May 2019

# From General Chemistry to Anatomy and Physiology: Revalidating and Adapting Assessments and Models

Victoria Fisher
*University of Wisconsin-Milwaukee*

# FROM GENERAL CHEMISTRY TO ANATOMY AND PHYSIOLOGY: REVALIDATING AND ADAPTING ASSESSMENTS AND MODELS

by

Victoria K. Fisher-Keough

A Thesis Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Master of Science

in Chemistry

at

The University of Wisconsin-Milwaukee

May 2019

ABSTRACT

# FROM GENERAL CHEMISTRY TO ANATOMY AND PHYSIOLOGY: REVALIDATING AND ADAPTING ASSESSMENTS AND MODELS

by

Victoria K. Fisher-Keough

The University of Wisconsin-Milwaukee, 2019
Under the Supervision of Professor Kristen L. Murphy, PhD

The 1980s saw an increasing demand for education standards that would create a scientifically literate society. In response, the American Association for the Advancement of Science (AAAS) published a report that outlined four themes that are characteristic of a scientifically literate individual: systems, models, constancy and change, and scale[1]. In 1993, the AAAS published the Benchmarks for Science Literacy which outlined common scientific skills that a student should be able to demonstrate by grades 2, 5, 8, and 12[2]. Beyond the AAAS scale was not included in national science educational standards until 2012 when the National Research Council released the Framework for K-12 science education followed by the Next Generation Science Standards in 2013. Scale was included as a cross-cutting concept titled "Scale, Proportion, and Quantity"[3]. Because proportion and quantity were included along with scale, some instructors who cover proportion and quantity believe that they also cover scale but may not have fully addressed the scale portion of the cross-cutting concept.

[1] American Association for the Advancement of Science, Project 2061; *Science for all Americans: a project 2061 report on literacy goals in science, mathematics, and technology*; Washington, D.C., 1989.
[2] American Association for the Advancement of Science, Project 2061; *Benchmarks for science literacy*; New York, New York: Oxford University Press, 1993.
[3] National Research Council; *Next Generation Science Standards: for states, by states*; Washington, D.C., National Academies Press: Washington, D.C., 2013.

Previous research in general chemistry I and scale led to the development of two instruments: the Scale Literacy Skills Test (SLST) and Scale Concept Inventory (SCI)[4]. The average of the two assessments generated a Scale Literacy Score for a student providing a measure of their scale ability. Previous research has shown that scale literacy is a better predictor for success in chemistry than traditional measures. Scale has been systematically integrated as a theme in the undergraduate chemistry curriculum in lecture, laboratory, and supplemental instruction activities. When scale was integrated in all components of the course there was an increase in student learning as measured by final exam performance. Scale as a cross-cutting concept has applications beyond that of only chemistry, e.g. biology. When transferring disciplines from chemistry to biological sciences, the existing scale instruments, SLST and SCI, cannot be assumed to be valid. Before investigating students' ability in scale in biological sciences the existing instruments were tested for reliability and validity. Once this was complete, the SLST and SCI were used to measure scale ability in Anatomy and Physiology I.

The goal of this project is studying student scale understanding across STEM disciplines. This continues the previous research in General Chemistry II and adapts the research for Anatomy and Physiology I[5]. This thesis contains the details of three studies between two courses covering student scale conception and scale's relation, if any, to final exam performance. The first (**Chapter 3**) discusses the development and implementation of two supplemental instruction online adaptive activities for General Chemistry II students. **Chapter 4** details semi-structured interviews with Anatomy and Physiology I students with regards to their scale

[4] Gerlach, Trate, Blecking, Geissinger, and Murphy 2014: 1538-1545
[5] Trate 2017: 1-205

conception. **Chapter 5** details the building of a multiple regression model to predict cumulative final exam score for the Anatomy and Physiology I course.

## References:

American Association for the Advancement of Science, Project 2061; *Science for all Americans: a project 2061 report on literacy goals in science, mathematics, and technology*; Washington, D.C., 1989.

American Association for the Advancement of Science, Project 2061; *Benchmarks for science literacy*; New York, New York: Oxford University Press, 1993.

Gerlach, K.; Trate, J.; Blecking, A.; Geissinger, P.; Murphy, K. (2014). Valid and Reliable Assessments to Measure Scale Literacy of Students in Introductory College Chemistry Courses. *Journal of Chemical Education*. *91*, 1538-1545.

National Research Council; *Next Generation Science Standards: for states, by states*; Washington, D.C., National Academies Press: Washington, D.C., 2013.

Trate, J. (2017). *Integrating Scale-Themed Instruction Across the General Chemistry Curriculum and Selected In-Depth Studies* (Doctoral dissertation). University of Wisconsin-Milwaukee, Milwaukee, WI.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

## LIST OF ABBREVIATIONS

| | |
|---|---|
| ACT COMP | ACT Composite |
| A & P | Anatomy and Physiology |
| AAAS | American Association for the Advancement of Science |
| ACT ENGL | ACT English |
| ACT MATH | ACT Mathematics |
| ACT READ | ACT Reading |
| ACT SCIRE | ACT Writing and Science |
| adj | Adjusted (threats to validity removed) |
| adult | adult height |
| ALG | Algebra |
| atom nuc | atomic nucleus |
| bcm | bacterium |
| CTT | Classical Test Theory |
| field | football field |
| hair | hair width |
| jet | cruising altitude of 747 jet |
| math | Math algorithm questions |
| MBSC | Math basics |
| moon | Earth to moon |
| NCISE | National Center for Improving Science Education |
| NGSS | Next Generation Science Standards |
| pencil | new pencil length |
| post | End of the semester |
| pre | Beginning of the semester |
| SAO | Scale Anchoring Objects assessment |
| SCI | Scale Concept Inventory |
| SI | Supplemental Instruction |

| | |
|---|---|
| SLS | Scale Literacy Score |
| SLST | Scale Literacy Skills Test |
| SOQ | Scale of Objects Questionnaire |
| sun | Earth to sun |
| text | textbook |
| Theo | Theoretical questions |
| TRG | Trigonometry |
| virus | virus |
| WI | width of Wisconsin |

# ACKNOWLEDGMENTS

A special thanks to Dr. Kristen Murphy, my advisor, for her constant support, help, and advice. Additional heartfelt thanks to all members of my committee, Dr. Anja Blecking, Dr. Peter Geissinger, Dr. John Kirk, and Dr. Alan Schwabacher, for their feedback, contributions, and kindness throughout this process. Thank you to all former and current research group members: Jaclyn Trate, Shalini Srinivasan, David Schreurs, and Leandro Inacio De Souza along with the instructors, teaching assistants, and students at UWM without which this research would not be possible. A large thank you to my family and close friends for their never-ending support. To Julie, Gary, and Olivia for always inspiring me and Skyler, Sara, Megan, Lina, and Simon for keeping me focused on the goal.

# Chapter 1: Introduction and Literature Review

## 1.1 Introduction

### 1.1.1 Curriculum standards

The year 1989 brought changes for education. This was the beginning of a curriculum reform for K-12 backed by frameworks developed for mathematics, science, and technology curricula[6]. Two influential frameworks include the development of the Benchmarks for Science Literacy by the American Association for the Advancement of Science (AAAS) and reports by The National Center for Improving Science Education (NCISE)[7,8].

In 1989 AAAS's Project 2061 published a report titled *Science for All Americans* where the need for science literacy as well as recommendations of steps to take to form a scientifically literate society were discussed[9]. Four themes, including systems, models, constancy and change, and scale, were established as being important in science, mathematics, and technology while at the same time transcending the traditional focus of the subjects. Of the four themes AAAS identified, scale had no explicit scientific literature supporting its inclusion. In 1993, Project 2061 published The Benchmarks for Science Literacy which details specific targets for each theme. These targets were further broken down by grade level, 2, 5, 8, and 12, students should be able to demonstrate the targets for the themes in science, mathematics, and technology.

---

[6] Bybee 1995: 12-13
[7] Project 2061 American Association for the Advancement of Science 1989
[8] Bybee 1995: 12-13
[9] Project 2061 American Association for the Advancement of Science 1989

Meeting these benchmarks meant that students were on the path to becoming scientifically literate adults[10]. Again, scale was included with no basis in the existing scientific literature[11].

The NCISE created a framework based on *organizing concepts* for elementary school that was extended to middle and high schools[12]. The organizing concepts include cause and effect, change and conservation, diversity and variation, energy and matter, evolution and equilibrium, models and theories, probability and prediction, structure and function, systems and interaction, and time and scale[13]. NCISE used the organizing concepts to connect disciplines and provide curriculum learning objectives.

AAAS and NCISE both developed frameworks to improve how science is taught in schools as a way of developing scientifically literate adults but there continues to be a call to update and unify the science, mathematics, and technology standards across the United States of America. In 2011 the National Research Council (NRC) released a report titled: *A Framework for K-12 Science Education: Practices, Crosscutting Concepts and Core Ideas* which outlined three "dimensions" that students would "build on and revise" over many years[14]. The three dimensions are broken down into scientific and engineering practices, crosscutting concepts which transcend disciplines, and disciplinary core ideas in physical, life, earth, and space sciences, engineering, technology, and science applications[15]. The Next Generation Science Standards (NGSS) were written keeping the NRC report as the backbone. Where AAAS and NCISE defined frameworks, the NGSS provided instructors with expectations for students.

