				
Constant	66.24	2.34		
Recall homework problems	2.56	1.17	.19	.03

[a]No significant model.

The exposures variables did not predict a significant amount of success for the recall and algorithmic levels for the PCA-All student group. Recall homework problems accounted for 3% of the differences in exam grades for conceptual level questions, $F(1, 125) = 4.76$, $p = .03$, VIF < 2.00

The results of the analyses of the PCA-Upper percentile are in Table 33.

Table 33

PCA-Upper Percentile Regressed on the Levels of Question Type

Level	B	*SE*(B)	β	*p*
Recall[a]	-	-	-	
Algorithmic				
Constant	92.18	1.33		
Algorithmic lecture exposure	.79	.38	.19	.04
Conceptual				
Constant	80.91	2.73		
Recall homework problems	2.70	1.06	.23	.01

[a]No significant models.

For this PCA-Upper student percentile, there was not a regression model that would predict a significant amount of student success on recall questions. For the algorithmic level, algorithmic lecture exposures explained a significant amount of variance for the achievement of this percentile, $F(1, 116) = 4.40$, $p = .04$, $R^2 = .03$, VIF < 2.00. For conceptual questions, laboratory exposures and recall homework problems accounted for 4% of the differences in exam scores, $F(2, 124) = 3.80$, $p = .02$, VIF < 2.00.

The same process used on the two previous student percentiles or groups was applied to the PCA-Middle percentile. The results are in Table 34.

Table 34

PCA-Middle Percentile Regressed on the Levels of Question Type

Level	B	*SE*(B)	β	*p*
Recall				
Constant	76.73	3.43		
Algorithmic homework problems	1.53	.70	.24	.03
Algorithmic				
Constant	80.60	2.93		
Recall homework problems	2.90	1.38	.20	.04
Conceptual				
Constant	67.11	2.45		
Recall homework problems	2.94	1.28	.20	.02

For the recall level of question difficulty, algorithmic homework problems accounted for 6% of the variance in exam scores for this PCA-Middle student percentile, $F(2, 78) = 3.43$, $p = .04$, VIF < 2.00.

Recall homework problems resolved 5% of the variance in exam performance at the algorithmic difficulty level, $F(3, 114) = 2.86$, $p = .04$, VIF < 2.00.

For conceptual level problems, recall homework problems explained 3% of the difference in exam scores for the PCA-Middle student percentile, $F(1, 125) = 5.26$, $p = .02$, VIF < 2.00.

The results for the analysis of the PCA-Lower student percentile are in Table 35. The analysis method used on the previous percentiles was also used for this percentile.

Table 35

PCA -Lower Percentile Regressed on the Levels of Question Type

Recall/Algorithmic/Conceptual	B	*SE*(B)	β	*p*
Recall				
Constant	59.00	4.09		
Algorithmic Homework Problems	2.30	1.02	.25	.03
Algorithmic[a]	-	-	-	
Conceptual[a]	-	-	-	

[a]No significant models.

For recall questions, algorithmic homework problems explained 5% of the variance in the PCA-Lower student percentile exam scores, $F(1, 79) = 5.14$, $p = .03$, VIF < 2.00.

For the levels of algorithmic and conceptual type questions, there were no regression models that explained a significant amount of the variance in student performance.

Cognitive skill level. The cognitive skill level classifies question difficulty by the cognitive processes required to answer questions. Each student percentile was analyzed by linear regression selecting for a certain level of the measurement scale. All predictor variables were entered simultaneously and then a backward elimination was performed.

The results for the PCA-All student group are in Table 36.

Table 36

PCA-All Student Group Regressed on Cognitive Skill Level

LOCS/HOCS	B	*SE*(B)	*β*	*p*
LOCS				
Constant	72.76	2.06		
Recall homework problems	2.02	.68	.20	.00
Algorithmic homework problems	1.41	.43	.22	.00
HOCS[a]	-	-	-	

[a]No significant model.

Recall and algorithmic homework problems explained 6% of differences in exam scores for this percentile, $F(2, 219) = 7.44$, $p = .00$, VIF < 2.00. The predictor variables did not explain a significant amount of the variance in scores for HOCS questions.

For the PCA-Upper student percentile, the categorized predictor variables were entered simultaneously and then a backward regression was performed. Results are in Table 37.

Table 37

PCA-Upper Student Percentile Regressed on Cognitive Skill Level

LOCS/HOCS	B	*SE*(B)	β	*p*
LOCS				
Constant	89.65	1.38		
Recall homework problems	.94	.46	.14	.04
Algorithmic homework problems	.87	.29	.21	.00
HOCS[a]	-	-	-	

[a]No significant model.

The categorized exposures variables of recall and algorithmic homework problems accounted for 4% of the difference in exam scores on questions requiring lower-order cognitive skills, $F(2, 219) = 5.26$, $p = .01$, VIF < 2.00. The predictor variables did not explain a significant amount of variance of scores on questions requiring higher-order cognitive skills.

The same process used to analyze the PCA-Upper percentile was used on the PCA-Middle student percentile. Results are in Table 38.

Table 38

PCA-Middle Student Percentile Regressed on Cognitive Skill Level

LOCS/HOCS	B	*SE*(B)	β	*p*
LOCS				
Constant	74.41	2.21		
Recall homework problems	2.54	.73	.24	.00
Algorithmic homework problems	1.56	.46	.23	.00
HOCS[a]	-	-	-	

[a]No significant models.

For the PCA-Middle student percentile, recall and algorithmic homework problems predicted a statistically significant amount of student success and explained 7% of the variance in exam scores on lower-order cognitive skill questions, $F(2, 219) = 9.07$, $p = .00$, VIF < 2.00. For the higher-order cognitive skill questions, none of the predictor variables explained a significant amount of differences in student performance.

For the PCA-Lower percentile, the categorized exposure variables were entered simultaneously and then a backward elimination was performed to determine the amount of

variance the variables could predict. There were no significant effects of exposure variables on cognitive skill levels.

Number of steps. The number of steps measure question difficulty by the complexity of the questions. Each student percentile was analyzed by multiple regression. All predictor variables were entered simultaneously and then a backward elimination was performed. Results are in Table 39.

Table 39

All Student Group Regressed on the Number of Steps

Variable	B	*SE*(B)	β	*t*	*p*
Model					
Constant	78.46	3.62		21.70	.00
Number of steps	-2.19	.78	-.19	-2.92	.00
Recall homework problems	2.27	.52	.21	2.98	.00
Algorithmic homework problems	1.43	.60	.19	2.36	.02

Note. VIF < 1.00.

For the PCA-All student group, recall, β = .18, $t(325) = 3.10$, $p = .00$, and algorithmic, β = .14, $t(325) = 2.39$, $p = .02$, homework problems along with the number of steps, β = -.16, $t(325) = -3.02$, $p = .00$, predicted a statistically significant amount of variance in exam scores, $F(3, 322) = 7.15$, $p = .00$, $R^2 = .06$, VIF < 1.00.

The same procedure was used to analyze the PCA-Upper student percentile. Results are in Table 40.

Table 40

PCA-Upper Percentile Regressed on the Number of Steps

Variable	B	*SE*(B)	β	*t*	*p*
Model					
Constant	86.47	1.62		53.55	.00
Recall homework problems	1.49	.55	.16	2.73	.01
Algorithmic homework problems	.87	.34	.15	2.60	.01

Note. VIF < 1.00.

For the PCA-All upper one-fourth group, recall homework problems predicted a significant amount of exam success, β = .16, $t(325) = 2.73$, $p = .01$. Algorithmic homework problems predicted a statistically significant amount of exam success, β = .15, $t(325) = 2.60$, $p = .01$. These two variables explain 3% of variance in exam scores, $F(2, 323) = 5.36$, $p = .00$, VIF < 1.00.

The PCA-Middle percentile was analyzed by multiple regression with all predictor variables entered simultaneously and then a backward elimination was performed. Results are in Table 41.

Table 41

PCA-Middle Percentile Regressed on the Number of Steps

Variable	B	*SE*(B)	β	*t*	*p*
Model					
Constant	76.83	2.75		27.95	.00
Number of steps	-2.44	.79	-.17	-3.08	.00
Recall homework problems	2.56	.74	.20	3.47	.00
Algorithmic homework problems	1.05	.45	.13	2.23	.02

Note. VIF < 1.00.

For the PCA-All middle percentile, recall, β = .20, $t(325) = 3.47$, $p = .00$, and algorithmic, β = .13, $t(325) = 2.23$, $p = .02$, homework problems along with the number of steps, β = -.17, $t(325) = -3.08$, $p = .00$, predicted a statistically significant amount of variance in exam scores, $F(3, 322) = 8.35$, $p = .00$, $R^2 = .06$, VIF < 1.00.

The PCA-Lower one-fourth percentile was analyzed by multiple regression with all predictor variables entered simultaneously and then a backward elimination was performed. Results are in Table 42.

Table 42

PCA-Lower Percentile Regressed on the Number of Steps

Variable	B	*SE*(B)	β	*t*	*p*
Model					
Constant	58.85	3.73		15.76	.00
Number of steps	-2.66	.97	-.15	-2.73	.01
Recall homework problems	1.87	.91	.12	2.06	.04
Algorithmic homework problems	2.02	.68	.21	2.97	.00

Note. VIF < 1.00.

For the PCA-Lower one-fourth percentile, recall homework problems, β = .12, $t(325) = 2.06$, $p = .04$, and algorithmic homework problems, β = .21, $t(325) = 2.97$, $p = .00$, along with the number of steps, β = -.15, $t(325) = -2.73$, $p = .01$, predicted a statistically significant amount of variance in exam scores, $F(4, 321) = 5.25$, $p = .00$, $R^2 = .05$, VIF < 1.00.

Discussion

The goal of this research project was to find ways to help novice learners be more successful in learning chemistry by examining the relationship between the exposure variables and exam performance. Both the cognitive load and constructivist theories are used to understand both the source of problems with learning chemistry and possible solutions to these problems.

First, the effects of the exposure variables and the difficulty scales will be discussed for all students in the class. Next, the effects of these variables on the upper one-fourth, middle one-half and the lower one-fourth student percentiles will examined.

Exposure Variables

The exposure variables are based on the traditional methods of teaching. For chemistry, this means lectures, laboratory experiments, homework assignments, and textbooks. These exposure variables were tested to determine their effect on student exam scores for a freshman chemistry course at a large Midwestern university.

Number of exposures. For the main exposure variables, the number of exposures did not affect student success on exam questions. In addition, interaction terms composed of these exposure variables had no effects on student exam success.

A possible reason for the lack of significant effects is the inability to determine the number of exposures a student actually received in each learning environment. The number of students attending lecture fluctuated throughout each semester so the number of lecture exposures received is unclear. This is the same problem with textbook examples. There is no way to determine the number of times the students used the sample exercises in the textbook. Both variables have uncertain values for the number of exposures resulting in an imprecise measuring tool.

In addition, certain topics that are deemed difficult by the instructor may have more exposures than those they consider less cognitively demanding. This could decrease the effects of the lecture and homework exposure variables.

Location of exposures. The location of the exposure did not affect student exam scores.

Categorized variables. The categorized variables represent the method of exposure for lecture and homework variables. Of the categorized variables, recall and algorithmic homework problems increase student exam success.

For every one increase in the number of recall homework problems, there is a 2.20 increase in percent correct answers for the all student group while holding all other variables constant. For each one increase in algorithmic homework problems, there is a 1.54 increase in the percent correct answers in the dependent variable while controlling for all other variables.

Recall homework problems have a positive effect on student exam success by increasing the understanding of the terminology and symbolic representations of a topic. When students do not have this understanding, they will have to use the limited amount of working memory to search for this information (Garnett et al., 1995; Nakhleh, 1992; Tabor, 2013). Recall homework problems can decrease the demand on available working memory.

Algorithmic homework problems are the means to practice problem solving. Most college students can process five pieces of information at once in their short term memory (Bunce, 2005). A question requiring six pieces of information to solve will lead to cognitive overload. Practice will consolidate steps into chunks of information that will decrease the drain on the working memory (Gulacar et al., 2013; Sirhan, 2007). This could be one reason algorithmic homework problems have a positive effect on student exam success.

Interactions of categorized variables. Recall and algorithmic homework problems have a positive effect on student exam scores. Now, the effects of combining the categorized exposure variables into interaction terms will be examined. The percent correct answers for the all student group were regressed on the different interaction terms.

Recall lectures interacting with laboratory exposures. The visual representation of this interaction is in Figure 20. The regression equation with unstandardized beta coefficients is $\hat{y} = 2.64(\text{recall homework}) + 1.75(\text{alg. homework}) + 2.05(\text{recall lectures x lab})$.

When the number of laboratory experiments is large as compared to a medium or small number of experiments, an increase in recall lectures increases student success on chemistry exam questions. Each laboratory experiment has pre-labs, brief introductions to the topic, and the lab itself. Each interaction uses the terminology of the topic. When the topic is discussed in the lecture, the students have been exposed to the terms, variables, and symbolic representations several different times and in various ways, and this should reduce the demand on short-term memory allowing students to focus on other aspects of the lecture.

Algorithmic lectures interacting with conceptual homework. The interaction term of algorithmic lectures with conceptual (conc.) homework (hwk.) problems increases student success on exam questions. Visual representation of the interactions is in Figure 21. The regression equation is $\hat{y} = 2.88(\text{recall hwk}) + 1.45(\text{alg. hwk}) + .58\ (\text{alg. lecture x conc. hwk})$.

When there are a large number of conceptual homework problems, increasing the number of algorithmic lectures increases student exam success. Conceptual problems require students to understand the connections between subtopics. As students work through the conceptual homework problems, the relationships between the variables become more understandable, and the cognitive demand of the topic decreases. Conceptual understanding of a topic help students to understand the reasons for each step of solving a problem (Phelps, 1996).

Conceptual homework problems interacting with textbook examples. The effect of increasing the number of conceptual homework problems raises student performance when there are a large number of textbook examples. Figure 22 has the visual representation of the interaction.

Textbooks provide worked examples, definitions of the symbols and terminology of a topic. Graphical representations demonstrate the relationships between variables, and visual representations can be used as a basis to form mental models. Each of these could increase a student's conceptual understanding of a topic.

Question Difficulty Scales

Difficulty scales classify problems by different cognitive criteria. Each scale starts at a lower level of cognitive demand and increases as the tiers of the scale increase. The effects on the student exam success of the difficulty scale and the effects of exposure variables on each level of question difficulty are discussed in this section.

Number of steps. The number of steps is a continuous variable that measures question difficulty by the cognitive complexity of a problem. As the number of steps increases, the success on exam questions decreases. These results are similar to those reported in the Hartman and Lin (2011) paper.

Recall and algorithmic homework problems decrease the effects of increasing cognitive complexity. The unstandardized beta coefficient (B) for the number of steps is -2.40. For every one increase in the number of steps, there is a 2.40 decrease in the percent correct answers. When the effects of recall and algorithmic homework problems are added to the regression equation, the unstandardized beta coefficient becomes -2.19 while controlling for all other variables.

The exposure variables reduce the negative effects of increasing the number of steps. The cognitive complexity of an exam question can be decreased when students recognize and comprehend the terminology of the topic, which reduces the cognitive demand on working memory. Students can be more successful at problem solving when they do not have to search for the meaning of words and symbols as they work to solve a problem.

In addition, homework problems increase students' time on task, and on-line algorithmic homework problems provide practice solving problems with immediate feedback on their performance. Practice solving these problems can reduce cognitive demand by incorporating more information into an element that frees up more working memory to solve problems.