---

[10] Project 2061 American Association for the Advancement of Science 1993
[11] Project 2061 American Association for the Advancement of Science 1993
[12] Bybee 1995: 12-13
[13] Bybee 1995: 12-13
[14] National Research Council, Board on Science Education, Division of Behavioral and Social Sciences and Education 2011
[15] National Research Council, Board on Science Education, Division of Behavioral and Social Sciences and Education 2011

These performance expectations were accompanied by examples of how a student could demonstrate understanding within a particular standard[16].

### 1.1.2 Scale definition and expert perspectives

The Oxford Dictionary defines scale in a variety of ways including “the relative size or extent of something” and “a ratio of size”[17]. Both definitions refer to scale as a relationship or as a mathematical concept. Gary Lock and Brian Molyneaux describe how scale can be seen as a “mathematical abstraction”[18] and other definitions of scale include “any quantification of a property that is measured”[19]. Lock and Molyneaux discuss, for archeologists, how analysis and interpretation require “multiple scales” which can often be done using technology and allow scale to be ignored by the researcher.

An added layer of scale complexity comes from scale relating to “space, time and social position” and humans tend to use themselves as a means to create relationships between space, time, and objects and communicate this with one another making scale “a human phenomenon [that] is culturally constructed”[20]. Scale, then, is important to cultivate for it impacts how humans interact with each other and their careers. Thomas R. Tretter states that “In spite of the centrality of scale to many science disciplines, the pressure to cover specific content in a course may make it easy to overlook this unifying theme”[21]. Being able to understand life on different scales allows one to understand the world around them and use this knowledge in nearly every aspect of life.

---

[16] National Research Council; *Next Generation Science Standards: for states, by states* 2013
[17] Scale 2018
[18] Lock and Molyneaux 2006: xi-xii
[19] Jones and Taylor 2009: 191-221
[20] Lock and Molyneaux 2006: xi-xii
[21] Tretter and Jones 2003: 22-25

In addition to being present in every aspect of life and being identified as a cross-cutting concept, scale is important in both chemistry and biology specifically. M. Gail Jones and Amy R. Taylor interviewed 50 professionals about scale use in their careers as well as their scale development throughout their lives including learning that happened in school as well as out[22]. Of the 50 professionals interviewed, one chemist said "A lot of this you take for granted after a while in your work. You just are so comfortable with it that you don't pay too much attention to it. But it is obviously in the background of everything you do"[23]. All 50 professionals stressed the importance of scale. When prompted by the interviewer to discuss scale in their work, a cell biologist said "Everything. Absolutely everything. But it's really exciting to work with all those different scales"[24]. Scale, both in a mathematical sense and an abstract sense, has been considered by experts to be integral to a variety of careers and identified as an important concept for students to master.

# 1.2 Literature Review

## 1.2.1 Instructors and scale

While there are debates as to whether earlier research of scale exists[25], this review will begin with Roger David Trend's report investigating scale in a science context 2001. Trend published a report examining primary school teachers' conceptions of geological time[26]. In this experiment, primary school teachers were given two instruments used to identify their personal interest in certain topics, how often they touch on those topics in the classroom, and their deep

[22] Jones and Taylor 2009: 460-475
[23] Jones and Taylor 2009: 460-475
[24] Jones and Taylor 2009: 460-475
[25] Golledge, Gale, Pelligrino, and Doherty 1992: 223-244
[26] Trend 2001: 191-221

time perceptions. For personal interest and use in the classroom the teachers were given a questionnaire, on a 5-point response scale. For identifying their deep time perceptions, the teachers were given a “responding-to-objects” instrument which had them identify a list of events such as “the first fish appeared” on a 9-point scale ranging from “less than one thousand years ago” to “more than approx. a million million years ago”[27]. The teachers were found to be more comfortable with relative time than with absolute time and were more accurate with relative compared to absolute time. An example of relative time Is the big bang occurred before the extinction of the dinosaurs. An example of absolute time is the big bang occurred over 13 billion years ago and the extinction of dinosaurs occurred about 165 millions years ago.

### 1.2.2 Scale conception

Using Trend’s (2001) geological time instrument as a model, Thomas Tretter, Gail Jones, Thomas Andre, Atsuko Negishi, and James Minogue studied 5th, 7th, 9th, and 12th grade students along with doctoral students’ conceptions of scale[28]. Instead of using geological time Tretter et al. used linear distances. Students in 5th, 7th, and 9th grade were classified as novices, 12th grade as advanced students, and doctoral students as experts. This was done to see how scale conception changes as expertise develops. The first part of the study was the Scale of Objects Questionnaire (SOQ). Students were given 26 objects, such as “the distance from the Earth to the Moon,” and asked to determine a size range on a 12-point scale ranging from “<1 nm” to “>1 billion meters”. This was followed by a card sort activity where students were given 31 cards with the name of an object on them ranging in size from the subatomic to the galactic. The students sorted the cards into piles according to size. Similar to Trend, Tretter, et al., found

[27] Trend 2001: 218
[28] Tretter, Jones, Andre, Negishi, and Minogue 2006a: 282, 288

that students were more accurate when dealing with relative scale compared to absolute scale. Student's utilized landmarks to help establish size. Landmarks are objects that students use to determine the size of other items. The most explicit landmark that came through the interviews and SOQ was human height. The more advanced students utilized more landmarks than the novice students and the experts utilized more landmarks than the advanced students. Another way of saying this is that novice students had fewer distinct size categories and with increasing expertise, there was an increasing number and distinctiveness of the size categories.

Thomas Tretter, Gail Jones, and James Minogue continued to study scale perception of different expertise levels[29]. Students, grades 5th, 7th, 9th, 12th, and doctoral, were given the Scale Anchoring Objects assessment (SAO). The SAO listed a range of sizes, from "1 meter" to "1,000,000,000 meters (one billion meters)" in part A and "equal to your body length" to "1/1,000,000,000 your body length (one billionth your body length)" in part B. Next to each size was a space where students were instructed to write an object that they identified with being that size in both parts A and B. The researchers compiled a list of commonly identified objects, such as "virus" or "skyscraper". The data showed that for small lengths, such as virus, students tended to identify objects that were too large for the given length and for large lengths, such as skyscraper, students tended to identify objects that were too small for that length, but with increasing expertise there was an increase in correctly identifying objects for a given length. Interview data was collected along with the SAO regarding the strategies used to identify objects of a given length. An example of a specific strategy used was "atomic radii were listed in Angstroms and that's close to nanometer size"[30]. The interview and SAO data found that the more experienced a student was in scale conception, as measured by the number of objects

[29] Tretter, Jones, and Minogue 2006b: 1061-1085
[30] Tretter 2006b: 1067

correctly listed on the SAO, the greater number of "specific strategies" they were able to articulate during the interview. Advanced students and experts primarily used two types of strategies: mathematical computations, such as metric system use, or object comparison, such as adult height. Novice students gave vague answers, such as the smallest thing they could think of, when identifying strategies. The more experienced a student was in scale, the more comfort they expressed with the metric system. The experts also were able to transition between large and small objects by defining a new unit of measurement, either based on a measurement or an object. The more advanced students tended to demonstrate a "transition to thinking like the experts but was not as rich and detailed as the experts' descriptions"[31].

### 1.2.3 Scale in the undergraduate level

Su Swarat, et al., studied scale and size conception with undergraduate engineering students by conducting three exploratory studies: two interview and one survey[32]. The first interview study was a think-aloud interview with participants ordering objects according to size, e.g. human hair width. Participants were then instructed to "apply a numerical scale to the line to represent their size differences". In the second interview participants were provided three different options of number lines with objects placed on them as well as the option for the participant to create their own. Participants were instructed to choose the most appropriate option and explain their reasoning. The 3-item survey contained an item nearly identical to interview 2 and two items looking at fragmented versus continuous scale conception. Results showed four categories of student conceptions of scale: fragmented, linear, proportional, and logarithmic. A fragmented conception meant students do not understand scale is continuous.

[31] Tretter 2006b: 1077
[32] Swarat 2011: 512-533

While a number line may end the linear distances continue beyond the physical number line. A linear conception was defined by students placing objects on a scale based on their physical experience or observation of the object's size. A student had a proportional conception of scale when they exhibit descriptions or understanding that was a hybrid of the linear and logarithmic conceptions. The logarithmic conception was the "the most sophisticated conception of size and scale" that was observed during the interviews and survey. Students, as they improve or continue to improve their understanding of scale, are able to move through these conceptions of scale as they become more experienced.

### 1.2.4 Scale in chemistry

In 2014 scale began to be studied by Karrie Gerlach and colleagues through the adaptation of Tretter, Jones, and Taylor's SOQ and SAO activities[33]. Preparatory and general chemistry (novices) and chemistry graduate (experienced) students participated in a one-on-one interview activity consisting of four parts: bin creation and item sort (part I), item ordering within bins (part II), item ordering with measurements (part III), and item ordering on a number line (part IV). Parts I-IV examined students understanding of relative scale and parts III-IV additionally examined absolute scaling. In part I students constructed bins to sort object cards by length. After creating the bins the students then sorted the cards into their bins. In part II students organized the cards in order from smallest to largest. In part III students handed back the initial cards and were given a new set of cards with identical objects along with the object length to sort. Part IV had students place the objects with measurements at their proper size on the number line. Results showed that novice students demonstrated a lower scale conception

[33] Gerlach, Trate, Blecking, Geissinger, and Murphy 2014b: 1526-1537

than experienced students in both relative and absolute scaling which is the same results found by Tretter et al[34,35].

Knowing that chemistry students were struggling with scale, Gerlach, et al. developed and tested two different assessments to measure scale at the class-wide level, the Scale Concept Inventory (SCI) and the Scale Literacy Skills Test (SLST)[36]. These measures were rigorously tested for reliability and validity with interviews, trial testing, and content validation by experts in the field. The combined average of the SCI and SLST generates a student's Scale Literacy Score (SLS).

Jaclyn Trate expanded upon Gerlach's work by developing multiple regression models for general chemistry final exams[37]. When the scale measures, as well as traditional course measures such as ACT and sub-scores, were correlated with two ACS exams the students take as a final, the scale measures correlated similarly to, or better than, the traditional measures. In the final models for general chemistry I, scale was a greater predictor than a student's ACT composite score. These models are one of the ways to determine whether integrating scale as a theme in the course affected students' final exam performance.

The multiple regression models provided a way to measure learning gains in the course as scale was systematically integrated into a general chemistry I lecture, active learning, laboratory, and online supplemental instruction (SI) activities. In general chemistry I students demonstrated content learning gains with the integration of four aspects of scale into the course. These gains were seen over multiple semesters of testing in general chemistry I[38].

---

[34] Tretter 2006a: 282-319
[35] Tretter 2006b: 1061-1085
[36] Gerlach, Trate, Blecking, Geissinger, and Murphy 2014a: 1538-1545
[37] Trate 2017: 17-33
[38] Trate 2017: iii

The research questions this thesis focuses on are divided by chapter. For supplemental instruction, does supplemental instruction support student learning of their course content through use of scale as a framework? In order to do this, we need to know at what level of scale understanding the students start at, build a predictive model to use to predict their score without scale interventions (with the hope that scale interventions would have students score higher on the predicted measure), and integrate scale interventions such as supplemental instruction activities built with a scale framework.

# 1.3 Citations


Bybee, R. W. (1995). Science Curriculum Reform in the United States. In Bybee, R.W. & McInerney, J.D. (Eds.), *Redesigning the Science Curriculum: Report on the Implications of Standards and Benchmarks for Science Education* (p. 12-13). Colorado Springs, CO. BSCS.

Chi, M. T. H.; Feltovich, P.J.; and Glaser, R. (1981) *Cognitive Science*, *5*, 121-152.

Gerlach, K.; Trate, J.; Blecking, A.; Geissinger, P.; Murphy, K. (2014a). Valid and Reliable Assessments to Measure Scale Literacy of Students in Introductory College Chemistry Courses. *Journal of Chemical Education*. *91*, 1538-1545.

Gerlach, K; Trate, J; Blecking, A; Geissinger, P; Murphy, K. (2014b). Investigation of Absolute and Relative Scaling Conceptions of Students in Introductory College Chemistry Courses. *Journal of Chemical Education. 91*, 1526-1537

Golledge, R. G.; Gale, N.; Pelligrino, J. W.; Doherty, S. (1992). Spatial Knowledge Acquisition by Children: Route Learning and Relational Distances. *Annals of the Association of American Geographers. 82*, 223-244.

Jones, M. G. & Taylor, A. R. (2009). Developing a sense of scale: Looking backward. *Journal of Research in Science Teaching*, *46*, 460-475

Lock, G. & Molyneaux, B. L. (Eds.). (2006). *Confronting scale in archeology*. New York, NY: Springer Science & Business Media, LLC.

Trate, J. (2017). *Integrating Scale-Themed Instruction Across the General Chemistry Curriculum and Selected In-Depth Studies* (Doctoral dissertation). University of Wisconsin-Milwaukee, Milwaukee, WI.

Trend, R. D. (2001). Deep time framework: A preliminary study of U.K. primary teachers' conceptions of geological time and perceptions of geoscience. *Journal of Research in Science Teaching. 38*, 191-221.

Tretter, T. R. (2003). A sense of scale. *The Science Teacher*, 70(1). 22-25.

Tretter, T. R.; Jones, M. G; Andre, T.; Negishi, A; and Minogue, J. (2006a) Conceptual Boundaries and Distances: Students' and Experts' Conceptions of the Scale of Scientific Phenomena. *Journal of Research in Science Teaching*, *43*(3), 282-319.

Tretter, T. R.; Jones, M. G.; and Minogue, J. (2006b) Accuracy of Scale Conceptions in Science: Mental Maneuverings across Many Orders of Spatial Magnitude. *Journal of Research in Science Teaching*, *43*(10), 1061-1085.

Scale. (n. d.). In *Oxford Dictionary online.* Retrieved from https://en.oxforddictionaries.com/definition/scale on 10/20/2018.

Swarat, S.; Light, G.; Park, E. J.; Drane, D. (2011). A Typology of Undergraduate students' conceptions of size and scale: Identifying and characterizing conceptual variation. *Journal of Research in Science Teaching 48*, 512-533.

# Chapter 2: General Statistics

## 2.1 Introduction

The general statistics chapter is broken down into general methods, courses of interest, and data cleaning. The courses of interest are General Chemistry II and Anatomy and Physiology I with the rest of the courses of interest section containing descriptive statistics for the courses. Specific methods can be found in each chapter: supplemental instruction activities methods are in **Chapter 3**, Anatomy and Physiology I scale activity interview methods are in **Chapter 4**, and the Anatomy and Physiology I multiple regression model methods are in **Chapter 5**.

## 2.2 General methods

This research was conducted at a large public, doctoral, R1 research university in the Midwest. The academic calendar followed two 16-week semesters in fall and spring. Courses are available over winter (2-week session) and summer (4-, 6-, 8-, and 12-week sessions available) break. The university has approximately 21,000 undergraduate students. The student population is 48% male and 52% female.

All data reported here were obtained via IRB approval # 14.404.

All statistical analyses presented in this work were performed using IBM®SPSS Statistics®.

## 2.3 Courses of interest

### 2.3.1 General Chemistry II

General Chemistry II is a five-credit sixteen-week course with lecture, laboratory, and discussion taken primarily by science majors, engineering majors, and students from the College of Health Sciences. The course consists of three 50-minute lectures, a three-hour laboratory, and a 50-minute discussion per week. The course instructor teaches the lecture portion of the course and the laboratory and discussion sections are led by teaching assistants. The course prerequisite set by the university includes earning a letter grade of C or better in General Chemistry I, or a score of 4 or greater on the AP® Chemistry exam.

General Chemistry II instruction begins with a review of intermolecular forces and ends with electrochemistry. In total the course covers 8 chapters covering the topics of:

- Solutions
- Colligative properties and kinetics
- Mechanisms and catalysts
- Equilibrium
- Acids and bases
- Buffers and solubility equilibria
- Enthalpy and entropy
- Spontaneity and Gibbs Free Energy
- Redox reactions and cell potentials
- Corrosion, and batteries

Seventy-five percent of a student's course grade comes from performance on four hourly exams, lecture assignments, and assessments including: online homework, weekly in-class quizzes, and two nationally standardized final exams. Exams alone contribute 62.5% of a student's course grade. Weekly laboratory quizzes, laboratory reports, and an end of semester laboratory practical contribute to 18.75% of a student's course grade. Discussion accounts for the remaining 6.25% of the student's course grade.

The university institutional research data collected for General Chemistry II participants including sex and ACT composite (ACT COMP) score, ACT reading (ACT READ), ACT English (ACT ENGL), ACT mathematics (ACT MATH), and ACT science and reasoning (ACT SCIRE), for students who have an ACT score and IRB consented are listed in **Table 2.1**.

**Table 2.1 General Chemistry II descriptive statistics for ACT composite and sub-scores for fall 2017 and spring 2018 semesters**

| | Male | Female | ACT COMP | ACT READ | ACT ENGL | ACT MATH | ACT SCIRE |
|---|---|---|---|---|---|---|---|
| *n* | *149* | *166* | *315* | *315* | *315* | *315* | *315* |
| Minimum | | | 15 | 12 | 12 | 15 | 13 |
| Maximum | | | 34 | 35 | 35 | 33 | 35 |
| Mean | | | 23.84 | 24.22 | 23.23 | 23.37 | 24.04 |
| Median | | | 23 | 24 | 23 | 24 | 24 |
| Mode | | | 23 | 23 | 21 | 24 | 24 |
| Std. Deviation | | | 3.7 | 4.9 | 4.7 | 3.9 | 3.9 |
| Skewness | | | 0.132 | 0.199 | 0.212 | -0.253 | 0.205 |
| Kurtosis | | | -0.472 | -0.686 | -0.024 | -0.507 | 0.442 |

At the beginning of the semester General Chemistry II students complete the Scale Literacy Skills Test (SLST) and the Scale Concept Inventory (SCI)[39]. The SLST is 45 multiple-choice items assessing student scale skills. The SLST is administered online, via the course management site (D2L), at the beginning and end of a semester with students receiving extra

[39] Gerlach 2014a: 1538-1545

credit for its completion. The SLST is scored based on the total correct answers out of the total number of items.

The SCI deals with misconceptions regarding scale. The SCI is administered online at the beginning and end of a semester with students receiving extra credit for its completion. The SCI consists of 40 items each on a 5-point Likert scale containing objective items, subjective items and a verification item. The SCI has both positive statements (questions developed to evoke a positive response such as strongly agree) and negative statements (questions developed to evoke a negative response such as strongly disagree). A student's SCI score does not include responses to the subjective items or the verification item.