Furthermore, practice can increase automation of certain steps in problem solving. As an example is calculating the molar mass of a substance. When students are first introduced to the concept, some can struggle with calculating this quantity. Then, students have to find the molar mass to solve percent composition questions and then, the molar mass becomes a step in finding the ratios of reactants and products. The act of calculating the molar mass is no longer a challenge. It has become an automated step.

Cognitive skill level. This question difficulty measure is divided into levels of higher-order cognitive skills (HOCS) and lower-order cognitive skills (LOCS). Students scored higher of LOCS questions than on HOCS. These results are similar to those found in the literature (Zoller, 1993; Zoller & Tsaparlis, 1997). For the students in this study, for every one increase in the number of HOCS questions, there is a -12.49 decrease in student success.

Effects of exposures on each level of the cognitive skill level. The variables of recall and algorithmic homework problems increase student success on LOCS exam questions. By

definition, LOCS questions include definitions and algorithmic problems. Recall and algorithmic homework problems should increase student exam success.

For the HOCS level, there are no exposure variables that affect student exam success. This level is difficult because the questions focus on the conceptual understanding of topics and the use of abstract models. The cognitive demand of these questions can exceed the cognitive abilities of students at the concrete operational or transitional stage and lead to confusion and unsuccessful attempts at solving the problem (Bird, 2010; Shayer, 1987).

Bloom's taxonomy. The effects of this difficulty measure decrease the number of percent correct answers on exam questions. The exposure variables of recall and algorithmic homework problems decrease the negative effects of the taxonomy.

The taxonomy classifies questions by the type of knowledge required to solve the problem. There are four levels in this modified version. The lowest level is recall, followed by comprehension, application, and then the analysis level, considered the most cognitively demanding. The results show the levels of recall and comprehension have no significant difference in means, which is true for the means of application and analysis questions. In effect, the taxonomy has been reduced to two levels with recall and the comprehension levels as the lower-level tier and application and analysis as the higher-level tier. These results are similar to those found in the research literature written by Nevid and McClelland (2013). These researchers used the terms of identify and describe to represent the levels of knowledge and comprehension questions. Explain represents the application level, and apply denotes the remaining tiers of analysis, synthesis, and evaluation. Their research showed the same grouping of the lower-level tiers and the top-level tiers (Nevid & McClelland, 2013).

The different levels of Bloom's taxonomy have different effects on student exam success. Recall questions increase student exam success, for every one-unit increase in the number of recall questions, there is 5.37 increase in percent correct answers while holding all other variables constant. Recall questions are lowest in cognitive demand, and success on this level would be within reach of all students in the study (Johnsone & El-Banna, 1986). Comprehension questions increase student exam success, for every one-unit increase in comprehension questions, there is a 3.14 increase in percent correct answers while holding all other variables constant.

Both the levels of application and analysis decrease student success. For the application level, every one-unit increase in the number of application questions, there is a 9.91decrease in percent correct answers while holding all other variables constant. For every one-unit increase in analysis questions, there is a 12.87 decrease in percent correct answers while holding all other variables constant. These levels of the taxonomy are the most cognitively demanding focusing on the higher-order cognitive skills of conceptual understanding and theoretical models, and the cognitive demand of the questions may exceed the students' processing abilities of their working memory.

Effects of exposure variables on each level of Bloom's taxonomy. Student exam success was regressed on the exposure variables for each level of the taxonomy. The exposure variables that predict a significant amount of the variance in student achievement will be discussed.

For the recall level of the taxonomy, algorithmic homework problems increase student exam success, and textbook examples decrease student exam success. The end-of-chapter questions and homework problems for most chemistry textbooks are predominately written as word problems (Bunce, 2005), and therefore, the terminology of the topic is used in wording the

questions. Increasing the number of algorithmic homework problems can help students become familiar with the definitions and symbolic representations of the topic.

In contrast, textbook examples have a negative effect on student success on recall questions. Since the textbook is written using the same wording as is used in the algorithmic homework problems, it would seem the two exposure variables would complement each other. One explanation could be that there is no objective measure of how many times a student took advantage of the textbook, which could contribute to the contradiction of the results.

For comprehension level questions, algorithmic homework problems increase student exam success. Solving word problems on a specific topic require students to understand which variables are relevant and which are not. Some comprehension questions can require substitution of variables in formulas. The more students practice solving problems the more likely they are to become familiar with the relationships between variables, which can increase success on comprehension questions (Solaz-Portoles & Lopez, 2008).

For application and analysis questions, no exposure variable predicted a significant amount of the variance in student exam scores. For these levels, a student's ability to think abstractly may not be developed enough for exposure variables to have any effect (Bird, 2010; Bunce, 2005; Cowan, 2014; Igaz & Proksa, 2012; Sweller et al., 1998).

Question type. Recall and algorithmic questions increase student exam success. There is not a statistically significant difference in the means of recall and algorithmic questions. Conceptual problems decrease student exam success. These results are similar to those found by other researchers in chemical education (Nurrenbern & Pickering, 1987; Papaphotis & Tsaparlis, 2008).

For every one increase in recall questions, there is an 8.17 increase in percent correct answers while holding all other variables constant. For every one increase in the number of algorithmic level questions, there is an increase of 4.70 percent correct answers for all students while holding all other variables constant. For every one increase in the number of conceptual questions, there is a 10.98 decrease in student exam success while controlling for all other variables.

Conceptual questions require abstract thinking. As it is for the other measures of question difficulty, some students at this stage in their cognitive development may not have a fully developed ability to think abstractly (Bird, 2010).

Effects of exposure variables on each level of question type. For the level of recall questions, there are no categorized exposure variables that have a significant effect on student exam success. This is true for the algorithmic level. For the conceptual level, recall homework problems has an effect on student exam success.

For both the cognitive-skill level and the modified Bloom's taxonomy difficulty scales of cognitive, questions from the more cognitively demanding levels decreased student exam success. In addition, there are no exposure variables that increased student success. For the HOCS and the application and analysis levels, questions require students to apply their pre-existing knowledge to new situations and to be able to build abstract models. For students in the transitional or concrete operational stages, these questions would be very difficult and confusing for them. Recall and algorithmic homework problems would not have an effect on the success of these levels of question difficulty.

For the lower levels of these question scales, the cognitive demand is lower. Questions from the lower tiers increase student exam success. The exposure variables that increase success on these questions are recall and algorithmic homework problems.

For the question-type, difficulty scale, the levels of recall and algorithmic increase student success while the conceptual level questions decrease student success. Categorized exposure variables do not have a significant effect on both the recall and algorithmic question levels. The means between these levels are not statistically different, and for these two question types, students' exam scores are sufficiently high they cannot be increased in a significant amount.

Recall homework problems decreased the difficulty of the conceptual level questions. Knowing the terminology of a topic could increase student understanding.

Topics. Some topics are more difficult for students compared to others because of the high degree of interactivity between elements of the topic (Gulacar et al., 2013; Johnstone, 1983; Sirhan, 2007). For this section, the topics with which students had the most success and those topics deemed the most difficult will be examined. Lastly, the effects of the exposure variables on topic difficulty will be reported.

Students scored higher on the topics of matter, electron configuration, nuclear chemistry and chemical equilibrium compared to the topics of acid-base and additional aqueous equilibria. In Appendix A, there is a list of topics with their subtopics.

Matter, electron configuration and nuclear chemistry are topics that should be taught in high school chemistry courses (American Association for the Advancement of Science, 1993; Kansas State Department of Education, 2007; Texas Essential Knowledge and Skills). Prior exposure to these topics should reduce the cognitive demand (Rittle-Johnson et al., 2009; Seery & Donnelly, 2012).

The most difficult topics for these students are those with high element interactivity and those requiring prior knowledge that the students do not have. The topics of equilibrium, acid-base equilibria, and additional aqueous equilibria will be used to illustrate the idea of interactivity between elements.

In the Chem 175 course, the topic of equilibrium is divided into three chapters: equilibrium, acid-base equilibria and additional aspects of aqueous equilibria (Brown et al., 2012). Each chapter builds upon the previous one. This is illustrated in Figure 23.

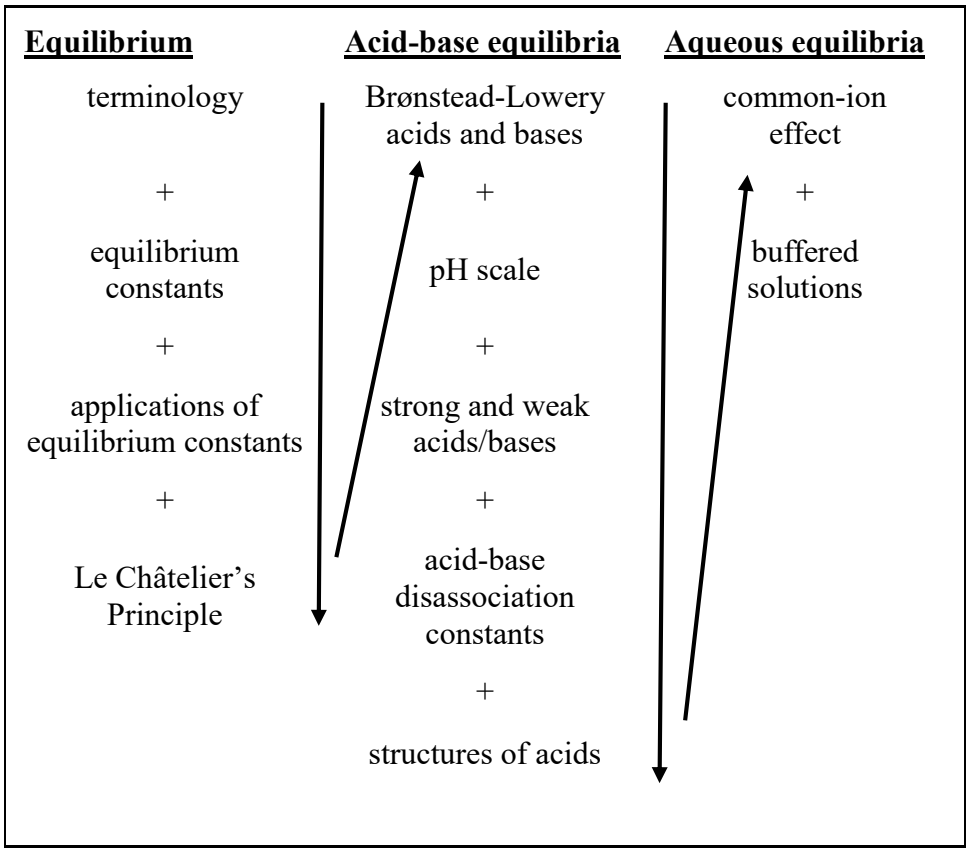


Figure 23. Cognitive complexity of the topics based on the concept of equilibrium. Each topic builds upon the previous topic resulting in increasing element interactivity.

Each new topic builds on the concepts from the previous topics. The interconnectivity of the different subtopics increases as each chapter is studied. To calculate the pH of a solution, the student has to understand the concept of equilibrium constants. To understand how to make and use buffers, the learner has to comprehend Le Châtelier's Principle. The working memory of students can become overwhelmed, and pupils become confused. Then, students are less successful answering questions on these topics.

In addition, these topics are very abstract, and students are more likely to form misconceptions when incorporating the material into their pre-existing conceptual frameworks.

One major misconception about equilibrium is that when a system is at equilibrium no further reactions occur (Akkus, Kadayifçi, Atasoy, & Geban, 2003; Nakhleh, 1992; Niaz, 2001).

Topic and effects of categorized exposures. The topics affected by the exposure variables are not necessarily the ones with which students have the most difficulty. These are the topics that exposure variables can affect and either increase or decrease student exam success.

Recall homework problems predict a significant amount of student test achievement on the topic of thermochemistry. This topic discusses heat energy and the variables that describe its properties. For many students these are new concepts, and the definitions and symbolic representations are unfamiliar. Recall homework problems could be beneficial to students by decreasing the cognitive demand on these novice learners (Gulacar et al., 2013). Students can focus on the conceptual or algorithmic aspects of the topic without having to search their memory for the definitions.

Conceptual lecture presentations have a negative effect on questions about electron configuration. Students can be successful on questions about this topic by either using rote learning to produce the electron configurations of elements or to understand the underlying concepts of the topic. When lectures present new concepts about atomic structure, and the new information does not fit into the student's pre-existing conceptual framework, the students will experience disequilibration. To reduce this stress, students may choose to ignore the new concepts in favor of their already established frameworks they have constructed (Nakhleh, 1992; Niaz, 2001).

Both conceptual lectures and laboratory experiments have an effect on student grades when testing on the topic of gases, but conceptual lectures increase student success while the laboratory experiments decrease success. In the lecture portion of the course, the course

instructor used demonstrations and hands-on student activities to illustrate the different properties of gases. For students at the concrete operational stage, these activities can provide a foundation for students to build abstract models.

Next, the instructor attached symbols to the different physical properties of gases. She used symbolic language to describe the physical properties of gases the students had just observed. The course instructor provided physical models to support students in building conceptual understanding.

The chemistry laboratory decreased student exam performance on the topic of gases. There was one lab experiment on gases, and the physical properties of gases were used to determine the average molecular weight of air. The lab was confusing to students not because of the topic but because the process of finding the mass of air using helium gas as a reference. This procedure was confusing to them, and the confusion negated any benefits of the lab.

Comparison of Student Percentiles

The effects of the exposure variables, topics, and question difficulty scales on the class as whole has been investigated. Now, the analyses will focus on how exposures affect the success of students in the upper one-fourth, the middle one-half, and the lower one-fourth student percentiles.

In the previous section, unstandardized beta coefficients were used to describe the relationships between the exposure variables and student exam success. Unstandardized variables cannot be used across models, but standardized beta coefficients can. For the following comparisons between student percentiles, the standardized beta coefficients are used.

Exposure variables. The exposure variables did not predict a significant amount of variance in exam scores for any of the student percentiles. As it was for the all students'

analyses, the lecture and textbook exposure variables are not precise enough to produce valid results.

Learning environment. The location of exposures did affect student exam success. Homework problems have a positive effect on student exam scores for each student percentile. The on-line assignments provide immediate feedback on student's understanding of a topic, provide for more time spent on task, and increase student's time management skills (Butler & Zerr, 2005; Zerr, 2007).

Question difficulty. Each difficulty measuring scale was analyzed to determine the effect on each student percentile. In addition, the categorized variables regressed on each level of the difficulty scales to determine which exposure variables have the effect of decreasing the complexity of these questions.

Question type on each student percentile. For the upper one-fourth and middle one-half student percentiles, there is no statistical difference in the means for the recall and algorithmic levels. This implies the cognitive demand of recall questions is approximately the same as the difficulty of algorithmic questions. For the lower one-fourth student percentile, the means for the recall (M = 66.11, SD = 24.18), algorithmic (M = 57.47, SD = 22.44), and conceptual (M = 49.08, SD = 24.36) levels of questions differ. For this percentile, students score higher on recall questions than on algorithmic questions. The cognitive demand of algorithmic questions is greater than for recall questions.

Recall questions are the least cognitively demanding for this measurement scale. Three-fourths of the class scored approximately the same on algorithmic questions as on recall ones indicating algorithmic questions are not difficult for these students. Whereas, the lower one-fourth of students still find algorithmic questions more challenging than recall questions.