The SLST and SCI are complementary to each other by assessing different areas of student scale ability. The combined average of the SLST and SCI generates a Scale Literacy Score (SLS) for a student. The SLST and SCI items are available upon request. The descriptive statistics of the SLST is in **Table 2.2**, SCI in **Table 2.3**, and SLS in **Table 2.4**.

**Table 2.2 Scale Literacy Skills Test descriptive statistics for General Chemistry II for fall 2017 and spring 2018 semesters**

| | Scale Literacy Skills Test score pre | Scale Literacy Skills Test score post |
|---|---|---|
| *n* | *327* | *206* |
| Minimum | 0.200 | 0.156 |
| Maximum | 0.978 | 0.956 |
| Mean | 0.617 | 0.592 |
| Median | 0.644 | 0.600 |
| Mode | 0.644 | 0.600 |
| Std. Deviation | 0.15 | 0.17 |
| Skewness | -0.173 | -0.115 |
| Kurtosis | -0.543 | -0.602 |

**Table 2.3 Scale Concept Inventory descriptive statistics for General Chemistry II for fall 2017 and spring 2018 semesters**

| | Scale Concept Inventory score pre | Concept Inventory score post |
|---|---|---|
| *n* | *270* | *178* |
| Minimum | 0.550 | 0.556 |
| Maximum | 0.917 | 0.917 |
| Mean | 0.687 | 0.679 |
| Median | 0.670 | 0.672 |
| Mode | 0.656 | 0.650 |
| Std. Deviation | 0.066 | 0.063 |
| Skewness | 0.954 | 1.121 |
| Kurtosis | 0.730 | 1.489 |

**Table 2.4 Scale Literacy Score descriptive statistics for General Chemistry II for fall 2017 and spring 2018 semesters**

| | Scale Literacy score pre | Post Scale Literacy score post |
|---|---|---|
| *n* | *236* | *153* |
| Minimum | 0.436 | 0.4361 |
| Maximum | 0.892 | 0.9139 |
| Mean | 0.661 | 0.6478 |
| Median | 0.6611 | 0.6361 |
| Mode | 0.6750 | 0.600, 0.614, 0.628, 0.700, and 0.761 |
| Std. Deviation | 0.094 | 0.10 |
| Skewness | 0.128 | 0.285 |
| Kurtosis | -0.464 | -0.314 |

During the first week of the semester, the General Chemistry II course instructor administers the *ACS Exams 2005 First Term General Chemistry Paired Questions Exam* as a secure, low-stakes diagnostic test. The diagnostic test is used as the first part of their final exam taken at the end of the semester. The second part of their final exam is the *ACS Exams 2008 General Chemistry Conceptual Exam – Second Term*. The descriptive statistics for the diagnostic test and final exams are in **Table 2.5**.

**Table 2.5 Placement exam and final exams descriptive statistics for General Chemistry II for fall 2017 and spring 2018 semesters**

| | Placement Exam | Paired Final Exam | Conceptual Final Exam |
|---|---|---|---|
| *n* | *376* | *333* | *333* |
| Minimum | 0.125 | 0.250 | 0.200 |
| Maximum | 0.950 | 1.000 | 0.900 |
| Mean | 0.603 | 0.710 | 0.518 |
| Median | 0.6125 | 0.725 | 0.500 |
| Mode | 0.650 | .700 and .800 | 0.500 |
| Std. Deviation | 0.15 | 0.15 | 0.15 |
| Skewness | -0.248 | -0.462 | 0.255 |
| Kurtosis | -0.188 | -0.164 | -0.488 |

## 2.3.2 Anatomy and Physiology I

Anatomy and Physiology I is a four-credit sixteen-week course with lecture and laboratory with no university prerequisites. The large-enrollment lecture is two and a half hours a week either divided into three 50-minute lectures or two 75-minute lectures. Each three-hour laboratory section is taught by a teaching assistant once a week. The course is taken primarily by students with nursing (19.46%), biomedical sciences (16.21%), undecided (11.89%), and kinesiology (10.00%) intended majors. Other majors account for 5% or less of the students in the course with intended majors ranging from art history to mechanical engineering, see **Appendix A** for the distribution of majors in Anatomy and Physiology I. The course typically consists of 29% male students and 71% female students.

A student's course grade is determined by the in-class activities, online assessments, take-home exams, and laboratory. In-class activities including attendance, participation, and worksheets account for 20% of a student's course grade. Online assessments accounted for 20% of a student's course grade and include online quizzes, approximately two per week, and an online activity taken at the beginning of the semester. Three take-home exams and an online cumulative final exam accounts for 20% of a student's course grade. Each exam individually contributed to 5% of the course grade. The remaining 40% of a student's course grade is the laboratory component of the course including weekly lab worksheets, participation, a midterm laboratory practical and an end of the semester laboratory practical. Extra credit was given for completing the scale assessments at the beginning and end of the semester.

The university institutional research data collected for Anatomy and Physiology I participants' information, including sex and ACT composite score and sub-scores are listed in **Table 2.6**.

**Table 2.6 Anatomy and Physiology I descriptive statistics for ACT composite and sub-scores for fall 2017 and spring 2018 semesters**

| | Male | Female | ACT COMP | ACT READ | ACT ENGL | ACT MATH | ACT SCIRE |
|---|---|---|---|---|---|---|---|
| *n* | *180* | *445* | *625* | *625* | *625* | *625* | *625* |
| Minimum | | | 13 | 11 | 7 | 12 | 12 |
| Maximum | | | 34 | 36 | 35 | 34 | 34 |
| Mean | | | 22.23 | 22.53 | 21.73 | 21.82 | 22.30 |
| Median | | | 22 | 22 | 22 | 22 | 22 |
| Mode | | | 20 and 22 | 22 | 21 | 24 | 21 |
| Std. Deviation | | | 3.5 | 4.7 | 4.5 | 4.0 | 3.6 |
| Skewness | | | 0.244 | 0.367 | 0.158 | 0.117 | 0.343 |
| Kurtosis | | | -0.179 | -0.245 | 0.097 | -0.751 | 0.483 |

The university institutional research data also collected math placement scores and sub-scores for students. This information was not included in the previous table because of the drastic difference in sample size. While the math placement exam is standardized and the same exam is administrated by any university in the system, the method of storing and reporting a student's math placement scores varies dependent on the university at which the exam was administered and not necessarily on which university the student is enrolled (see **Table 2.7** for descriptive statistics). Prior to spring, 2017, the math placement exam consists of three sections:

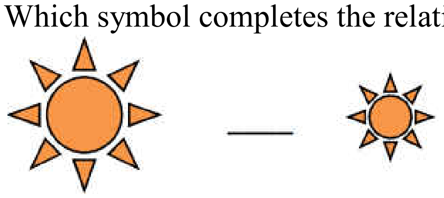
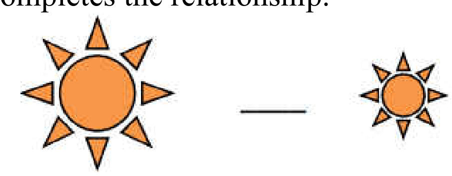
algebra (ALG), trigonometry (TRG), and math basics (MBSC). The items within each section are different every year. Starting in spring, 2017, the math placement exam sub sections changed from ALG, TRG, and MBSC to math fundamentals (MFUND), advanced algebra (AALG), and trigonometry and analytic geometry (TAG). The sections are ranked in difficulty with MFUND being the lowest and TAG being the highest. Each section of the math placement exam is scored and converted separately to a normalized score, using a conversion table, with values between 150 and 850 for each section. The combination of the sections of the math placement exam determine into which math class a student may enroll via a nominal code. Descriptive statistics for the math placement sections are given in **Table 2.7**.

**Table 2.7 Anatomy and Physiology I math placement sections scores descriptive statistics for fall 2017 and spring 2018 semesters**

| | Male | Female | algebra | trigonometry | math basics |
|---|---|---|---|---|---|
| *n* | *94* | *202* | *296* | *296* | *296* |
| Minimum | | | 150 | 150 | 150 |
| Maximum | | | 850 | 820 | 850 |
| Mean | | | 462.33 | 452.09 | 489.36 |
| Median | | | 450 | 420 | 480 |
| Mode | | | 440 | 420 | 430 |
| Std. Deviation | | | 110 | 120 | 130 |
| Skewness | | | 0.583 | 0.362 | 0.137 |
| Kurtosis | | | 0.790 | 0.518 | 0.095 |

When transitioning to Anatomy and Physiology I all existing scale instruments developed in chemistry were retested for reliability and validity. In spring, 2016, the SLST and SCI were administered only at the end of the semester in Anatomy and Physiology I. In summer 2016, domain experts (biological science professors) received copies of the assessments along with the answers. Two biological science professors commented on existing SLST and SCI items and suggested changes. The results from spring and summer 2016 administrations resulted in changes to the SLST and SCI for Anatomy and Physiology I. One of the changes that was made to all assessments was adjusting statements to make them more domain specific. For example, the statement cell, virus, or bacteria was changed to average cell, virus, or bacteria size. In total six items were changed. Four items were changed on the SLST, and two items were changed on the SCI. These changes were implemented starting in fall of 2016. Additional examples are listed in **Table 2.8** with two examples of changes made to the SLST and one change to the SCI. The full Anatomy and Physiology I version of the SLST and SCI are available upon request. Item statistics are given in **Appendix B**.

**Table 2.8 Examples of SLST and SCI item changes from General Chemistry to Anatomy and Physiology**

| **General Chemistry-SLST** | **Anatomy and Physiology I-SLST** |
|---|---|
| 22. Between a cell, a bacterium and a virus, which if any is the smallest? | 22. Considering their average sizes, which if any of the following is the smallest: a cell, a bacterium and a virus? |
| 24. Which symbol completes the relationship?  | 24. Fill in the blank with the symbol that completes the relationship.  |
| **General Chemistry-SCI** | **Anatomy and Physiology I-SCI** |
| 16. Magnifying a virus 100 times will not make it visible to the unaided human eye | 16. Magnifying an average virus 100 times will not make it visible to the unaided human eye |

During fall 2016 a response process validity study was conducted for the SLST with 20 students currently in Anatomy and Physiology I[40]. The study showed that 4 items posed a threat to the validity of the SLST. These items were removed when scoring the SLST creating an SLST adjusted score (SLST adj). Because the SLST directly contributes to the SLS, students

[40] Trate, Fisher, Geissinger, Blecking, and Murphy 2018

also have a SLS adjusted score (SLS adj). Descriptive statistics for the SLST pre, SCI pre, SLS pre, and SLST pre adjusted and SLS pre adjusted are given in **Table 2.9**.

**Table 2.9 Anatomy and Physiology I scale measure descriptive statistics for fall 2017 and spring 2018 semesters**

| | Scale Literacy Skills Test pre | Scale Concept Inventory pre | Scale Literacy Score pre | Scale Literacy Skills Test pre adjusted | Scale Literacy Score pre adj |
|---|---|---|---|---|---|
| *n* | *391* | *375* | *246* | *391* | *246* |
| Minimum | 0.089 | 0.522 | 0.339 | 0.098 | 0.334 |
| Maximum | 0.911 | 0.911 | 0.833 | 0.902 | 0.825 |
| Mean | 0.430 | 0.642 | 0.545 | 0.421 | 0.540 |
| Median | 0.400 | 0.633 | 0.529 | 0.390 | 0.522 |
| Mode | 0.378 | 0.611 | 0.472 | 0.366 | .416, 0.442, 0.452, 0.502, 0.505, 0.506, 0.645[a] |
| Std. Deviation | 0.15 | 0.051 | 0.091 | 0.15 | 0.091 |
| Skewness | 0.687 | 1.270 | 0.808 | 0.644 | 0.778 |
| Kurtosis | 0.182 | 3.265 | 0.576 | 0.119 | 0.491 |

[a]Three individuals for each mode

The laboratory survey is a survey about techniques or practices students were taught in the laboratory[41]. The laboratory survey was distributed and collected by the laboratory TAs during the first and last laboratory periods of the semester. Each item on the 20-item laboratory survey was on a 5-point Likert scale from strongly agree to strongly disagree. Subjective items, objective items, and a verification item were included on the survey.

The laboratory survey used in General Chemistry laboratory was adjusted for Anatomy and Physiology I to accommodate for domain specific knowledge and make the survey applicable for experiments performed in the Anatomy and Physiology I laboratory. Items were edited, deleted, or created to make the survey more relevant to Anatomy and Physiology I students. Seven items were identified as needing to be edited. Two ways questions were edited

[41] Trate 2017: 36-37

were to adjust the wording for the course and to adjust the question to cover similar content but using information taught in the Anatomy and Physiology I laboratory. Examples are in **Table 2.10**. Five items were deleted from the existing General Chemistry laboratory survey items and five new items were created for the Anatomy and Physiology version to make the survey applicable to the techniques used in the Anatomy and Physiology I laboratory. An example is in **Table 2.10**. The laboratory survey was not used in the work presented in this thesis, but data was collected.

**Table 2.10 Examples of survey statement changes from General Chemistry to Anatomy and Physiology I**

| Changes | General Chemistry | Anatomy and Physiology I |
|---|---|---|
| Wording for the course | 1 – I expect the lab will help reinforce the chemistry concepts taught in lecture. | 1 – I expect the lab will help reinforce the concepts taught in lecture. |
| Technique change, content same | 17 – Overfilling a volumetric flask while making a solution would result in a higher calculated concentration. | 17 – Adding more water while making a solution would result in a higher calculated concentration. |
| New Item | 20 – Using a volumetric flask instead of an Erlenmeyer flask to make a solution will make the measurement more precise. | 20 – Microscopes are used to view features that are not visible to the naked eye. |

At the end of the semester Anatomy and Physiology I students take a cumulative final exam. The exam is administered online through the book publisher website (McGraw Hill). They have one two-hour attempt to complete the 97-item exam during a one-week time frame beginning the last day of the course. The descriptive statistics for the cumulative final exam are in **Table 2.11**.

**Table 2.11 Cumulative final exam percent descriptive statistics for Anatomy and Physiology I for fall 2017 and spring 2018 semesters**

| | Male | Female | Cumulative final exam percent |
|---|---|---|---|
| *n* | *172* | *444* | *616* |
| Minimum | | | 0.068 |
| Maximum | | | 0.925 |
| Mean | | | 0.663 |
| Median | | | 0.680 |
| Mode | | | 0.631, 0.652, 0.693, 0.699, 0.708, 0.713, 0.734, 0.747, 0.846[a] |
| Std. Deviation | | | 0.12 |
| Skewness | | | -0.678 |
| Kurtosis | | | 0.789 |

[a]Two individuals for each mode

# 2.4 Cleaning data sets

A verification item was used on the SCI, and students failing the verification item were removed. The verification item was written to elicit a positive response. Students failed the verification item by selecting a neutral or negative response. Students who completed the instruments in less time than reading each statement on the SCI would take were removed (less than 3 minutes). For both the SCI and SLST students who had a variance of 0 were removed. Questions left blank on the laboratory survey were reverse coded. Students who did not take the final exam were removed for not completing the course.

To be included in the analysis students had to have an ACT composite and sub-scores, beginning of semester scale measures, and take the final exam or have an ACT composite and sub-scores, math placement and sub-scores, and beginning of semester scale measures depending on the analysis.

## 2.5 Citations


Gerlach, K.; Trate, J.; Blecking, A.; Geissinger, P.; Murphy, K. (2014). Valid and Reliable Assessments to Measure Scale Literacy of Students in Introductory College Chemistry Courses. *Journal of Chemical Education*. *91*, 1538-1545.

Trate, J. (2017). *Integrating Scale-Themed Instruction Across the General Chemistry Curriculum and Selected In-Depth Studies* (Doctoral dissertation). University of Wisconsin-Milwaukee, Milwaukee, WI.

Trate, J.M.; Fisher, V. K.; Geissinger, P.; Blecking, A.; Murphy, K.L. "Response Process Validity Studies of the Scale Literacy Skills Test" in revision *Journal of Chemical Education*, 2018.

# Chapter 3: Supplemental instruction in a General Chemistry II course

## 3.1 Introduction

Supplemental instruction is an academic support model developed for students for a topic or course, for example chemistry. Adaptive learning is the use of technology to provide a more individual experience to students for a topic. Supplemental instruction can be created using technology in order to make an adaptive learning supplemental instruction for students to receive more targeted instruction in a topic, such as self-efficacy or multiplication.

The purpose of supplemental instruction and adaptive learning is to support student learning of the course content. One way to frame supplemental instruction is by integrating a theme such as scale or models. A theme can be integrated explicitly by making connections between different areas of the course content and the particular theme in instruction. NGSS has seven cross-cutting concepts one of which is "scale, proportion, and quantity"[42]. Scale was found to be one of the lesser studied cross-cutting concepts and many instructors who cover "scale" may not cover the entire breadth of the cross-cutting concept.

Investigation into General Chemistry I revealed that chemistry students struggle with scale[43]. Instruments were developed to measure student scale ability[44]. Where to use scale as a theme in a General Chemistry I course was determined through comparison of student scaling ability, hourly exams, and course content. Results from a different study showed that the areas in the course that would have the greatest benefit of integration would be in lecture and laboratory. Scale was integrated into those points in the form of active learning, reworking the

---

[42] National Research Council 2013
[43] Gerlach 2014b: 1526-1537
[44] Gerlach 2014a: 1538-1545

experiments, and pre-lab quizzes. Integration of scale as a theme in General Chemistry I lecture, laboratory, and supplemental instruction impacted students understanding of chemistry. An increase in student performance occurred on the final exam for semesters with scale integration. Scale has been systematically integrated as a theme in lecture and laboratory in General Chemistry II. This chapter details the development, implementation, and results of supplemental instruction being integrated into General Chemistry II at two time-points during the semester.

# 3.2 Background

## 3.2.1 Adaptive learning systems

A version of adaptive learning made its first appearance in an experiment by Sidney L. Pressley to present a stimulus, adapt to a response, and provide reinforcement based on the response[45]. While this was progress for technology, building the machine, and teaching, the ability for a student to progress at their own pace, Skinner opposed Pressley's learning machine because Skinner claimed the machine recorded how students performed and allowed them to take their time but did not actively participate in teaching the student new information[46]. In order to be considered a teaching tool the machine should be built with a theoretical basis and teach the students information.

Skinner built his own machine with James G. Holland based on the idea that animals can learn behavior from reinforcement[47]. This machine was programmed with a course textbook that students would spend an average of 15 hours working through. As technology advanced full

---

[45] Stolurow and Davis 1965: 162-212
[46] Kara and Sevim 2013: 108-120
[47] Skinner 1960: 189-191

machines no longer needed to be devoted to teaching and instead technology could be used to build adaptive learning systems with existing machines such as computers.