The exposure variables that affect the difficulty of each level are in Table 43.

Table 43

Question Type and Significant Effects of Exposure Variables

Exposure variables	PCA-Upper 1/4		Middle 1/2		Lower 1/4	
	β	ρ	β	ρ	β	ρ
Recall Level						
Algorithmic homework			.24	.03	.25	.03
Algorithmic Level						
Algorithmic lectures	.19	.04				
Recall homework			.20	.04		
Conceptual Level						
Recall homework	.23	.01	.20	.02		

For the upper one-fourth student percentile, exposure variables do not have a significant effect on the means of recall questions. For these students, scores on recall exam questions are sufficiently high enough that exposure variables will not increase exam success in any significant amount.

For the middle one-half percentile, the exposure variables have a significant effect on each level of question type. These students are successful on the recall level questions, but exposure variables can still affect their scores.

For the lower-one-fourth student percentile, the exposure variables have a significant effect on the least cognitively demanding level of recall. However, exposure variables have no statistically significant effect on algorithmic and conceptual questions. The algorithmic and conceptual question levels are too cognitively demanding for the exposure variables to have an effect for this percentile.

Bloom's taxonomy. Each level of the taxonomy was analyzed to determine the effect each has on student exam success. Then, the exposure variables were investigated to determine what effect they had on exam success for the individual levels of question difficulty.

For the upper one-fourth and middle one-half percentile, there is not a statistically significant difference in means between the recall and comprehension levels of questions. For these students, the cognitive demand of recall and comprehension questions are approximately the same.

For the lower one-fourth percentile, there is a significant difference in the means between these two levels. For these students, comprehension questions (M = 56.77, SD = 21.33) are more difficult than the recall questions (M = 66.49, SD = 24.37). A summary of the effects of exposure variables on each level of the taxonomy are in Table 44.

Table 44

Bloom's Taxonomy and Significant Effects of Exposure Variables

Exposure variables	Student percentile: Upper 1/4th β	Upper 1/4th ρ	Middle 1/2 β	Middle 1/2 ρ	Lower 1/4th β	Lower 1/4th ρ
Recall level						
Recall homework			.22	.05		
Algorithmic homework	.88	.04	.34	.00	.54	.00
Textbook examples					-.41	.00
Comprehension level						
Recall homework	.22	.02	.22	.01		
Algorithmic homework	.21	.03				
Application level						
Analysis level						

For the application and analysis levels of question difficulty, exposure variables did not have a statistically significant effect on student exam success for any of the student percentiles. One possible reason for this is the cognitive demand of these two levels is greater than the cognitive resources of the students, and the exposure variables cannot bridge this gap.

For the lower one-fourth student percentile, exposure variables do not have a significant effect on student exam success above the recall level. In contrast to the upper and middle student percentiles, comprehension questions exert a high degree of cognitive demand on the working memory of these students, and the exposure variables do not have a significant effect on student exam success.

Cognitive skill level. Students scored higher on LOCS questions than on HOCS questions for every student percentile. Table 45 has a summary of the effects of exposure variables for each cognitive skill level.

Table 45

Cognitive Skill Level and Significant Effects of Exposure Variables

Exposure variables	Student percentile: Upper 1/4th		Middle 1/2		Lower 1/4th	
	β	*ρ*	β	*ρ*	β	*ρ*
LOCS						
Recall homework	.21	.00	.23	.00	.15	.00
Algorithmic homework	.14	.04	.24	.00	.20	.03
HOCS						
	-	-	-	-	-	-

Exposure variables have a statistically significant effect on LOCS questions for each student percentile, but they do not have a significant effect on the HOCS level of questions. Both algorithmic and recall homework problems affect exam success, but the contribution each makes to that success varies between the student percentiles.

For the middle student percentile, recall and algorithmic homework problems contribute equally to the exam success of these students. For the lower percentile, algorithmic homework problems contribute more to test achievement. For this set of students, algorithmic homework problems contribute to exam success for the lowest tier of the question type scale and the lowest level of the modified Bloom's taxonomy scale. In addition, these are the only levels affected by

the exposure variables in a significant way. It reasons algorithmic homework problems would contribute most to the exam success of lower one-fourth percentile.

For the upper percentile, recall homework questions contribute more to the exam success of this percentile than algorithmic homework problems do. Students in this percentile are successful solving algorithmic problems. The more students understand the terminology and symbols of a topic the lower the cognitive demand is for questions on that subject.

Number of steps. The number of steps did have a significant negative correlation with the middle and lower student percentiles. There was not a significant correlation between the number of steps and the upper one-fourth percentile. Correlation values are in Table 46.

Table 46

Pearson Correlation between All Student Percentiles and the Number of Steps

Student percentile	Number of steps	Significance	Coefficient of Determination (r^2)
PCA-Upper[a]	$r(325) = -.09$	.09	.01
PCA-Middle	$r(325) = -.18$**	.00	.03
PCA-Lower	$r(325) = -.17$*	.00	.03

[a]Results for the PCA-Upper student percentile are not significant.
* $p < .05$ (two-tailed).
** $p < .01$ (two-tailed).

The number of steps is a measure of cognitive complexity. For the upper percentile, the means for the levels of recall and algorithmic for question type and the levels of recall and comprehension for the modified Bloom's taxonomy were not statistically different. This indicates the difficulty of these lower levels are similar. The effects of the number of steps is not significant for this group. This implies the students in the upper percentile are combining steps

and decreasing the cognitive complexity of problems. Figure 24 has an example of combining steps when calculating the number of moles of substance from the mass.

Calculating moles from mass of a substance:	Number of steps before chunking:	Number of steps after chunking:
Count atoms	1	
Find atomic weights	2	1
Calculate molar mass	3	
Divide by the molar mass	4	2

Figure 24. Chunking of steps to calculate the moles of a substance. When students combine several steps into one, the cognitive demand of the problem decreases.

By chunking several steps into one, the cognitive complexity of the question decreases from four steps to two. This decrease in the number of steps increases student success in answering the question.

The exposure variable that influenced student exam success for each student percentile was recall homework problems. The effects of the exposure variables on the number of steps for each student percentile are in Table 47.

Table 47

Number of Steps and Significant Effects of Exposure Variables

Exposure variables	Student percentile					
	Upper 1/4th		Middle 1/2		Lower 1/4th	
	β	*ρ*	β	*ρ*	β	*ρ*
Number of steps						
Recall homework	.16	.01	.20	.00	.15	.02
Algorithmic homework					.22	.01

For each student percentile, recall homework problems increase student exam success. In addition, algorithmic homework problems increase the exam success of the lower student percentile. For this percentile, algorithmic homework problems increase student success for every difficulty measuring scale. From these results, it can be assumed the lower one-fourth student percentile are challenged solving algorithmic homework problems. For this group, algorithmic homework problems increases exam success.

Topics. For each student percentile acid-base and aqueous equilibria were the most difficult topics compared to the other subjects in Chem 170/175 courses. The topics with statistically significant difference in means are listed in Table 48.

Table 48

Topics for All Student Percentiles with Significant Differences in Means

Student percentile	Scored higher on the topic	than the topic	*p*
Upper 1/4th percentile			
	electron configuration	acid-base equilibria.	.03
Middle ½ percentile			
	matter	aqueous equilibria.	.04
	mole	acid-base equilibria.	.03
		aqueous equilibria.	.01
	electron configuration	acid-base equilibria.	.02
		aqueous equilibria.	.01
	nuclear chemistry	aqueous equilibria.	.02
Lower 1/4th percentile			
	nuclear chemistry	acid-base equilibria.	.02
		aqueous equilibria.	.02

The topics of acid-base and aqueous equilibria were the topics with which these students had the most difficulty. These are topics have a high degree of element interconnectivity which increases the cognitive complexity and increase cognitive demand on the working memory.

In comparison, students score higher on topics that appear in the curriculum for high school chemistry courses and should have been studied in high school (American Association for the Advancement of Science, 1993; Kansas State Department of Education, 2007; Texas Essential Knowledge and Skills). Students should be more familiar with these topics, and the cognitive demand could be lower due to the previous exposure.

Effects of exposure variables on the difficulty of topics were investigated for each student percentile. In Table 49, the effects of the exposure variables on topics for the upper one-fourth student percentile are listed.

Table 49

Effects of Categorized Variables on Topics for the Upper Percentile

Topic	Categorized exposure variable	Statistics β	Statistics *p*
Acid-base equilibria	Lab exposures	.51	.01
Equilibrium	Algorithmic lectures	.67	.00
Aqueous Equilibria	Conceptual homework	.63	.02
Nuclear chemistry	Textbook examples	-.47	.03
Molecular geometry	Algorithmic lecture	-.65	.00

For molecular geometry, algorithmic lecture presentations have a negative effect on the exam success of the upper one-fourth student percentile. To determine a molecule's geometry, there is a systematic process the leads to the geometry. The algorithmic lectures on this topic used the same process. It seems illogical that the algorithmic lectures on this topic would reduce student success; therefore, this might be attributed to the number of times the student attended a chemistry lecture, which as noted before cannot be ascertained.

Textbook exposures on nuclear chemistry had a negative effect on student exam success for this student percentile. The lecture instructor presented material that is applicable to real world problems associated with nuclear chemistry, while the textbook did not. If the student missed the lectures on this topic, he or she would not be able to replace it from the material in the textbook.

Algorithmic lectures had a positive effect on exam questions about equilibrium. Equilibrium and the formulas and equations would be new topics for most of these students. Demonstrating how to use these algorithms through worked examples would increase student success.

Lab exposures had a positive impact on questions about acid-base equilibria. In the laboratory course associated with the Chem 175 course, four out of ten labs were associated in with acid-base reactions.

Conceptual homework problems increased student success on questions about aqueous equilibria. Aqueous equilibria is topic with high element connectivity. Conceptual homework questions ask about the different subtopics and their relationships. These homework problems should decrease the cognitive demand of this topic.

Table 50 has the effects of various exposure variables on topics for the middle one-half student percentile.

Table 50

Effects of Categorized Variables on Topics for the PCA-Middle Student Percentile

		Statistics	
Topic	Categorized exposure variable	β	*p*
Electron configuration	Conceptual lecture	-.55	.01
Gases	Algorithmic lecture	.59	.04
Equilibrium	Algorithmic lecture	.48	.01
Acid-base equilibria	Algorithmic lecture	-1.79	.00
	Conceptual homework	-.97	.00
	Textbook example	.78	.01
Electrochemistry	Textbook example	.62	.02

For the middle one-half student percentile, both algorithmic and conceptual homework problems decreased student success on exam questions about acid-base equilibria. The topic is so confusing and difficult for some students the exposures that should help their understanding only

adds to the confusion. Textbook exposures increase exam success on acid-base equilibria. The worked examples provided by the textbook and the visual representations could help students reduce the cognitive load of this topic.

For the topics of equilibrium and gases, algorithmic lectures increase the success on these topics. Both topics require a high degree of algorithmic problem solving. The modeling of problem solving and the ability to ask questions in the lecture plus the diagrams produced by the instructor all help a student's understanding.

For the lower one-fourth student percentile, the exposure variables had an effect on two topics. The topics of thermochemistry and electron configuration were affected by the exposure variables. The effects of these exposure variables are in Table 51.

Table 51

Effects of Categorized Variables on Topics for the PCA-Lower Student Percentile

		Statistics	
Topic	Categorized exposure variable	β	*p*
Thermochemistry	Recall homework	.65	.03
Electron configuration	Algorithmic lecture	.44	.01
	Conceptual lecture	-.78	.01

The discussion on the effects of recall homework problems on the topic of thermochemistry is in the results section for all student group. The discussion on electron configuration is also in that section. For this percentile, conceptual lectures decreases exam success on the topic of electron configuration. Conceptual lectures could decrease student success for this percentile if the students are not able to construct abstract models about this topic.

Recall and Algorithmic Homework Variables

The two variables that increased student success were recall and algorithmic homework problems. Recall homework problems increase student success by increasing the symbolic knowledge of the topics and is effective for both conceptual and algorithmic learning.

Algorithmic homework problems provide opportunities to practice problem solving, and graded homework assignments provide feedback on the student's progress in understanding the topic material. However, students must know the symbolic representations and the terminology to solve these problems.

The following questions are similar to those used for homework assignments and should demonstrate how these relationships work together to increase student exam success.

Question 1: Calculate the <u>standard enthalpy of formation</u>, using the following thermochemical information:

$$aA_{(s)} + bB_{(\ell)} \rightarrow cC_{(s)} + dD_{(g)} \qquad \underline{\Delta H^{\circ}} = -125.62$$

Question 2: Calculate $\underline{\Delta H_f^{\circ}}$ from the following thermochemical information:

$$aA_{(s)} + bB_{(\ell)} \rightarrow cC_{(s)} + dD_{(g)} \qquad \underline{\Delta H^{\circ}} = -225.02$$

For question one, the definition of the standard enthalpy of formation plus the symbolic representation and definition of standard enthalpy (ΔH°) adds two to three steps to the number of steps for a question that already requires three steps to solve. Question 2 is the same question as number one, but uses the symbolic representation of the standard enthalpy of formation for the variable. In both cases, the knowledge of the symbolic representations and definitions reduce the cognitive demand of questions.

Some information about a topic has to be memorized to be effective. An example is the names and ionic charges of polyatomic ions (Gulacar et al., 2013). Students must know the name

of the polyatomic ion, the formula of the ion, and the ionic charge to be successful in naming compounds and writing balanced chemical equations.

For novice learners, the knowledge of the terminology and the symbolic representations has a large effect on the success of these students in chemistry.

Effects of Categorized Variable on Each Student Percentile

The categorized variables count the method of exposures. These are recall, algorithmic, and conceptual homework problems, recall, algorithmic, and conceptual lecture exposures, and the textbook and laboratory exposures. Table 52 presents the effects of the categorized variables on each student percentile.

Table 52

Categorized Exposure Variables on Student Success for Each Percentile

Variable	β	*p*
Upper 1/4th percentile		
Recall homework problems	.16	.01
Algorithmic homework problems	.15	.01
Middle ½ percentile		
Recall homework problems	.21	.00
Algorithmic homework problems	.14	.02
Lower 1/4th percentile		
Recall homework problems	.13	.02
Algorithmic homework problems	.22	.02
Textbook examples	-.15	.03

The difficulty of exam questions are a product of the cognitive demand and the interactivity of the elements of a topic. Each difficulty measure scale determines this difficulty

by different cognitive criteria. The categorized exposure variables of recall and algorithmic homework problems increase exam success for all student percentiles. The amount each one contributes to this success varies for each student percentile. The effects of the exposure variables on each difficulty scale will be used to explain these differences. Furthermore, only the levels of question difficulty that are effected by exposure variables will be discussed.

Middle one-half student percentile. For the overall success of this percentile, recall homework questions contribute more to exam achievement than algorithmic homework problems. The different difficulty scales and the exposure variables that have an effect on student exam success are in Table 53.

Table 53

Effects of Exposure Variables on Question Difficulty Scales for Middle One-half Percentile

Cognitive skill	β	*p*	Question type	β	*p*	Bloom's taxonomy	β	*p*
LOCS			**Recall**			**Recall**		
Recall hwk	.23	.00	Alg.hwk	.24	.03	Recall hwk	.22	.05
Alg. hwk	.24	.00				Alg. hwk	.34	.00
HOCS			**Algorithmic**			**Comprehension**		
			Recall hwk	.20	.04	Recall hwk	.22	.01
			Conceptual			**Application**		
			Recall hwk	.20	.02			
						Analysis		

For the middle student percentile, the exposure variables affect student exam success at each level of question difficulty. This indicates the students in the middle percentile have a wider range of cognitive abilities than either the upper or the lower percentiles.