Chieu defines adaptability as "the ability of a learning system to provide each learner with appropriate learning conditions to facilitate his or her own process of knowledge construction and transformation"[48]. Chieu gives five techniques or ways for learning systems to be adaptive: presentation of learning contents, pedagogical devices, communication support, problem-solving support, and assessment. Adaptability in the presentation of learning contents allow students to open a new sequence of web pages if a student has "proven" to the system, usually by answering questions, that they have an adequate amount of knowledge. Pedagogical devices are a means to support student learning. Adaptability in pedagogical devices allows students to be supported in the way that would benefit the most by providing a more individualized approach such as students receiving different levels of instruction about a topic depending on their current knowledge level of the topic. Adaptability in communication support allows students who are struggling to contact peers. The system provides the student with a list of peers who appear to have mastered the concept and the student is able to select one or more students to contact. Adaptability in problem-solving provides the support students who are struggling need to learn the concept. Assessment adaptability allows students at different learning levels to be graded at their current level. An example is if a project is due the system would choose different projects for students based on their knowledge of the topic that was demonstrated previously in questions or other assignments.

An example of an adaptive learning system is online flashcards[49]. As a student gets the answer correct additional or new flashcards are shown. If a student gets a card incorrect the card

---

[48] Chieu 2005: 70-96
[49] Kerr 2015: 88-93

will show up again until the student gets the card correct a set number of times. The technology, for example a computer or website, determines the order of the flashcards, frequency, etc. while an instructor, or a student, selects the topic of instruction. Flashcards sets can be written for a specific topic or theme.

In 2017 the Australian Government Department of Education and Training published a report about assessing an online adaptive tool in a large undergraduate first-year psychology course[50]. Students had access to the LearnSmart tool in two psychology courses. In one course (Course A) LearnSmart was recommended and in Course B LearnSmart was required. LearnSmart was an adaptive tool that adjusted "the difficulty of the assessment to suit the understanding of individual students". LearnSmart usage was found to be the most significant predictor of the end-of-semester exam performance for both courses. Similar results was found for both courses despite the different motivations for students to use the adaptive tool.

### 3.2.2 Learning theories

Jean Piaget spent his life studying the psychology of children. His research, and the idea that humans have the ability to do "abstract symbolic reasoning" where animals do not, led him to develop the theory of cognitive development[51]. Piaget realized that at different points in a human lifespan, humans think qualitatively different than in previous stages. Piaget's cognitive theory can be broken down into two main parts: schemas and cognitive developmental stages.

Schemas are "organized packets of knowledge" located in the long-term memory[52]. A schema is a mental concept that helps a person know what to expect from a variety of situations, for example what to attend to during a conversation or lecture. These packets of knowledge are

---

[50] Dry 2017
[51] Piaget and Cook 1952
[52] Eysenck 2012: 159

linked to form an intricate web of information and connections. Because of these connections schemas affect how new knowledge is processed and stored. As a person is presented with new information Piaget describes one of two things happen: assimilation or accommodation[53]. Assimilation happens when the new information is integrated into the existing schema, such as a child's schema of a tree may be brown with green leaves but as the child experiences different types of trees or trees during different seasons (such as without leaves) the schema of a tree is enriched. Accommodation is when the schema is changed to accommodate the new information such as a child seeing a donkey for the first time may say it fits their current schema for a horse. As the information about the donkey is learned the child's schema adapts to incorporate the new information and separate donkey from horse.

As schemas are developed and undergo the process of assimilation and accommodation they become more complex. This increasing complexity of cognitive thinking leads to the development of stages of cognitive development[54]:

- Sensorimotor stage (birth to age 2; infancy)
- Pre-operational stage (from 2 to 7; toddler and early childhood)
- Concrete operational stage (from 7 to 11; elementary and early adolescence)
- Formal operational stage (11+; adolescence and adulthood)

In the sensorimotor stage intelligence is gained through physical experiences. As mobility develops more intelligence can be gained. The major achievement at this level is object permanence. In the pre-operational stage the use of symbols, language, memory, and imagination are developed but thinking is egocentric and not logical. The concrete operational stage is what Piaget considered the beginning of where logical thought begins to happen. For

---

[53] Piaget and Cook 1952). *The origins of intelligence in children*. New York, NY: International University Press.
[54] Piaget and Cook 1952). *The origins of intelligence in children*. New York, NY: International University Press.

number, mass, and weight, for example, conservation happens and manipulation of symbols that refer to concrete objects occurs. In the formal operational stage, the ability to logically test hypotheses and conceive abstract concepts is formed.

Zone proximal development (ZPD) measures the difference between what a learner is able to do by themselves and what a learner cannot do[55]. The area between the two is what a student is able to do with guidance from an expert. The concept was introduced by Lev Vygotsky. Providing learners with guidance in the zone of proximal development provides support so that the student is able to complete the task and thereby help advance the learner's skills. Once a task is mastered by the learner, the task becomes part of the area that a learner is able to do by themselves[56].

Traditional chemistry instruction involves lecturing to students about specific reactions or experiments while they take notes. Johnstone and others looked for a new way of teaching chemistry that would focus on larger topics with more emphasis of the students making discoveries about their chemistry understanding. Johnstone was on the forefront in incorporating educational psychology and learning theory into chemistry instruction. He developed a representational framework focusing on three main components: macroscopic, symbolic, and microscopic, see **Figure 3.1** [57].

[55] Warford 2011: 1-12
[56] Siyepu 2013: 1-13.
[57] Johnstone, A. H. (1993). The development of chemistry teaching: A changing response to changing demand. *Journal of Chemical Education, 70*(9), 701-705.

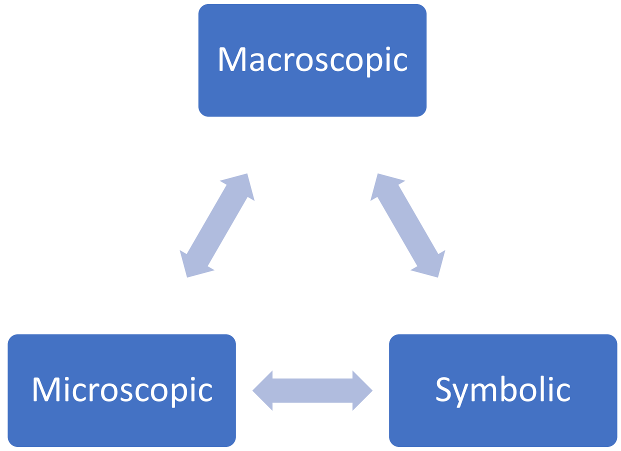


*Figure 3.1 Three components of "new chemistry" recreated from The Development of Chemistry Teaching*[58].

Experts can move between these representations easily while novices have difficulties[59]. Meaningful learning happens when a student understands a topic and all the pieces that fit together within that topic. Applying Johnstone's three components of chemistry, a student would need to master the macroscopic, symbolic, and microscopic levels of a topic.

The main research question for this chapter is using scale and Johnstone's triangle as a framework, does the development of an adaptive online supplemental instruction aid students in understanding the topics of solutions and fuel cells from beginning the activity to completing the activity. Another research question is how does completion of both supplemental instruction activities impact student performance in the course.

# 3.3 Methods

The development of supplemental instruction activities for General Chemistry II was completed during the fall 2017 semester. Development and implementation of the activities began in the university's course management system (Desire2Learn) and has subsequently been

[58] Johnstone 1993: 701-705
[59] Johnstone 1993: 701-705

moved to a free-standing website (web.uwm.edu/scale/). The first activity was completed at the beginning of the semester, after the first textbook chapter had been taught in lecture. The second activity was completed at the end of the semester.

In continuation of previous research in General Chemistry I and II, fall 2017 (semester 1) had active learning, and supplemental instruction. Spring 2018 (semester 2) had active learning with scale integrated as a theme, and supplemental instruction. Scale was not integrated as a theme in lecture or laboratory either semester. The difference in treatments between semesters was semester 1 was active learning control while semester 2 was scale active learning in class workbooks.

### 3.3.1 Content Selection

When determining the concepts addressed in the activities, the current lecture topics were taken into consideration along with determining whether the topic could easily be divided into levels of difficulty for scenario 1 (lowest level of difficulty), 2, and 3 (highest level of difficulty). When choosing the chemistry content for the scenarios a variety of criteria had to be met. The topic had to be relevant to lecture topics, easily and fluently transferred between the three representations, and relevant to themes of scale. In General Chemistry II supplemental instruction, each activity had an over-arching situation or experiment to link the scenarios together for those students who completed more than one scenario. However, each scenario would have to be stand-alone so that if a student placed in any scenario the content and fictionally posed situation was comprehensible. For example, if a student placed into scenario 2 they would not have to know details from scenario 1 other than those already provided in scenario 2.

At the beginning of the semester only one unit has been completed (chapter 13 of their textbook: solutions). Usability studies of active learning showed students held misconceptions with regards to solutions as well as understanding what a calculated number means within the solutions unit. Solutions lends itself to different representations from the macroscopic (e.g. a beaker), symbolic (e.g. chemical equations), and microscopic (e.g. particulate) level. Solution chemistry was selected as the topic for activity 1.

Activity 2 is completed at the end of the semester. The course topics covered at the end of the semester include enthalpy, entropy, spontaneity, and Gibbs free energy, redox and cell potentials, thermodynamics, and corrosion and batteries. A fuel cell activity can utilize all of these topics from information about a battery to energy calculations. Fuel cells also lends itself to different representations with macroscopic battery function, chemical and mathematical equations, and particle level redox reactions. Fuel cells was selected as the topic for activity 2.