Recall homework problems contribute to student exam success for each level of question difficulty except for the recall level of question type. For students in the middle one-half percentile, the knowledge of the symbolic representations and definitions of the topic increases student exam success. Recall homework problems contribute more to the exam success for this student percentile.

Upper one-fourth student percentile. For this student percentile, recall and algorithmic homework problems contribute almost equally to exam success. The effects of exposure variables on each level of the question difficulty scales are in Table 54.

Table 54

Effects of Exposure Variables on Question Difficulty Scales for Upper One-fourth Percentile

Cognitive skill	β	*p*	Question type	β	*p*	Bloom's taxonomy	β	*p*
LOCS			**Recall**			**Recall**		
Recall hwk	.21	.00				Alg. hwk	.88	.04
Alg. hwk	.14	.04						
HOCS			**Algorithmic**			**Comprehension**		
			Alg. hwk	.19	.04	Recall hwk	.22	.02
						Alg. hwk	.21	.03
			Conceptual			**Application**		
			Recall hwk	.23	.01			
						Analysis		

For this percentile, the effects of exposure variables on the lower levels of the difficulty scales does not influence exam success as much as for the other student percentiles. Exposure variables do not affect exam success for the recall level of question type, and the relationship between the number of steps and exam success is not significant. For the more cognitively

demanding levels of the difficulty scales, both recall and algorithmic homework problems contribute to exam success for this percentile.

Lower one-fourth student percentile. Algorithmic homework problems affect exam success than recall homework problems. The relationships between exposure variables and the different levels of the difficulty measurement scales are in Table 55.

Table 55

Effects of Exposure Variables on Question Difficulty Scales for Lower One-fourth Percentile

Cognitive skill	β	*p*	Question type	β	*p*	Bloom's taxonomy	β	*p*
LOCS			**Recall**			**Recall**		
Recall hwk	.15	.00	Alg.hwk	.25	.03	Alg. hwk	.54	.00
Alg. hwk	.20	.03				Textbook	-.41	.00
HOCS			**Algorithmic**			**Comprehension**		
			Conceptual			**Application**		
						Analysis		

For the lower one-fourth student percentile, exposure variables affect the least cognitively demanding levels of each question difficulty scale. Algorithmic homework problems increase exam success for each question scale. For these students, there is a statistically significant difference in the means between the recall and algorithmic levels of question type and between the means of recall and comprehension levels of Bloom's taxonomy. Algorithmic homework problems contribute more to the success of these students.

The effects of question difficulty impacts each student percentile in a different manner. For the upper student percentile, their level of achievement benefits slightly from higher order cognitive questions. For the middle percentile, their achievement benefits from a broad range of question types. For the lower percentile, the least cognitively demanding levels of exam questions affect their success.

Limitations

A major limitation of this research project was the variables of lecture and textbook exposures. The manner in which these exposure variables were counted was flawed. Students completed the homework assignments and attended lab sessions because these grades were used to determine a portion of their course grade. The number of homework exposures would be the number of exposures most students received. For the laboratory exposure variable, the number of lab exposures was the number of exposures. For both of these exposure variables, using the total number of possible exposures was similar to the actual number of exposures.

This was not the case for the lecture and textbook exposures. Lecture attendance was not consistent and the actual use of the textbook could not be counted. For these two exposure variables, the maximum possible number of exposures was not the same as the actual number experienced by the student. Each time the variable for lecture or for textbook was used, all possible exposures were counted. This decreases the effects of these variables.

Future Research

The object of the new pedagogy is to increase a pupil's developmental level from concrete operational to formal operational through increased exposure to the higher-order thinking skills. The intervention is called cognitive acceleration through science education (CASE).

Challenging students at the concrete operational stage with difficult and complex problems increases the transition from the concrete to formal operational stage (Endler & Bond, 2007). Future research would be to determine the effects of interventions instituted in the lecture portion of an introductory chemistry course on a student's Piagetian cognitive stage.

Interventions would be both conceptual and algorithmic exposures in the lecture portion of the class using personal response systems and in-class assignments. During lecture presentations, students would be asked conceptual questions about the topic being studied. The students would answer using a personal response system. The instructor would know immediately if students understand the concepts of the topic, and students would receive immediate feedback on their understanding. This would allow an instructor to reteach any concept the majority of students responded to incorrectly.

In-class assignments are short, group exercises used at the end of a class session. These would be algorithmic questions covering topics presented in recent lecture presentations. They would start with simple questions leading to more challenging ones. The students would work in small groups to complete the assignment before the class is over, and then, the assignments would be graded and returned to the students the next lecture session.

Exam questions would be similar to those presented in the lecture. The effects of each type of exposure and their interactions would be measured by exam success.

Students would be challenged, but would have support systems in place to help them be successful. The teach-reteach would support their conceptual understanding, and the small group and peer teaching would support their algorithmic understanding.

Another project would be to design a lecture exposure variable that measures the actual number of lecture exposures by tracking attendance. Daily quizzes or attendance points would

increase the accuracy of the lecture exposure variable and produce a better measuring tool. For the textbook exposure, additional material could be posted on Blackboard or other on-line programs to keep track of which students accessed the material.

Conclusions

Each measure of question difficulty produced results similar to those found in the chemical education research literature. For each scale, students scored higher on the lower tiers of the scale, and student exam success decreased as the levels of the difficulty scales increased.

The difficulty of topics depends on the prior knowledge of the student and the degree of element interactivity of the topic. Students scored higher on topics listed in the curriculum for high school chemistry courses. The prior exposure to these topics decreased the cognitive demand of the topic. Students scored lower on the topics of acid-base equilibria and additional aqueous equilibria when compared to other topics taught in the course because of high interconnectivity of the subtopics.

For novice learners, the knowledge of the terminology and the symbolic representations of a topic increases student success for all student percentiles. The student reduces the load on the working memory by understanding the symbolic language of a topic. The student does not have to use his or her working memory to search for the meaning of the terminology and symbols.

The effects of the categorized exposure variables differed by student percentile. For the upper one-fourth, exposure variables are not a major factor in their success on the lower levels of cognitive demand. For the middle one-half percentile, the exposure variables can affect success on each level of cognitive demand because the distribution of cognitive development is broader for this level. For the lower-one fourth student percentile, for these students exposure variables increase student success only on the lower tiers of question difficulty.

There were five research questions that needed answers to determine if the number of exposures affects the success of students. The questions and their answers follow.

- Does the number of times a student is exposed to a chemical concept affect the success rate of that student on exam questions covering the same topic?

  The results indicate there is not a statistically significant relationship between the number of exposures and student success on exam questions.

- Does the location where the student is shown a concept affect the success rate for students on exam questions testing the same topic?

  Homework problems increased student exam success for all students in this study.

- Does the method in which a concept is presented affect the success rate of students on exam questions?

  Recall and algorithmic homework problems promoted the most success on student exam scores.

- Do topics have different levels of difficulty, and if so, does the number of exposures decrease the difficulty of the more challenging topics?

  Topics have different difficulty levels. The most difficult topics were those with a high degree of connectivity between the elements of the topic. The type of exposures with a significant effect on student success depends on the topic.

- Do the learning environment and the number and type of exposures affect students differently at different cognitive levels?

  Students are affected differently by topics, and exposure variables. For the lower one-fourth student percentile, exposure variables increased student success for the least cognitively demanding questions. Above this level, there were no

variables that increased the exam success of these students. For the middle-one half student percentile and upper-one-fourth percentile, exposure variables increased exam success for the lower tiers of each measure of question difficulty.

References

Akkus, H., Kadayifçi, H., Atasoy, B., & Geban, Ö. (2003). Effectiveness of instruction based on the constructivist approach on understanding chemical equilibrium concepts. *Research in Science & Technological Education, 21*(2), 209-227. doi: http://dx.doi.org/10.1080/0263514032000127248

American Association for the Advancement of Science. (1993, 24 Oct 2014). *Benchmarks on-line*. *Project 2061 Benchmarks On-line.* Retrieved 14 March 2016, Retrieved from http://www.project2061.org/publications/bsl/online/index.php?chapter=4

Ayres, P.L. (1993). Why goal-free problems can facilitate learning. *Contemporary Educational Psychology, 18*(3), 376-381. doi: http://dx.doi.org/10.1006/ceps.1993.1027

Bannert, M. (2002). Managing cognitive load—recent trends in cognitive load theory. *Learning and Instruction, 12*(1), 139-146. doi: http://dx.doi.org/10.1016/S0959-4752(01)00021-4

Bauer, C.F. (2008). Attitude toward chemistry: A semantic differential instrument for assessing curriculum impacts. *Journal of Chemical Education, 85*(10), 1440-1445. doi: 10.1021/ed085p1440

Benjah-bmm27 (Own work). (2012). Sodium-chloride-3d-ionic.Png: [Public domain] via Wikimedia Commons.

Bird, L. (2010). Logical reasoning ability and student performance in general chemistry. *Journal of Chemical Education, 87*(5), 541-546. doi: 10.1021/ed8001754

Bodner, G., Klobuchar, M., & Geelan, D. (2001). The many forms of constructivism. *Journal of Chemical Education, 78*(8), 1107. doi: 10.1021/ed078p1107.4

Bodner, G.M. (1986). Constructivism: A theory of knowledge. *Journal of Chemical Education, 63*(10), 873. doi: 10.1021/ed063p873

Bretz, S.L. (2005). All students are *not* created equal: Learning styles in the chemistry classroom. In N. J. Pienta, M. M. Cooper & T. J. Greenbowe (Eds.), *Chemists' guide to effective teaching* (pp. 28-40). Upper Saddle River, NJ: Pearson Education, Inc.

Brown, T.L., LeMay Jr., H.E., Bursten, B.E., Murphy, C.J., & Woodward, P.M. (2012). *Chemistry: The central science* (12th ed.). United States of America: Prentice Hall.

Bunce, D. (2001). Does piaget still have anything to say to chemists? *Journal of Chemical Education, 78*(8). doi: 10.1021/ed078p1107.2

Bunce, D.M. (2005). Solving word problems in chemistry: Why do students have difficulty and what can be done to help. In N. J. Pienta, M. M. Cooper & T. J. Greenbowe (Eds.), *Chemists' guide to effective teaching*. Upper Saddle River, New Jersey: Pearson Education, Inc.

Butler, M.B., & Zerr, R.J. (2005). The use of online homework systems to enhance out-of-class student engagement. *The International Journal for Technology in Mathematics Education, 12*(2).

Buzzetto-More, N., & Ukoha, O. (2009). The efficacy of a web-based instruction and remediation program on student learning. Retrieved 6 Feb. 2013 http://go.galegroup.com/ps/i.do?id=GALE%7CA229896128&v=2.1&u=ksstate_ukans&it=r&p=AONE&sw=w

Carver, L. Evolution of the dog. In T. W. s. w. p. f. E. b. W. Interactive (Ed.), *Evolution Library*: PublIc Broadcasting Station.

Cheng, M., & Gilbert, J. (2009). Towards a better utilization of diagrams in research into the use of representative levels in chemical education. In J. Gilbert & D. Treagust (Eds.), *Multiple representations in chemical education* (Vol. 4, pp. 55-74): Springer Science+Business Media B.V.

Choppin, B.H., & Postlethwaite, T.N. (2014). *Evaluation in education, volume 1*: Elsevier.

Contributors, W., 12/22/2015). *Problem-based learning*. Retrieved 01/11/2016, Retrieved from https://en.wikipedia.org/w/index.php?title=Problem-based_learning&oldid=696351620

Costu, B. (2010). Algorithmic, conceptual and graphical chemistry problems: A revisited study. *Asian Journal of Chemistry, 22*(8), 6013-6025.

Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review, 26*(2), 197-223. doi: 10.1007/s10648-013-9246-y

Cracolice, M.S. (2005). How students learn: Knowledge construction in college chemistry courses. In N. J. Pienta, M. M. Cooper & T. J. Greenbowe (Eds.), *Chemists' guide to effective teaching*. Upper Saddle River, NJ: Prentice Hall.

Cracolice, M.S., & Deming, J.C. (2001). Peer-led team learning. *The Science Teacher, 68*(1), 20-24.

Crippen, K.J., & Brooks, D.W. (2009). Applying cognitive theory to chemistry instruction: The case for worked examples. *Chemistry Education Research and Practice, 10*(1), 35-41. doi: 10.1039/B901458F

Crippen, K.J., & Earl, B.L. (2004). Considering the efficacy of web-based worked examples in introductory chemistry. *Journal of Computers in Mathematics and Science Teaching, 23*(2), 151-167.

Dávila, K., & Talanquer, V. (2010). Classifying end-of-chapter questions and problems for selected general chemistry textbooks used in the united states. *Journal of Chemical Education, 87*(1), 97-101. doi: 10.1021/ed8000232

Deming, J.C., O'Donnell, J.R., & Malone, C.J. (2012). Scientific literacy: Resurrecting the phoenix with thinking skills. *Science Educator, 21*(2), 10-17.

Di Leonardi, B.C. (2007). Tips for facilitating learning: The lecture deserves some respect. *Journal of Continuing Education in Nursing, 38*(4), 154-175.

Elkind, D. (1988). Acceleration. *Young Children, 43*(4).

Endler, L.C., & Bond, T.G. (2007). Changing science outcomes: Cognitive acceleration in a us setting. *Research in Science Education, 2008*(38), 149-166. doi: 10.1007/s11165-007-9042-0

Fleming, N. (2001). *Vark-a guide to learning styles*. Retrieved October, 3, 2014, 2014, Retrieved from http://vark-learn.com/english/index.asp

Forehand, M. (2005). *Bloom's taxonomy: Original and revised*. *Emerging perspectives on learning, teaching, and technology*. Retrieved 04/19/2015, Retrieved from http://projects.coe.uga.edu/epltt/index.php?title=Bloom%272_Taxonomy

Fynewever, H. (2008). A comparison of the effectiveness of web-based and paper-based homework for general chemistry. *Chemical Educator, 2008*(13), 264-269.

Gabel, D.L. (1993). Use of the particle nature of matter in developing conceptual understanding. *Journal of Chemical Education, 70*(3), 193. doi: 10.1021/ed070p193

Garnett, P.J., Garnett, P.J., & Hackling, M.W. (1995). Students' alternative conceptions in chemistry: A review of research and implications for teaching and learning. *Studies in Science Education, 25*(1), 69-95.

Gasiewski, J.A., Eagan, M.K., Garcia, G.A., Hurtado, S., & Change, M.J. (2012). From gatekeeping to engagement: A multicontextual, mixed method study of student academic engagement in introductory stem courses. *Research in Higher Education, 53*(2), 229-261. doi: 10.1007/s1162-011-9247-y

Gould, M., & Howson, A. (2011). Piaget's stages of cognitive development. In T. E. o. S. Press (Ed.), *Sociology reference guide: The process of socialization* (1st ed., pp. 29-38): Salem Press.