### 3.3.2 Overview of activity

Supplemental instruction was developed to support students' understanding of two specific content areas of chemistry: solutions and fuel cells. The format of the supplemental instruction activities mirrored that of General Chemistry I. The adaptive activities were developed in the form of multiple quizzes that students have access to based on their performance. Each activity contained eight subsections: three scenarios, three post-scenario questions, and initial and final questions. All students complete the initial questions and, based on their score, are placed into either scenario 1, 2, or 3. Each scenario had to be fully contained so students would not need information from any previous scenario if they were placed in scenario 2 or 3 and were not required to complete scenario 1. Once the students complete the scenario they have access to the scenario questions. If a student receives a perfect score on the scenario 1 questions, they are moved on to scenario 3. If a student scores less than 100%, the student moves on to scenario 2 and then the scenario 2 questions. The student has completed the activity once they have finished the final questions. If a student failed to meet the minimum required score, they repeated the scenario or scenario questions until they met the minimum required score. **Figure 3.2** describes the paths through the activity the student may take.

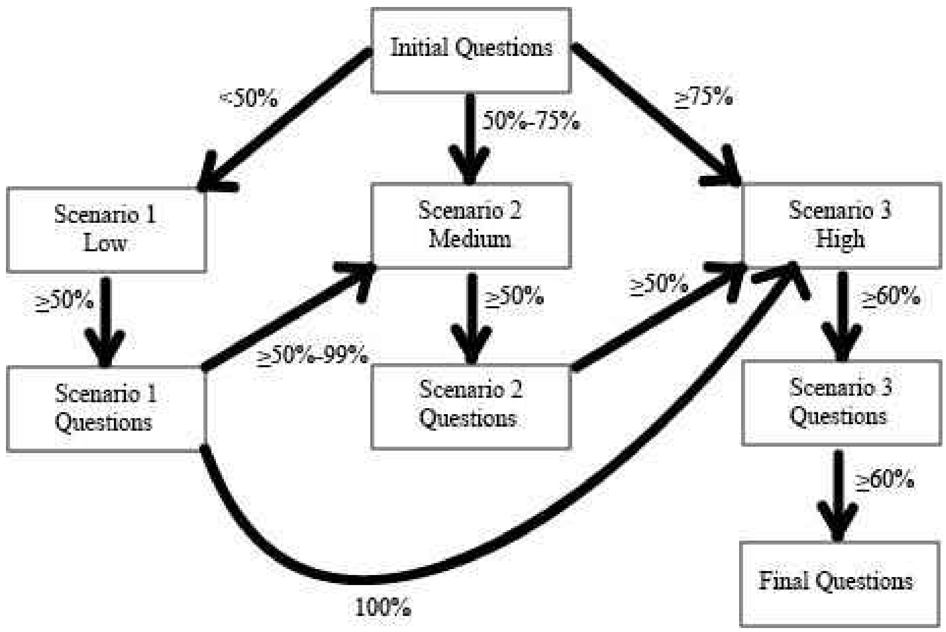


*Figure 3.2 Supplemental instruction pathways*

Once the scope of the scenario, content, and format for the scenarios was defined the scenarios were outlined. After the scenarios were outlined, scenario questions, and post scenario questions were written and vetted by experts. Each scenario was further broken down into sub-topics. For each sub-topic a database of questions was written, typically 5-8 questions, from which the initial, final, and post scenario questions were pulled from. The subcategories for the solutions activity, the number of questions in each pool, and how many questions are pulled from each pool are in **Table 3.1** and **Table 3.2** for the fuel cells activity. Questions were vetted by four chemistry experts who also wrote questions for the activities. Concepts within each scenario where students may struggle to answer the questions within the scenario were identified and hints were created to teach students the concept.  The design of the hints was general but specific examples were provided to assist students.  See **Appendix B** for the scenarios.

**Table 3.1 The solutions activity subcategories, number of multiple-choice questions in each pool, pulled for the initial and final questions, and pulled for each post scenario questions**

| Subcategory | Number of questions in pool | Number of questions pulled for initial/final questions | Number of questions pulled for scenario 1 questions | Number of questions pulled for scenario 2 questions | Number of questions pulled for scenario 3 questions |
|---|---|---|---|---|---|
| Heating and Cooling Curves | 5 | 1 | 1 | | |
| Intermolecular forces | 8 | 1 | 2 | | |
| Phase Change | 5 | 1 | 1 | | |
| Phase Diagram | 5 | 1 | 1 | | |
| Solution Amounts | 8 | 1 | | 2 | |
| Intermolecular forces in Solutions | 7 | 1 | | 3 | |
| Vapor Pressure Lowering | 5 | 1 | | | 2 |
| Boiling Point Elevation | 6 | 1 | | | 2 |
| Phase Diagrams of Solutions | 5 | 1 | | | 1 |
| Total | 54 | 9 | 5 | 5 | 5 |

**Table 3.2 The fuel cells activity subcategories, number of multiple-choice questions in each pool, pulled for the initial and final questions, and pulled for each post scenario questions**

| Subcategory | Number of questions in pool | Number of questions pulled for initial/final questions | Number of questions pulled for scenario 1 questions | Number of questions pulled for scenario 2 questions | Number of questions pulled for scenario 3 questions |
|---|---|---|---|---|---|
| Galvanic cell | 10 | 1 | 1 | | |
| Voltage/cell potential | 8 | 1 | 1 | | |
| System/surroundings | 8 | 1 | 1 | | |
| Gases-macroscopic | 13 | 1 | 2 | | |
| Symbolic reactions | 15 | 1 | | 1 | |
| Stoichiometry | 7 | | | 1 | |
| Nernst equation calculations | 8 | 1 | | 1 | |
| Spontaneity and temperature | 15 | 1 | | 1 | |
| Ideal gas law calculations | 5 | 1 | | 1 | |
| Particulate ideal gas law and kinetic energy | 5 | 1 | | | 1 |
| Energy diagrams | 12 | 1 | | | 1 |
| Particulate mechanism drawings | 4 | 1 | | | 1 |
| Mechanism of fuel cells | 4 | | | | 1 |
| Energy/bonds | 15 | 1 | | | 1 |
| Total: | 129 | 12 | 5 | 5 | 5 |

### 3.3.3 Data analysis

Semesters 1 and 2 were combined for analysis in how students utilized the supplemental instruction activities. The pathways students took to complete each activity as well as where students placed from the initial questions were examined. Descriptive statistics for each activity were provided along with analysis regarding initial and final questions for both activities. Pearson correlations were used to support student placement into scenario 1, 2, or 3 for each activity but due to different treatments the semesters were analyzed independently for the correlations.

Data analysis regarding supplemental instruction's impact on student learning was also performed. Due to different treatments the semesters were analyzed independently. Independent samples t-tests were run at the beginning and end of the semester to determine if the samples had significantly different means.

### 3.3.4 Data cleaning

Statistical analysis was performed to determine if semester 1 and 2 could be combined when analyzing how students utilized the supplemental instruction activities. Depending on the analysis for how students used the activities, students were excluded if they did not start the supplemental instruction activities or complete the supplemental instruction activities. Each activity was treated separately so students were not excluded if they did not complete both activities.

Semester 1 and semester 2 received different treatments. When looking at how the supplemental instruction activities impacted student performance the semesters must be treated differently. For how supplemental instruction activities impacted student performance, students were excluded if they did not have the beginning of the semester scale measures, ACT composite and sub-scores, and completed the final exam.

# 3.4 Results

## 3.4.1 Supplemental instruction activity 1 and 2 results

Ideally when comparing semester 1 and semester 2 there will be no significant difference between beginning of semester measures and supplemental instruction activity performance. Comparison of student performance between the two semesters of data showed no significant differences that exist between either groups for an independent t-test run for the initial questions, all scenarios, all scenario questions, and final questions. The only exception was activity 1 scenario 3 questions that had a significance at the 0.001 level. Independent sample t-tests were also conducted for ACT composite score and sub-scores, SLST, SCI, SLS, and placement exam to see if there was a difference between the two semesters and the results were not significant. Tables for the independent sample t-tests are included in **Table 3.3**. The data supports the case that the semesters are equivalent and semesters 1 and 2 may be used as a combined sample.