Gulacar, O., Overton, T.L., Bowman, C.R., & Fynewever, H. (2013). A novel code system for revealing sources of students' difficulties with stoichiometry. *Chemistry Education Research and Practice, 14*(4), 507-515. doi: 10.1039/C3RP00029J

Haley, S., & Good, R.G. (1976). Concrete and formal operational thought: Implications for introductory college biology. *The American Biology Teacher, 38*(7), 407-430. doi: 10.2307/4445652

Hartman, J.A.R., & Lin, S. (2011). Analysis of student performance on multiple-choice questions in general chemistry. *Journal of Chemical Education, 88*(9), 1223-1230. doi: 10.1021/ed100133v

Hockings, S.C., DeAngelis, K.J., & Frey, R.F. (2008). Peer-led team learning in general chemistry: Implementation and evaluation. *Journal of Chemical Education, 85*(7).

Holme, T., Knaus, K., Murphy, K., & Blecking, A. (2011). A valid and reliable instrument for cognitive complexity rating assignment of chemistry exam items. *Journal of Chemical Education, 88*(5), 554-560. doi: 10.1021/ed900070y

Holme, T., & Murphy, K. (2011). Assessing conceptual and algorithmic knowledge in general chemistry with acs exams. *Journal of Chemical Education, 88*(9), 1217-1222. doi: 10.1021/ed100106k

IBM. (2011). Ibm spss statistics for windows (Version 20.0). Armonk, NY: IBM Corp.

Igaz, C., & Proksa, M. (2012). Conceptual questions and lack of formal reasoning: Are the mutually exclusive? *Journal of Chemical Education, 2012*(89), 1243-1248. doi: dx.doi.org/10.1021/ed100895c |

Intel Teach Program. (2012). *Instructional strategies: Modeling*. *Designing Effective Projects.* Retrieved 1/02/16, Retrieved from http://www.intel.com/content/dam/www/program/education/us/en/documents/project-design/strategies/instructionalstrategies-modeling.pdf

Johnsone, A.H., & El-Banna, H. (1986). Capacities, demands and processes - a predictive model for science education. *EDUCATION IN CHEMISTRY*.

Johnstone, A.H. (1982). Macro- and micro-chemistry. *School Science Review, 64*(227), 377-379.

Johnstone, A.H. (1983). Chemical education research: Facts, findings, and consequences. *Journal of Chemical Education, 60*(11), 968-971. doi: 10.1021/ed060p968

Jones, K.O., Harland, J., Reid, J.M.V., & Bartlett, R. (2009, 18-21 Oct. 2009). *Relationship between examination questions and bloom's taxonomy.* Paper presented at the Frontiers in Education Conference, 2009.

Kansas State Department of Education. (2007). *Kansas science education standards*. Retrieved from http://community.ksde.org/LinkClick.aspx?fileticket=49KllvxS4nc%3D&tabid=4697&mid=11426.

Keselman, A. (2003). Supporting inquiry learning by promoting normative understanding of multivariable causality. *Journal of Research in Science Teaching, 40*(9), 898-921. doi: 10.1002/tea.10115

Kirschner, P.A., Sweller, J., & Clark, R.E. (2010). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-base, experiential, and inquiry--based teaching. *Educational Psychologist, 41*(2), 75-86. doi: 10.1201/s15326985ep4102 1

Kuhn, D. (1979). The significance of piaget's formal operations stage in education. *The Journal of Education, 161*(1), 34-50.

Lee, H.S., & Anderson, J.R. (2013). Student learning: What has instruction got to do with it? *Annual Review of Psychology, 64*(1), 445-469. doi: doi:10.1146/annurev-psych-113011-143833

Leech, N.L., Barrett, K.C., & Morgan, G.A. (2015). *Ibm spss for intermediate statistics use and interpretation* (5th ed.). New York, NY: Routledge.

Lewis, A., & Smith, D. (1993). Defining higher order thinking. *Theory into Practice, 32*(3), 131-137. doi: 10.2307/1476693

Lord, T., & Baviskar, S. (2007). Moving students from information recitation to information understanding: Exploiting bloom's taxonomy in creating science questions. *Journal of College Science Teaching, 36*(5), 40-44.

Madden, S.P., Jones, L.L., & Rahm, J. (2010). The role of multiple representations in the understanding of ideal gas problems. *Chemistry Education Research and Practice, 12*(3), 283-293.

Mayer, R.E. (2005). *The cambridge handbook of multimedia learning*: Cambridge University Press.

Mayer, R.E., Bove, W., Bryman, A., Mars, R., & Tapangco, L. (1996). When less is more: Meaningful learning from visual and verbal summaries of science textbook lessons. *Journal of Educational Psychology, 88*(1), 64-73. doi: http://dx.doi.org/10.1037/0022-0663.88.1.64

McKinnon, J.W., & Renner, J.W. (1971). Are colleges concerned with intellectual development. *American Journal of Physics, 39*, 1047-1052. doi: 10.1119/1.1986367

McTigue, E.M., & Flowers, A.C. (2011). Science visual literacy: Learners' perceptions and knowledge of diagrams. *The Reading Teacher, 64*(8), 578-589. doi: 10.2307/41203457

Moore, J.C., & Rubbo, L.J. (2012). Scientific reasoning abilities of nonscience majors in physics-based courses. *Physical Review Special Topics - Physics Education Research, 8*(1), 010106.

Morgan, G.A., Leech, N.L., Gloeckner, G.W., & Barrett, K.C. (2013). *Ibm spss for introductory statistics use and interpretation* (5th ed.). New York, New York: Routledge and imprint of the Taylor & Francis Group.

Mostyn, G.R. (2012). Cognitive load theory: What it is, why it's important for accounting instruction and research. *Issues in Accounting Education, 27*(1), 227-245.

Nakhleh, M.B. (1992). Why some students don't learn chemistry: Chemical misconceptions. *Journal of Chemical Education, 69*(3), 191. doi: 10.1021/ed069p191

Nakhleh, M.B. (1993a). Are our students conceptual thinkers or algorithmic problem solvers? *Journal of Chemical Education, 70*(1).

Nakhleh, M.B. (1993b). Concept learning versus problem solving: There is a difference *Journal of Chemical Education* (Vol. 70, pp. 190-192).

Nakhleh, M.B. (1994). Chemical education research in the laboratory environment how can research uncover what students are learning? *Journal of Chemical Education, 71*(3).

Nevid, J.S., & McClelland, N. (2013). *Using action verbs as learning outcomes: Applying bloom's taxonomy in measuring instructional objectives in introductory psychology* (Vol. 1).

Niaz, M. (2001). Response to contradictions: Conflict resolution strategies used by students in solving problems of chemical equilibrium. *Journal of Science Education & Technology, 10*(2).

Nurrenbern, S.C., & Pickering, M. (1987). Concept learning versus problem solving: Is there a difference? *Journal of Chemical Education, 64*(6), 508. doi: 10.1021/ed064p508

Ohiostandard at en.wikipedia. (2016, February 11). Galvanic cell: Wikipedia, The Free Encyclopedia.

Orgill, M., & Crippen, K. Teaching with external representations: The case of a common energy-level diagram in chemistry. *Journal of College Science Teaching, 40*(1), 78-84.

Papaphotis, G., & Tsaparlis, G. (2008). Conceptual versus algorithmic learning in high school chemistry: The case of basic quantum chemical concepts. Part 2. Students' common errors, misconceptions and difficulties in understanding. *Chemistry Education Research and Practice, 9*(4).

Pecher, D., & Boot, I. (2011). Numbers in space: Differences between concrete and abstract situations. *Frontiers in Psychology, 2*(Article 121), 1-11.

Phelps, A. (1996). Teaching to enhance problem solving. *Journal of Chemical Education, 73*(4), 301-304.

Price, I. (2000). *Chapter 5: Analysing the data part ii: Inferential statistics*. Retrieved 03/08/2016, Retrieved from http://webstat.une.edu.au/unit_materials/c5_inferential_statistics/independ_observe.html

Richards-Babb, M., Drelick, J., Henry, Z., & Robertson-Honecker, J. (2011). Online homework, help or hindrance? What students think and how they perform. *Journal of College Science Teaching, 40*(4), 81-93.

Rittle-Johnson, B., Star, J.R., & Durkin, K. (2009). The importance of prior knowledge when comparing examples: Influences on conceptual and procedural knowledge of equation solving. *Journal of Educational Psychology, 101*(4), 836-852.

Rovai, A.P., Baker, J.D., & Ponton, M.K. (2013). *Social science research design and statistics: A practitioner's guide to research methods and spss analysis* (1st ed.). Chesapeake, VA: Watertree Press LLC.

Rovai, A.P., Baker, J.D., & Ponton, M.K. (2014). *Social science research design and statistics: A practitioner's guide to research methods and ibm spss analysis* (second ed.). Chesapeake, Virginia: Watertree Press LLC.

Royan, J.o.w. (2006). Four salt shakers in a table with spilled sal. In S.-. 5394.jpg (Ed.): WikiMedia.

Sandi-Urena, S., Cooper, M.M., Gatlin, T.A., & Bhattacharyya, G. Students' experience in a general chemistry cooperative problem based laboratory. *Chemistry Education Research and Practice, 12*(4).

Santa Barbara City College Student Learning Outcomes Project. (2015, 2015). *Bloom's taxonomy learning domains*. 2015, Retrieved from http://slo.sbcc.edu/wp-content/uploads/bloomstaxonomyoflearningdomains.doc

Seery, M.K., & Donnelly, R. (2012). The implementation of pre-lecture resources to reduce in-class cognitive load: A case study for higher education chemistry. *British Journal of Educational Technology, 43*(4), 667-677. doi: 10.1111/j.1467-8535.2011.01237.x

Shayer, M. (1987). Neo-piagetian theories and educational practice. *International Journal of Psychology, 22*(5/6), 751-772.

Sirhan, G. (2007). Learning difficulties in chemistry: An overview. *Turkish Science Education, 4*(2), 2-20.

Smith, D.K. (2006). Use of the mid-lecture break in chemistry teaching: A survey and some suggestions. *Journal of Chemical Education, 83*(11), 1621-1624. doi: 10.1021/ed083p1621

Smith, K.C., Nakhleh, M.B., & Bretz, S.L. (2010). An expanded framework for analyzing general chemistry exams. *Chemistry Education Research and Practice, 11*(3), 147-153. doi: doi:10.1039/C005463C

Solaz-Portoles, J.J., & Lopez, V.S. (2008). Types of knowledge and their relations to problem solving in science: Directions for practice. *Educational Sciences Journal, May/August 2008*(6).

StatsToDo.). *Intraclass correlation coefficient explained*. Retrieved 12/04/15, 2015, Retrieved from https://www.statstodo.com/index.php

Sweller, J., & Levine, M. (1982). Effects of goal specificity on means-ends analysis and learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*(5), 463-474. doi: http://dx.doi.org/10.1037/0278-7393.8.5.463

Sweller, J., van Merriënboer, J.J.G., & Paas, F.G.W.C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251-296.

Tabor, K.S. (2013). Revisiting the chemistry triplet: Drawing upon the nature of chemical knowledge and the psychology of learning to inform chemistry education. *Chemistry Education Research and Practice, 2013*(14), 156-168. doi: 10.1039/c3rp00012e

Texas Essential Knowledge and Skills. *Texas essential knowledge and skills*. Austin, Texas: Retrieved from http://tea.texas.gov/About_TEA/Welcome_and_Overview/.

Thompson, T. (2011). *An analysis of higher-order thinking on algebra i end-of-course tests*. Retrieved October 13, 2014, 2014, Retrieved from http://www.cimt.plymouth.ac.uk/journal/default.htm

Tiemeier, A.M., Stacy, Z.A., & Burke, J.M. (2011). Using multiple choice questions written at various bloom's taxonomy levels to evaluate student performance across a therapeutics sequence. *Innov. Pharm, 2*(41), 1-11.

Tsui, L. (2002). Fostering critical thinking through effective pedagogy: Evidence from four institutional case studies. *The Journal of Higher Education, 73*(6), 740-763.

Wedderburn, K.M., Bililign, S., Levy, M., & Gdanitz, R. (2006). Geometries and stabilities of 3d-transition metal-cation benzene complexes, $m^{+}bz_{n}$(m=sc-cu, n=1,2). *Chemical Physics, 326*(2-3), 600-604.

Zerr, R. (2007). A quantitative and qualitative analysis of the effectiveness of online homework in first-semester calculus. *The Journal of Computers in Mathematics and Sciencing Teaching, 26*(1), 55-73.

Zoller, U. (1993). Are lecture and learning compatible? Maybe for locs: Unlikely for hocs. *Journal of Chemical Education, 70*(3), 195-197. doi: 10.1021/ed070p195

Zoller, U., Lubezky, A., Nakhleh, M.B., Tessier, B., & Dori, Y.J. (1995). Success on algorithmic and locs vs. Conceptual chemistry exam questions. *Journal of Chemical Education, 72*(11), 987. doi: 10.1021/ed072p987

Zoller, U., & Tsaparlis, G. (1997). Higher and lower-order cognitive skills: The case of chemistry. *Research in Science Education, 27*(1), 117-130. doi: 10.1007/bf02463036

Appendix A

The list of topics and subtopics is below. The listings under the subtopics are the concepts studied in that section.

Topic 1: Matter and Its Measure

a. Matter
   chemical and physical properties

   changes in matter

b. Units of measurement
   SI units

   density, temperature and derived SI units

c. Dimensional analysis
   conversions with one or more factors

Topic 2: Atoms, Molecules and Ions

a. Atomic Theory
   discovery of subatomic particles

   quantum mechanics

b. Atomic weights and the Periodic Table
   atomic weights

   arrangement of the Periodic table

c. Chemical compounds
   molecular compounds

   ionic compounds

   naming compounds

Topic 3: Chemical Equations and Stoichiometric Calculations

a. Chemical equations
   balancing equations

b. The Mole

molecular and formula weights

empirical formulas

c. Quantitative information from chemical equations
   molar mass

   interconverting between moles ↔ mass ↔ number of particles

   limiting reagents

Topic 4: Thermochemistry

a. Energy
   kinetic and potential energy

   energy units

b. First Law of Thermodynamics
   endothermic and exothermic reactions

   changes in energy, ΔE

c. Enthalpy
   changes in enthalpy, ΔH

   enthalpies of reactions

   Hess's Law

   enthalpies of formation

Topic 5: Atoms and Their Structure

a. Wave-particle duality of matter
   wave nature of energy

   photons

   line spectra

   Bohr's Model

   Uncertainty Principle

b. Quantum mechanics
   quantum numbers

   s, p, d, and f orbitals

   orbital energies

   Pauli Exclusion Principle

c. Electron configuration
   Hund's Rule

   condensed electron configurations

   electron configurations and the Periodic Table

Topic 6: Periodic Properties

a. Effective nuclear charge

b. Size of atoms and ions

c. Ionization energy and electron affinity

Topic 7: Chemical Bonding

a. Bonding and Lewis symbols
   Lewis symbols

   Octet rule

   electronegativity and bond polarity

   dipole moments

b. Drawing Lewis Structures
   formal charges

c. Exceptions to the octet rule
   resonance structures

   bond strength

Topic 8: Geometry of Molecules

a. Valence-Shell Electron-Pair Repulsion Model, VSEPR
   molecular shapes

   bond angles

   molecular polarity

b. Valence Bond Theory and covalent bonds
   hybrid orbitals

   sigma (σ) and pi (π) bonding

c. Molecular Orbitals
   period 1 and 2 diatomic molecules

Topic 9: Gases

a. The Gas Laws
   - pressure
   - Boyle's Law
   - Charles's Law
   - Avogadro's Law

b. Ideal Gas Equation
   - gases in chemical reactions
   - molar mass
   - partial pressure
   - mole fraction

c. Kinetic-Molecular Theory of Gases
   - molecular speeds
   - effusion and diffusion
   - real gases

Topic 10: Intermolecular Forces and the States of Matter

a. Intermolecular forces
   - dispersion forces
   - dipole-dipole forces
   - hydrogen bonding
   - strength of intermolecular forces

b. Phase changes
   - surface tension
   - vapor pressure
   - volatility
   - heating curves

Topic 11: Chemical Kinetics

a. Reaction rates
   - factors that affect rates

change of rate with time

reaction rates and stoichiometry

reaction orders

b. Change of concentration
first, second, and zero order reactions

half-life

activation energy

Arrhenius Equation

c. Reaction mechanisms
elementary reactions and rate laws

rate-determing step for multistep reactions

Topic 12: Chemical Equilibrium

a. Concept of equilibrium
equilibrium constants

stoichiometry and equilibrium constants

b. Equilibrium constants
calculating equilibrium constants

heterogeneous equilibrium

reaction quotient

finding concentrations at equilibrium

c. Le Châtelier's Principle
volume, pressure and temperature changes

Topic 13: Acid/Base Equilibria

a. Acids and bases
Brønsted – Lowery acids and bases

conjugate acid-base pairs

pH and pOH scales

b. Acid/base strength
strong acids and bases

weak acids and bases

acid-dissociation constant (K_a) and base-dissociation constant (K_b)

c. Acid/base structure and behavior
   structure and acid strength

Topic 14: Aqueous Equilibria

a. Common-ion effect
   buffered solutions and buffer capacity

   pH of buffer solutions

   acid/base tritations

b. Solubility equilibria
   Solubility-product constant (K_{sp})

   factors that affect solubility

   complex ions

c. Quantitative analysis
   precipitation of ions

Topic 15: Free Energy and Equilibria

a. Spontaneous processes
   reversible and irreversible processes

b. Entropy
   Second Law of Thermodynamics

   Boltzmann's Equation

   Third Law of Thermodynamics

   entropy changes in chemical reactions

c. Gibbs Free Energy
   temperature

   equilibrium constant

   standard free-energy change (ΔG°) and the equilibrium constant (K)

Topic 16: Electrochemistry

a. Oxidation-reduction reactions
   oxidation states

   balancing redox equations

b. Voltaic cells
   cell potentials under standard conditions

   strengths of oxidizing and reducing agents

c. Cell potentials
   standard reduction potentials

   electromotive force (emf)

   equilibrium constants

   Nernst Equation

Topic 17: Nuclear Chemistry

a. Radioactivity
   nuclear equations

   types of radioactive emissions

b. Nuclear decay
   nuclear stability

   half-life

   transmutations

c. Energy Changes and Nuclear Energy
   $E = mc^2$

   fission and fusion

   nuclear reactors and radioactive waste

Topic 18: Organic and Biological Chemistry

a. Hydrocarbons
   structures

   nomenclature

   properties

b. Unsaturated and aromatic hydrocarbons
   addition and substitution reactions

c. Functional groups
   alcohols

   ethers

   aldehydes and ketones

   carboxylic acids and esters

Appendix B

Appendix B Section 1

Descriptives

Semester

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Fall 2012 Chem 170	54	15.8	16.6	16.6
	Fall 2013 Chem 170	90	26.3	27.6	44.2
	Spring 2013 Chem 175	111	32.5	34.0	78.2
	Spring 2014 Chem 175	71	20.8	21.8	100.0
	Total	326	95.3	100.0	
Missing	111	16	4.7		
Total		342	100.0		

Wald-Wolfowitz Runs Test

Runs Test

	PCA - All Student Percentiles	Fall 2012 Semester	Spring 2013	Fall 2013 Semester	Spring 2014 Semester
Test Value[a]	93	80	93	92	70
Cases < Test Value	262	23	89	65	27
Cases >= Test Value	64	31	22	25	44
Total Cases	326	54	111	90	71
Number of Runs	100	29	38	37	31
Z	-.682	.448	.519	-.029	-.879
Asymp. Sig. (2-tailed)	.495	.654	.604	.977	.379

a. Mode

Kruskal-Wallis H Test

Ranks

	Semester	N	Mean Rank
PCA - All Student Percentiles	Fall 2012 Chem 170 (0)	54	173.41
	Fall 2013 Chem 170 (1)	90	174.77
	Spring 2013 Chem 175 (2)	111	160.76
	Spring 2014 Chem 175 (3)	71	145.96
	Total	326	

Test Statistics[a,b]

	PCA - All Student Percentiles
Chi-Square	4.44
df	3
Asymp. Sig.	.22

a. Kruskal Wallis Test
b. Grouping Variable: Semester

One-Way ANOVA

ANOVA

PCA - All Student Percentiles

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2302.581	3	767.527	2.235	.084
Within Groups	110580.943	322	343.419		
Total	112883.525	325			

Appendix B Section 2

Descriptives

Descriptives

	Exam			Statistic	Std. Error
PCA - All Student Percentiles	Exam One (1)	Mean		79.32	1.618
		95% Confidence Interval for Mean	Lower Bound	76.11	
			Upper Bound	82.53	
		5% Trimmed Mean		80.69	
		Median		82.00	
		Variance		261.816	
		Std. Deviation		16.181	
		Minimum		22	
		Maximum		99	
		Range		77	
		Interquartile Range		21	
		Skewness		-1.191	.241
		Kurtosis		1.237	.478
	Exam Two (2)	Mean		74.25	2.037
		95% Confidence Interval for Mean	Lower Bound	70.20	
			Upper Bound	78.29	
		5% Trimmed Mean		76.06	
		Median		77.00	
		Variance		385.753	
		Std. Deviation		19.641	
		Minimum		1	
		Maximum		100	
		Range		99	
		Interquartile Range		25	
		Skewness		-1.360	.250
		Kurtosis		2.546	.495
	Exam Three (3)	Mean		74.22	2.193
		95% Confidence Interval for Mean	Lower Bound	69.86	
			Upper Bound	78.59	
		5% Trimmed Mean		75.46	
		Median		80.00	
		Variance		389.675	
		Std. Deviation		19.740	
		Minimum		14	
		Maximum		99	
		Range		85	
		Interquartile Range		31	
		Skewness		-.885	.267
		Kurtosis		-.003	.529
	Exam Four (4)	Mean		73.08	2.624
		95% Confidence Interval for Mean	Lower Bound	67.81	
			Upper Bound	78.35	
		5% Trimmed Mean		74.42	
		Median		77.00	
		Variance		358.072	
		Std. Deviation		18.923	
		Minimum		21	
		Maximum		97	
		Range		76	
		Interquartile Range		29	
		Skewness		-.955	.330
		Kurtosis		.285	.650

Kruskal-Wallis H Test

Ranks

	Exam Per Semester	N	Mean Rank
PCA - All Student Percentiles	Exam 1 Fall 2012	20	177.45
	Exam 1 Spring 2013	25	178.54
	Exam 1 Fall 2013	28	175.48
	Exam 1 Spring 2014	25	159.14
	Exam 2 Fall 2012	14	169.43
	Exam 2 Spring 2013	25	167.18
	Exam 2 Fall 2013	27	159.43
	Exam 2 Spring 2014	21	105.50
	Exam 3 Fall 2012	15	144.83
	Exam 3 Spring 2013	16	209.78
	Exam 3 Fall 2013	24	133.98
	Exam 3 Spring 2014	22	145.48
	Exam 4 Spring 2013	23	127.96
	Exam 4 Fall 2013	27	146.81
	Total	312	

Test Statistics[a,b]

	PCA - All Student Percentiles
Chi-Square	21.494
df	13
Asymp. Sig.	.064

a. Kruskal Wallis Test
b. Grouping Variable: Exam Per Semester

Wald-Wolfowitz Runs Test

Comparisons Between Semesters

Comparison	Number of Runs		Z	Asymp. Sig (1-tailed)
Exam 1 Fall 2012	Minimum Possible	22	-.55	.29
Exam 1 Fall 2013	Maximum Possible	28	1.25	.90
Exam 2 Fall 2012	Minimum Possible	13	-2.10	.02
Exam 2 Fall 2013	Maximum Possible	19	.00	.50
Exam 3 Fall 2012	Minimum Possible	19	.00	.50
Exam 3 Fall 2013	Maximum Possible	23	1.9	.92
Exam 4 Fall 2012[a]	Minimum Possible	-	-	-
Exam 4 Fall 2013[a]	Maximum Possible	-	-	-
Exam 1 Spring 2013	Minimum Possible	20	-1.72	.04
Exam 1 Spring 2014	Maximum Possible	32	1.72	.96
Exam 2 Spring 2013	Minimum Possible	17	-1.90	.03
Exam 2 Spring 2014	Maximum Possible	21	-.70	.24
Exam 3 Spring 2013	Minimum Possible	15	-1.36	.09
Exam 3 Spring 2014	Maximum Possible	21	.67	.75
Exam 4 Spring 2013[b]	Minimum Possible	-	-	-
Exam 4 Spring 2014[b]	Maximum Possible	-	-	-

[a] No data for exam 4 Fall 2012.
[b]No data for exam Spring 2014.

One-Way ANOVA

ANOVA

PCA - All Student Percentiles

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2038.760	3	679.587	1.974	.118
Within Groups	110844.764	322	344.238		
Total	112883.525	325			

Directional Measures

			Value
Nominal by Interval	Eta	PCA - All Student Percentiles Dependent	.134
		Exam Dependent	.508

Appendix B Section 3

Descriptive

Topics

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Matter	13	3.8	4.0	4.0
	Atoms	17	5.0	5.2	9.2
	Mole	16	4.7	4.9	14.1
	Thermochemistry	11	3.2	3.4	17.5
	Electron Configuration	22	6.4	6.7	24.2
	Periodic Trends	8	2.3	2.5	26.7
	Bonding	13	3.8	4.0	30.7
	Geometry	19	5.6	5.8	36.5
	Gases	12	3.5	3.7	40.2
	Intermolecular Forces	13	3.8	4.0	44.2
	Kinetics	33	9.6	10.1	54.3
	Nuclear Chemistry	21	6.1	6.4	60.7
	Equilibria	27	7.9	8.3	69.0
	Acid Base	24	7.0	7.4	76.4
	Aqueous Equilibria	21	6.1	6.4	82.8
	Energy	28	8.2	8.6	91.4
	Electrochemistry	15	4.4	4.6	96.0
	Organic Chemistry	13	3.8	4.0	100.0
	Total	326	95.3	100.0	
Missing	111	16	4.7		
Total		342	100.0		

Wald-Wolfowitz Runs Test

Runs Test 3

	Topics
Test Value[a]	11
Cases < Test Value	144
Cases >= Test Value	182
Total Cases	326
Number of Runs	23
Z	-15.610
Asymp. Sig. (2-tailed)	.000

a. Mode

Kruskal-Wallis H Test

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of PCA - All Student Percentiles is the same across categories of Topics.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

One-Way ANOVA

ANOVA

PCA - All Student Percentiles

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	16925.419	17	995.613	3.196	.000
Within Groups	95958.106	308	311.552		
Total	112883.525	325			

Appendix B Section 4

Descriptive

Modified Bloom's Taxonomy

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Recall (0)	91	26.6	27.9	27.9
	Comprehension (1)	134	39.2	41.1	69.0
	Application (2)	70	20.5	21.5	90.5
	Analysis (3)	31	9.1	9.5	100.0
	Total	326	95.3	100.0	
Missing	111	16	4.7		
Total		342	100.0		

Descriptives

	Modified Bloom's Taxonomy			Statistic	Std. Error
PCA - All Student Percentiles	Recall (0)	Mean		82.84	1.67
		95% Confidence Interval for Mean	Lower Bound	79.52	
			Upper Bound	86.15	
		5% Trimmed Mean		84.40	
		Median		89.00	
		Variance		253.72	
		Std. Deviation		15.93	
		Minimum		27	
		Maximum		100	
		Range		73	
		Interquartile Range		19	
		Skewness		-1.45	.25
		Kurtosis		1.70	.50
	Comprehension (1)	Mean		77.46	1.24
		95% Confidence Interval for Mean	Lower Bound	75.01	
			Upper Bound	79.91	
		5% Trimmed Mean		78.29	
		Median		80.00	
		Variance		205.29	
		Std. Deviation		14.33	
		Minimum		29	
		Maximum		100	
		Range		71	
		Interquartile Range		22	
		Skewness		-.79	.21
		Kurtosis		.57	.42
	Application (2)	Mean		67.83	2.66
		95% Confidence Interval for Mean	Lower Bound	62.53	
			Upper Bound	73.13	
		5% Trimmed Mean		69.41	
		Median		72.00	
		Variance		494.23	
		Std. Deviation		22.23	
		Minimum		1	
		Maximum		99	
		Range		98	
		Interquartile Range		31	
		Skewness		-.94	.29
		Kurtosis		.62	.57
	Analysis (3)	Mean		63.97	4.02
		95% Confidence Interval for Mean	Lower Bound	55.77	
			Upper Bound	72.17	
		5% Trimmed Mean		64.60	
		Median		74.00	
		Variance		499.77	
		Std. Deviation		22.36	
		Minimum		21	
		Maximum		97	
		Range		76	
		Interquartile Range		32	
		Skewness		-.48	.42
		Kurtosis		-.95	.82

Wald-Wolfowitz Runs Test between Levels of the Taxonomy

Runs Test

	Modified Bloom's Taxonomy
Test Value[a]	1
Cases < Test Value	91
Cases >= Test Value	235
Total Cases	326
Number of Runs	118
Z	-1.958
Asymp. Sig. (2-tailed)	.050

a. Median

Kruskal-Wallis H Test between Levels of the Taxonomy

Ranks

	Modified Bloom's Taxonomy	N	Mean Rank
PCA - All Student Percentiles	Recall (0)	91	206.31
	Comprehension (1)	134	164.74
	Application (2)	70	128.64
	Analysis (3)	31	111.21
	Total	326	

Test Statistics[a,b]

	PCA - All Student Percentiles
Chi-Square	37.943
df	3
Asymp. Sig.	.000

a. Kruskal Wallis Test
b. Grouping Variable: Modified Bloom's Taxonomy

One-Way ANOVA

ANOVA

PCA - All Student Percentiles

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	13650.773	3	4550.258	14.765	.000
Within Groups	99232.752	322	308.176		
Total	112883.525	325			

Kruskall-Wallis H Test within Each Level Grouped by Exams

Ranks

	Exam	N	Mean Rank
Recall Bloom's	Exam One (1)	39	45.33
	Exam Two (2)	27	43.20
	Exam Three (3)	16	57.19
	Exam Four (4)	9	37.39
	Total	91	
Comprehension Bloom's	Exam One (1)	44	74.73
	Exam Two (2)	43	60.59
	Exam Three (3)	26	71.90
	Exam Four (4)	21	61.05
	Total	134	
Application Bloom's	Exam One (1)	11	41.09
	Exam Two (2)	17	30.74
	Exam Three (3)	26	33.38
	Exam Four (4)	16	40.16
	Total	70	
Analysis Bloom's	Exam One (1)	6	17.17
	Exam Two (2)	6	16.67
	Exam Three (3)	13	14.96
	Exam Four (4)	6	16.42
	Total	31	

Test Statistics[a,b]

	Recall Bloom's	Comprehension Bloom's	Application Bloom's	Analysis Bloom's
Chi-Square	4.164	3.804	2.882	.314
df	3	3	3	3
Asymp. Sig.	.244	.283	.410	.957

a. Kruskal Wallis Test
b. Grouping Variable: Exam

Kruskall-Wallis H Test within Each Level Grouped by Semesters

Ranks

	Semester	N	Mean Rank
Recall Bloom's	Fall 2012 Chem 170 (0)	19	51.71
	Fall 2013 Chem 170 (1)	35	45.51
	Spring 2013 Chem 175 (2)	22	43.66
	Spring 2014 Chem 175 (3)	15	43.33
	Total	91	
Comprehension Bloom's	Fall 2012 Chem 170 (0)	22	66.20
	Fall 2013 Chem 170 (1)	36	71.17
	Spring 2013 Chem 175 (2)	42	63.60
	Spring 2014 Chem 175 (3)	34	69.28
	Total	134	
Application Bloom's	Fall 2012 Chem 170 (0)	10	37.55
	Fall 2013 Chem 170 (1)	17	36.62
	Spring 2013 Chem 175 (2)	28	40.64
	Spring 2014 Chem 175 (3)	15	23.27
	Total	70	
Analysis Bloom's	Fall 2012 Chem 170 (0)	3	11.67
	Fall 2013 Chem 170 (1)	2	14.25
	Spring 2013 Chem 175 (2)	19	18.66
	Spring 2014 Chem 175 (3)	7	11.14
	Total	31	

Test Statistics[a,b]

	Recall Bloom's	Comprehension Bloom's	Application Bloom's	Analysis Bloom's
Chi-Square	1.228	.843	7.365	4.382
df	3	3	3	3
Asymp. Sig.	.746	.839	.061	.223

a. Kruskal Wallis Test
b. Grouping Variable: Semester

Appendix B Section 5

Descriptives

Statistics
Question Type

N	Valid	326
	Missing	16
Median		1.00
Skewness		-.26
Std. Error of Skewness		.14
Range		2

Question Type

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Recall (0)	81	23.7	24.8	24.8
	Algorithmic (1)	118	34.5	36.2	61.0
	Conceptual (2)	127	37.1	39.0	100.0
	Total	326	95.3	100.0	
Missing	111	16	4.7		
Total		342	100.0		

Descriptives

	Question Type			Statistic	Std. Error
PCA - All Student Percentiles	Recall (0)	Mean		81.75	1.79
		95% Confidence Interval for Mean	Lower Bound	78.19	
			Upper Bound	85.31	
		5% Trimmed Mean		83.05	
		Median		88.00	
		Variance		259.01	
		Std. Deviation		16.09	
		Minimum		35	
		Maximum		99	
		Range		64	
		Interquartile Range		22	
		Skewness		-1.13	.27
		Kurtosis		.37	.53
	Algorithmic (1)	Mean		78.61	1.33
		95% Confidence Interval for Mean	Lower Bound	75.98	
			Upper Bound	81.24	
		5% Trimmed Mean		79.61	
		Median		80.00	
		Variance		207.37	
		Std. Deviation		14.40	
		Minimum		14	
		Maximum		99	
		Range		85	
		Interquartile Range		20	
		Skewness		-1.17	.22
		Kurtosis		2.52	.44
	Conceptual (2)	Mean		68.91	1.91
		95% Confidence Interval for Mean	Lower Bound	65.14	
			Upper Bound	72.68	
		5% Trimmed Mean		70.18	
		Median		74.00	
		Variance		460.90	
		Std. Deviation		21.47	
		Minimum		1	
		Maximum		100	
		Range		99	
		Interquartile Range		32	
		Skewness		-.84	.21
		Kurtosis		.17	.43

Wald-Wolfowitz Runs Test between Levels of Question Type

Runs Test

	Question Type
Test Value[a]	1
Cases < Test Value	81
Cases >= Test Value	245
Total Cases	326
Number of Runs	98
Z	-3.680
Asymp. Sig. (2-tailed)	.000

a. Median

Kruskal-Wallis H Test between Levels of Question Type

Ranks

	Question Type	N	Mean Rank
PCA - All Student Percentiles	Recall (0)	81	199.54
	Algorithmic (1)	118	172.09
	Conceptual (2)	127	132.53
	Total	326	

Test Statistics[a,b]

	PCA - All Student Percentiles
Chi-Square	26.557
df	2
Asymp. Sig.	.000

a. Kruskal Wallis Test
b. Grouping Variable: Question Type

One-Way ANOVA

ANOVA

PCA - All Student Percentiles

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	9827.529	2	4913.764	15.401	.000
Within Groups	103055.996	323	319.059		
Total	112883.525	325			

Kruskal-Wallis H Test within Each Level of Question Type Grouped by Exam

Ranks

	Exam	N	Mean Rank
QT Recall	Exam One (1)	35	43.47
	Exam Two (2)	21	38.88
	Exam Three (3)	15	48.40
	Exam Four (4)	10	25.70
	Total	81	
QT Algorithmic	Exam One (1)	48	62.82
	Exam Two (2)	34	53.66
	Exam Three (3)	21	57.67
	Exam Four (4)	15	64.67
	Total	118	
QT Conceptual	Exam One (1)	17	63.53
	Exam Two (2)	38	62.25
	Exam Three (3)	45	64.33
	Exam Four (4)	27	66.20
	Total	127	

Test Statistics[a,b]

	QT Recall	QT Algorithmic	QT Conceptual
Chi-Square	6.283	1.848	.189
df	3	3	3
Asymp. Sig.	.099	.605	.979

a. Kruskal Wallis Test
b. Grouping Variable: Exam

Kruskal-Wallis H Test within Each Level of Question Type Grouped by Semester

Ranks

	Semester	N	Mean Rank
QT Recall	Fall 2012 Chem 170 (0)	17	48.65
	Fall 2013 Chem 170 (1)	31	45.35
	Spring 2013 Chem 175 (2)	23	32.50
	Spring 2014 Chem 175 (3)	10	34.05
	Total	81	
QT Algorithmic	Fall 2012 Chem 170 (0)	18	53.22
	Fall 2013 Chem 170 (1)	29	60.38
	Spring 2013 Chem 175 (2)	47	58.86
	Spring 2014 Chem 175 (3)	24	64.40
	Total	118	
QT Conceptual	Fall 2012 Chem 170 (0)	19	68.66
	Fall 2013 Chem 170 (1)	30	63.42
	Spring 2013 Chem 175 (2)	41	69.72
	Spring 2014 Chem 175 (3)	37	55.74
	Total	127	

Test Statistics[a,b]

	QT Recall	QT Algorithmic	QT Conceptual
Chi-Square	6.747	1.134	3.166
df	3	3	3
Asymp. Sig.	.080	.769	.367

a. Kruskal Wallis Test
b. Grouping Variable: Semester

Appendix B Section 6

Descriptives

Cognitive Level

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	LOCS (0)	222	64.9	68.1	68.1
	HOCS (1)	104	30.4	31.9	100.0
	Total	326	95.3	100.0	
Missing	111	16	4.7		
Total		342	100.0		

Wald-Wolfowitz Runs Test

Runs Test 2

	Cognitive Level
Test Value[a]	.32
Cases < Test Value	222
Cases >= Test Value	104
Total Cases	326
Number of Runs	133
Z	-1.232
Asymp. Sig. (2-tailed)	.218

a. Mean

Kruskal-Wallis H Test

Ranks

	Cognitive Level	N	Mean Rank
PCA - All Student Percentiles	LOCS (0)	222	181.65
	HOCS (1)	104	124.75
	Total	326	

Test Statistics[a,b]

	PCA - All Student Percentiles
Chi-Square	25.830
df	1
Asymp. Sig.	.000

a. Kruskal Wallis Test
b. Grouping Variable: Cognitive Level

One-Sample T Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
PCA - All Student Percentiles	326	75.61	18.637	1.032

One-Sample Test

	Test Value = 0					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
PCA - All Student Percentiles	73.252	325	.000	75.610	73.58	77.64

Kruskal-Wallis H Test within Each Level of Cognitive Skill Level Grouped by Exam

Ranks

	Exam	N	Mean Rank
LOCS	Exam One (1)	77	112.01
	Exam Two (2)	58	109.97
	Exam Three (3)	53	116.36
	Exam Four (4)	33	101.85
	Total	221	
HOCS	Exam One (1)	23	53.09
	Exam Two (2)	34	52.40
	Exam Three (3)	28	48.64
	Exam Four (4)	19	57.66
	Total	104	

Test Statistics[a,b]

	LOCS	HOCS
Chi-Square	1.084	1.023
df	3	3
Asymp. Sig.	.781	.796

a. Kruskal Wallis Test
b. Grouping Variable: Exam

Kruskal-Wallis H Test within Each Level of Cognitive Skill Level Grouped by Semester

Ranks

	Semester	N	Mean Rank
LOCS	Fall 2012 Chem 170 (0)	40	116.31
	Fall 2013 Chem 170 (1)	68	117.51
	Spring 2013 Chem 175 (2)	69	101.40
	Spring 2014 Chem 175 (3)	44	111.17
	Total	221	
HOCS	Fall 2012 Chem 170 (0)	14	50.75
	Fall 2013 Chem 170 (1)	21	55.10
	Spring 2013 Chem 175 (2)	42	59.24
	Spring 2014 Chem 175 (3)	27	40.91
	Total	104	

Test Statistics[a,b]

	LOCS	HOCS
Chi-Square	2.539	6.288
df	3	3
Asymp. Sig.	.468	.098

a. Kruskal Wallis Test
b. Grouping Variable: Semester

Appendix B Section 7

Descriptives

Number of Steps

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	166	48.5	50.9	50.9
	2	84	24.6	25.8	76.7
	3	30	8.8	9.2	85.9
	4	9	2.6	2.8	88.7
	5	33	9.6	10.1	98.8
	6	2	.6	.6	99.4
	7	2	.6	.6	100.0
	Total	326	95.3	100.0	
Missing	111	16	4.7		
Total		342	100.0		

Statistics
Number of Steps

N	Valid	326
	Missing	2
Mean		2.00
Std. Error of Mean		.076
Median		1.00
Mode		1
Std. Deviation		1.37
Variance		1.88
Skewness		1.45
Std. Error of Skewness		.14
Range		6
Percentiles	4	1.00

Wald-Wolfowitz Runs Test

Runs Test

	Number of Steps
Test Value[a]	2.00
Cases < Test Value	166
Cases >= Test Value	160
Total Cases	326
Number of Runs	125
Z	-4.32
Asymp. Sig. (2-tailed)	.000

a. Mean

Kolmogorov-Smirnov Test

One-Sample Kolmogorov-Smirnov Test

		Number of Steps
N		326
Normal Parameters[a,b]	Mean	2.00
	Std. Deviation	1.371
Most Extreme Differences	Absolute	.276
	Positive	.276
	Negative	-.234
Kolmogorov-Smirnov Z		4.976
Asymp. Sig. (2-tailed)		.000

a. Test distribution is Normal.
b. Calculated from data.

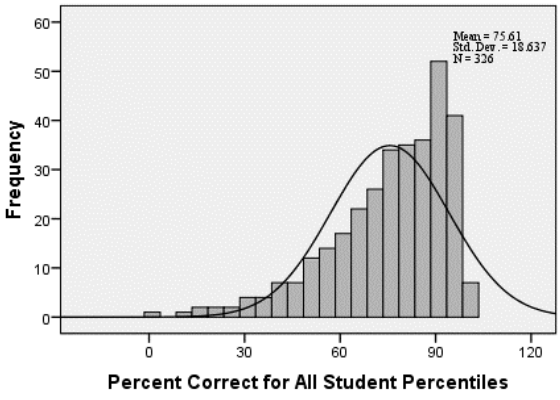
Appendix B Section 8

Descriptives

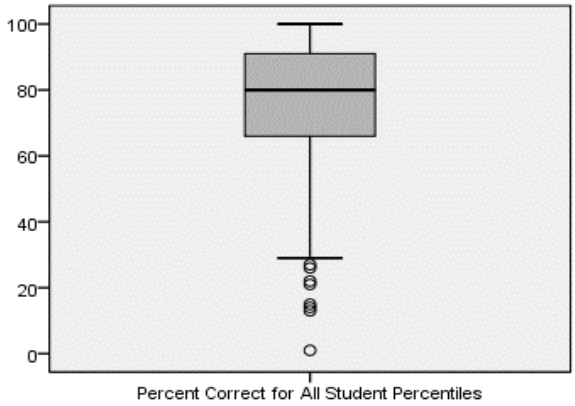
	N	Range	Minimum	Maximum	Variance	Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
PCA-All Percentiles	326	99	1	100	347.3	1.20	.27
PCA-Upper Percentile	326	100	0	100	215.1	11.97	.27
PCA-Middle Percentile	326	90	10	100	407.01	1.09	.27
PCA-Lower Percentile	326	96	4	100	599.98	-1.06	.27
Missing N	16						
Valid N (listwise)	342						

Visual Representations of the Distribution of Percent Correct Answers for all Student Percentiles

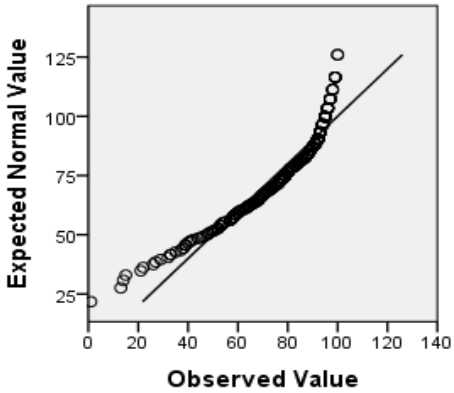
Histogram



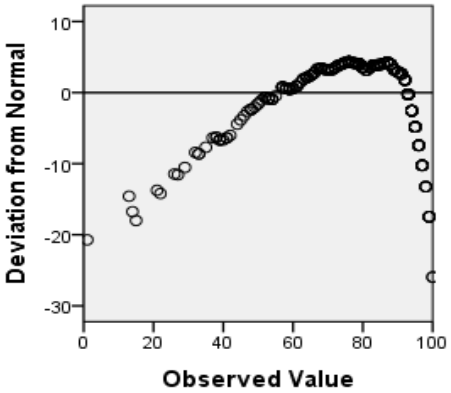
Box Plot



Q-Q Plot



Detrended Q-Q



Kolmogorov–Smirnov D Test

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
Percent Correct for All Student Percentiles	326	75.61	18.64	1	100

One-Sample Kolmogorov-Smirnov Test

		Percent Correct for All Student Percentiles
N		326
Normal Parameters[a,b]	Mean	75.61
	Std. Deviation	18.64
Most Extreme Differences	Absolute	.11
	Positive	.10
	Negative	-.11
Kolmogorov-Smirnov Z		1.92
Asymp. Sig. (2-tailed)		.001

a. Test distribution is Normal.
b. Calculated from data.

Wald-Wolfowitz Runs Test

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles 25th	50th (Median)	75th
Percent Correct for All Student Percentiles	326	75.61	18.64	1	100	65.75	80.00	91.00

Runs Test

	Percent Correct for All Student Percentiles
Test Value[a]	80
Cases < Test Value	162
Cases >= Test Value	164
Total Cases	326
Number of Runs	152
Z	-1.33
Asymp. Sig. (2-tailed)	.18

a. Median

Kruskal-Wallis H Test

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles 25th	50th (Median)	75th
Percent Correct for All Student Percentiles	326	75.61	18.637	1	100	65.75	80.00	91.00
Semester	326	1.61	1.004	0	3	1.00	2.00	2.00

Ranks

	Semester	N	Mean Rank
Percent Correct for All Student Percentiles	Fall 2012 Chem 170	54	173.41
	Fall 2013 Chem 170	90	174.77
	Spring 2013 Chem 175	111	160.76
	Spring 2014 Chem 175	71	145.96
	Total	326	

Test Statistics[a,b]

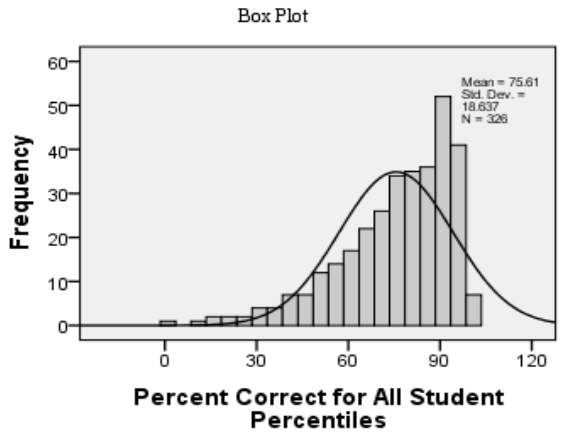
	Percent Correct for All Student Percentiles
Chi-Square	4.44
df	3
Asymp. Sig.	.218

a. Kruskal Wallis Test
b. Grouping Variable: Semester

Appendix B Section 9

Descriptive Data for the Original Data and the Transformed Data

No Transformation of Data

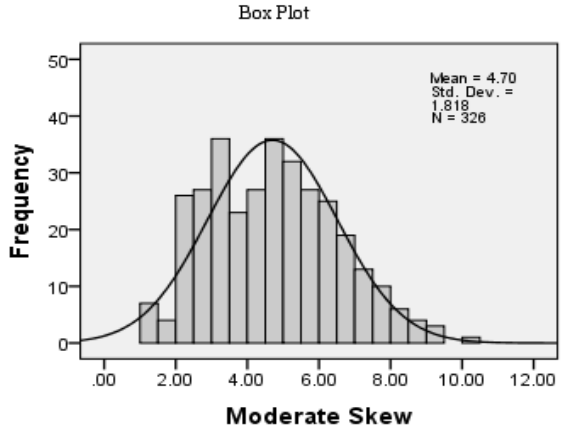


Statistics

		Percent Correct for All Student Percentiles
N	Valid	326.00
	Missing	16.00
Mean		75.61
Std. Error of Mean		1.03
Median		80.00
Std. Deviation		18.64
Variance		347.33
Mode		93
Skewness		-1.14
Std. Error of Skewness		.14
Kurtosis		1.20
Std. Error of Kurtosis		.27
Range		99.00

Moderate Skew Transformation:

$$\sqrt{((\text{highest exam score} + 1) - \text{each exam score})}$$

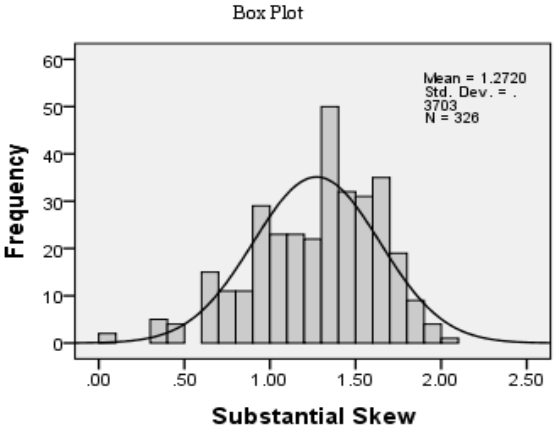


Statistics

		Moderate Skew
N	Valid	326.00
	Missing	16.00
Mean		4.70
Std. Error of Mean		.10
Median		4.58
Std. Deviation		1.82
Mode		2.83
Variance		3.31
Skewness		.34
Std. Error of Skewness		.14
Kurtosis		-.44
Std. Error of Kurtosis		.27
Range		9.00

Substatantial Skew Transformation:

$$\log\left((\text{highest exam score} + 1) - \text{each exam score}\right)$$

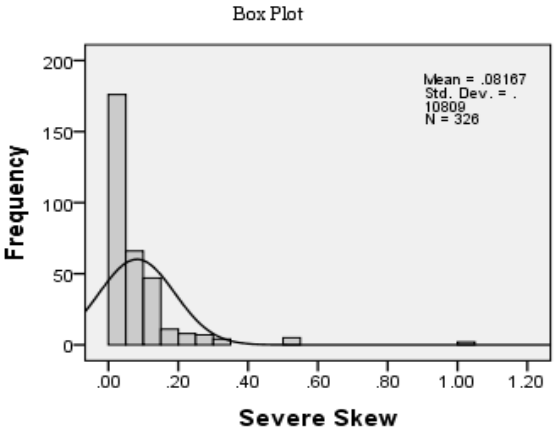


Statistics

		Substantial Skew
N	Valid	326.00
	Missing	16.00
Mean		1.27
Std. Error of Mean		.02
Median		1.32
Std. Deviation		.37
Mode		.90
Variance		.14
Skewness		-.60
Std. Error of Skewness		.14
Kurtosis		.16
Std. Error of Kurtosis		.27
Range		2.00

Severe Skew Transformation:

$$\frac{1}{\left((\text{highest exam score} + 1) - \text{each exam score}\right)}$$



Statistics

		Severe Skew
N	Valid	326.00
	Missing	16.00
Mean		.08
Std. Error of Mean		.01
Median		.05
Std. Deviation		.11
Mode		.13
Variance		.01
Skewness		4.92
Std. Error of Skewness		.14
Kurtosis		33.74
Std. Error of Kurtosis		.27
Range		.99

Appendix B Section 10

Lin-Hartman Paper

Case Processing Summary

		N	%
	Valid	121	35.4
a s e s	Excluded[a]	221	64.6
	Total	342	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	N of Items
.278	2

Intraclass Correlation Coefficient

	Intraclass Correlation[b]	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.162[a]	-.017	.330	1.385	120	120	.038
Average Measures	.278[c]	-.034	.496	1.385	120	120	.038

Two-way mixed effects model where people effects are random and measures effects are fixed.
a. The estimator is the same, whether the interaction effect is present or not.
b. Type C intraclass correlation coefficients using a consistency definition-the between-measure variance is excluded from the denominator variance.
c. This estimate is computed assuming the interaction effect is absent, because it is not estimable otherwise.

Question Type

Case Processing Summary

		N	%
	Valid	121	35.4
Cases	Excluded[a]	221	64.6
	Total	342	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	N of Items
.707	2

Intraclass Correlation Coefficient

	Intraclass Correlation[b]	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.547[a]	.409	.661	3.417	120	120	.000
Average Measures	.707[c]	.581	.796	3.417	120	120	.000

Two-way mixed effects model where people effects are random and measures effects are fixed.
a. The estimator is the same, whether the interaction effect is present or not.
b. Type C intraclass correlation coefficients using a consistency definition-the between-measure variance is excluded from the denominator variance.
c. This estimate is computed assuming the interaction effect is absent, because it is not estimable otherwise.

Modified Bloom's Taxonomy

Case Processing Summary

		N	%
Cases	Valid	121	35.4
	Excluded[a]	221	64.6
	Total	342	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	N of Items
.733	2

Intraclass Correlation Coefficient

	Intraclass Correlation[b]	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.578[a]	.447	.686	3.745	120	120	.000
Average Measures	.733[c]	.617	.814	3.745	120	120	.000

Two-way mixed effects model where people effects are random and measures effects are fixed.
a. The estimator is the same, whether the interaction effect is present or not.
b. Type C intraclass correlation coefficients using a consistency definition-the between-measure variance is excluded from the denominator variance.
c. This estimate is computed assuming the interaction effect is absent, because it is not estimable otherwise.

Cognitive Skill Level

Case Processing Summary

		N	%
Cases	Valid	121	35.4
	Excluded[a]	221	64.6
	Total	342	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	N of Items
.579	2

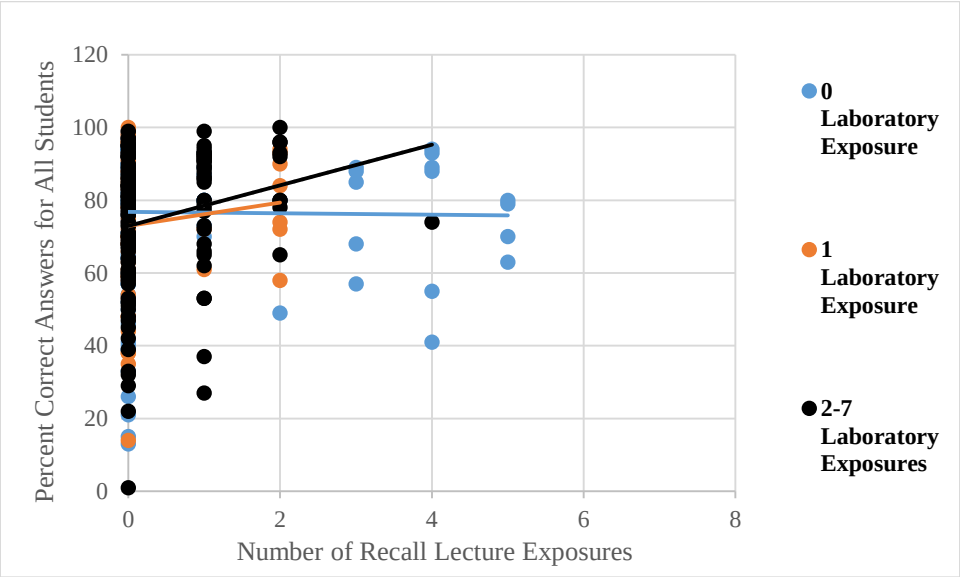
Intraclass Correlation Coefficient

	Intraclass Correlation[b]	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.407[a]	.248	.546	2.375	120	120	.000
Average Measures	.579[c]	.397	.706	2.375	120	120	.000

Two-way mixed effects model where people effects are random and measures effects are fixed.
a. The estimator is the same, whether the interaction effect is present or not.
b. Type C intraclass correlation coefficients using a consistency definition-the between-measure variance is excluded from the denominator variance.
c. This estimate is computed assuming the interaction effect is absent, because it is not estimable otherwise.

Appendix C

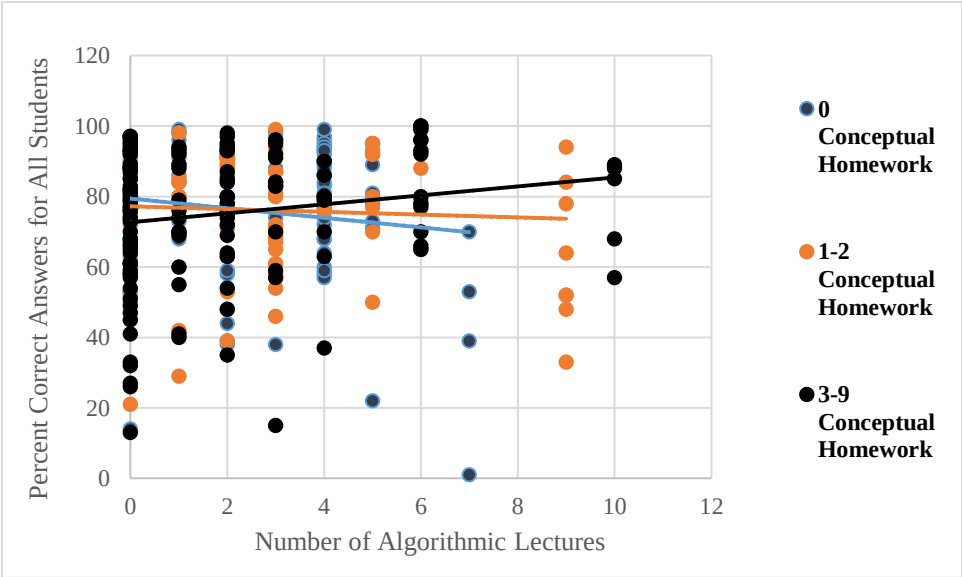
Appendix C Figure 1: Interaction of Recall Lecture Exposures by Laboratory Exposures for PCA-All Students Group



Regression Equations for Interaction of Recall Lecture with Laboratory Exposures

0 Laboratory Exposures	$\hat{y} = -0.019x + 76.79$	$R2 = 0.0002$
1 Laboratory Exposures	$\hat{y} = 3.18x + 72.96$	$R2 = 0.0125$
2-7 Laboratory Exposures	$\hat{y} = 5.60x + 72.89$	$R2 = 0.0433$

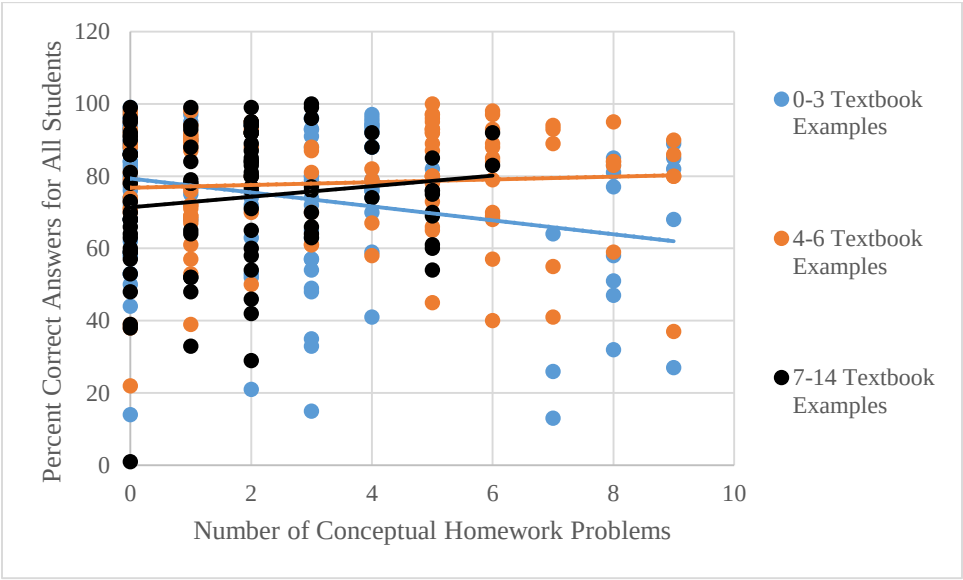
Appendix C Figure 2: Interaction of Algorithmic Lectures Exposures with Conceptual Homework Problems



Regression Equations for Interaction of Algorithmic Lectures with Conceptual Homework

0 Conceptual Homework Problems	$\hat{y}$= -1.36x + 79.44	R2= 0.0185
1-2 Conceptual Homework Problems	$\hat{y}$= -0.39x + 77.19	R2= 0.0033
3-9 Conceptual Homework Problems	$\hat{y}$= 1.27x + 72.72	R2= 0.0276

Appendix C Figure 3: Interaction of Textbook Examples with Conceptual Homework Problems



Regression Equations for Interaction of Conceptual Homework Problems with Textbook Examples

0 – 3 Textbook Examples	$\hat{y} = -1.92x + 79.33$ R2= 0.0648
4 – 6 Textbook Examples	$\hat{y} = 0.38x + 76.81$ R2= 0.0041
7 -14 Textbook Examples	$\hat{y} = 1.46x + 71.38$ R2= 0 .0174