**Table 3.3 Independent t-tests for beginning of semester measures and supplemental instruction activities (semester 1 minus semester 2)**

| | Semester | *n* | Mean | Std. Dev. | t | df | Sig. (2-tailed) |
|---|---|---|---|---|---|---|---|
| ACT COMP | 1 | *140* | 24.207 | 3.6 | 1.583 | 313 | 0.114 |
| | 2 | *175* | 23.543 | 3.8 | | | |
| ACT READ | 1 | *140* | 24.600 | 4.7 | 1.226 | 313 | 0.221 |
| | 2 | *175* | 23.914 | 5.1 | | | |
| ACT ENGL | 1 | *140* | 23.536 | 4.4 | 1.043 | 313 | 0.298 |
| | 2 | *175* | 22.983 | 4.9 | | | |
| ACT MATH | 1 | *140* | 23.600 | 3.9 | 0.935 | 313 | 0.350 |
| | 2 | *175* | 23.189 | 3.9 | | | |
| ACT SCIRE | 1 | *140* | 24.521 | 3.8 | 1.960 | 313 | 0.051 |
| | 2 | *175* | 23.651 | 4.0 | | | |
| Beginning of semester SLST | 1 | *158* | 0.623 | 0.14 | 0.625 | 325 | 0.532 |
| | 2 | *169* | 0.612 | 0.16 | | | |
| Beginning of semester SCI | 1 | *117* | 0.682 | 0.061 | -1.139 | 268 | 0.256 |
| | 2 | *153* | 0.691 | 0.070 | | | |
| Beginning of semester SLS | 1 | *110* | 0.656 | 0.084 | -0.748 | 234 | 0.456 |
| | 2 | *126* | 0.665 | 0.10 | | | |
| Placement Test | 1 | *174* | 0.597 | 0.15 | -0.814 | 374 | 0.416 |
| | 2 | *202* | 0.610 | 0.16 | | | |
| Initial questions activity 1 | 1 | *125* | 0.593 | 0.19 | -0.321 | 276 | 0.748 |
| | 2 | *153* | 0.601 | 0.20 | | | |
| Scenario 1 activity 1 | 1 | *40* | 0.818 | 0.092 | 2.097 | 83 | 0.039 |
| | 2 | *45* | 0.769 | 0.12 | | | |
| Scenario 1 questions activity 1 | 1 | *39* | 0.733 | 0.16 | 1.400 | 79 | 0.166 |
| | 2 | *42* | 0.676 | 0.20 | | | |
| Scenario 2 activity 1 | 1 | *65* | 0.627 | 0.11 | -0.297 | 144 | 0.767 |
| | 2 | *81* | 0.633 | 0.12 | | | |
| Scenario 2 questions activity 1 | 1 | *56* | 0.711 | 0.16 | 1.158 | 128 | 0.249 |
| | 2 | *74* | 0.678 | 0.16 | | | |
| Scenario 3 activity 1 | 1 | *87* | 0.786 | 0.095 | -2.240 | 201 | 0.026 |
| | 2 | *116* | 0.821 | 0.12 | | | |
| Scenario 3 questions activity 1 | 1 | *83* | 0.699 | 0.16 | -3.359 | 192 | 0.001 |
| | 2 | *111* | 0.782 | 0.18 | | | |
| Final questions activity 1 | 1 | *78* | 0.649 | 0.19 | -1.367 | 182 | 0.173 |
| | 2 | *106* | 0.690 | 0.21 | | | |
| Initial questions activity 2 | 1 | *140* | 0.576 | 0.17 | 1.295 | 274 | 0.196 |
| | 2 | *136* | 0.547 | 0.19 | | | |
| Scenario 1 activity 2 | 1 | *46* | 0.643 | 0.12 | 0.318 | 102 | 0.751 |
| | 2 | *58* | 0.635 | 0.13 | | | |
| Scenario 1 questions activity 2 | 1 | *41* | 0.751 | 0.15 | 0.508 | 86 | 0.613 |
| | 2 | *47* | 0.732 | 0.20 | | | |
| Scenario 2 activity 2 | 1 | *90* | 0.601 | 0.17 | -1.007 | 173 | 0.315 |
| | 2 | *85* | 0.626 | 0.15 | | | |
| Scenario 2 questions activity 2 | 1 | *67* | 0.648 | 0.16 | -0.592 | 138 | 0.555 |
| | 2 | *73* | 0.663 | 0.14 | | | |
| Scenario 3 activity 2 | 1 | *84* | 0.749 | 0.10 | 0.147 | 175 | 0.883 |
| | 2 | *93* | 0.746 | 0.13 | | | |
| Scenario 3 questions activity 2 | 1 | *84* | 0.674 | 0.15 | 1.961 | 170 | 0.052 |
| | 2 | *88* | 0.625 | 0.17 | | | |
| Final questions activity 2 | 1 | *79* | 0.563 | 0.17 | 1.330 | 154 | 0.186 |
| | 2 | *77* | 0.524 | 0.20 | | | |

Based on their initial questions score students were placed into scenario 1, 2, or 3. The number of students that were placed into each scenario are in **Table 3.4**. **Table 3.5** for the solutions activity, and **Table 3.6** for the fuel cells activity provides the number of students who completed the activity and how many scenarios they completed. The number of students who started the solutions activity was 278 with 66.19% completing the activity. The number of students who started the fuel cells activity was 276 with 56.52% completing the activity.

**Table 3.4 Number of students placed into each scenario based on initial questions score for semester 1 and 2**

| | Scenario 1 | Scenario 2 | Scenario 3 |
|---|---|---|---|
| Activity 1 (n = 278) | 95 (34.17%) | 106 (38.13%) | 77 (27.70%) |
| Activity 2 (n = 276) | 128 (46.38%) | 84 (30.43%) | 64 (23.19%) |

**Table 3.5 Number of students who completed each path and the final questions for the solutions activity for semester 1 and 2**

| | |
|---|---|
| placed into scenario 1 and completed all 3 scenarios and finished | 54 |
| placed into scenario 1 and skipped to 3 and finished | 12 |
| those who placed into 2 and finished | 62 |
| those who placed into 3 and finished | 56 |

**Table 3.6 Number of students who completed each path and the final questions for the fuel cells activity for semester 1 and 2**

| | |
|---|---|
| placed into scenario 1 and completed all 3 scenarios and finished | 57 |
| placed into scenario 1 and skipped to 3 and finished | 8 |
| those who placed into 2 and finished | 67 |
| those who placed into 3 and finished | 24 |

**Table 3.7 Number of scenarios completed by students in the solutions activity and the fuel cells activity for semester 1 and 2**

| | Completed 1 scenario | Completed 2 Scenarios | Completed 3 scenarios |
|---|---|---|---|
| Activity 1 (n = 184) | 56 (30.43%) | 74 (40.22%) | 54 (29.35%) |
| Activity 2 (n = 156) | 24 (15.38%) | 75 (48.08%) | 57 (36.54%) |

Of those that completed the initial questions for the solutions activity, 66.19% completed the final questions for the solutions activity and 56.52% of those who completed the initial questions for the fuel cells activity completed the final questions for the fuel cells activity. On average for the solutions activity and the fuel cells activity, those who completed the final questions of an activity completed 2 scenarios per activity. For example, a student completed scenario 1 and 3 or a student who completed scenario 2 and 3. The mode for the average number of scenarios completed was also 2 for both activities. The descriptive statistics as well as the number of students placed in each scenario within each activity supports the grouping of students based on score. The average, median and mode, number of scenarios completed by each student was 2 with fewer students completing all 3 scenarios or only 1 scenario. **Table 3.8** shows the descriptive statistics for each section for the solutions activity and **Table 3.9** for the fuel cells activity.

**Table 3.8 Solutions activity descriptive statistics for semesters 1 and 2**

| | Initial questions | Scenario 1 | Scenario 1 questions | Scenario 2 | Scenario 2 questions | Scenario 3 | Scenario 3 questions | Final questions |
|---|---|---|---|---|---|---|---|---|
| *n* | *278* | *85* | *81* | *146* | *130* | *203* | *194* | *184* |
| Minimum | 0.111 | 0.500 | 0.20 | 0.000 | 0.20 | 0.5000 | 0.4 | 0.11110 |
| Maximum | 1.000 | 1.000 | 1.0 | 0.88890 | 1.0 | 1.0000 | 1.0 | 1.00000 |
| Mean | 0.597 | 0.7921 | 0.7037 | 0.6305 | 0.6923 | 0.8057 | 0.7464 | 0.6724 |
| Median | 0.556 | 0.8333 | 0.6000 | 0.6111 | 0.600 | 0.8125 | 0.8000 | 0.7000 |
| Mode | 0.444 and 0.556 | 0.8333 | 0.6000 | 0.6111 | 0.600 | 0.8125 | 0.6000 | 0.7778 |
| Std. Dev. | 0.20 | 0.11 | 0.18 | 0.12 | 0.1578 | 0.11 | 0.18 | 0.20 |
| Variance | 0.039 | 0.012 | 0.034 | 0.013 | 0.025 | 0.012 | 0.031 | 0.041 |
| Skewness | 0.018 | -0.629 | -0.349 | -0.854 | 0.225 | -0.474 | 0.178 | -0.454 |
| Kurtosis | -0.611 | 0.162 | 0.562 | 5.119 | 0.650 | -0.059 | -1.071 | -0.161 |

**Table 3.9 Fuel cells activity descriptive statistics for semesters 1 and 2**

| | Initial questions | Scenario 1 | Scenario 1 questions | Scenario 2 | Scenario 2 questions | Scenario 3 | Scenario 3 questions | Final questions |
|---|---|---|---|---|---|---|---|---|
| *n* | *276* | *104* | *88* | *175* | *140* | *177* | *172* | *156* |
| Minimum | 0.083 | 0.313 | 0.200 | 0.000 | 0.200 | 0.111 | 0.000 | 0.083 |
| Maximum | 1.000 | 0.938 | 1.000 | 0.929 | 1.000 | 0.944 | 1.000 | 0.917 |
| Mean | 0.562 | 0.638 | 0.741 | 0.613 | 0.656 | 0.747 | 0.649 | 0.544 |
| Median | 0.583 | 0.625 | 0.800 | 0.619 | 0.600 | 0.722 | 0.600 | 0.583 |
| Mode | 0.417 | 0.625 | 0.800 | 0.524 and 0.571 | 0.600 | 0.667 and 0.778 | 0.600 | 0.667 |
| Std. Dev. | 0.18 | 0.13 | 0.18 | 0.16 | 0.15 | 0.12 | 0.16 | 0.19 |
| Variance | 333.132 | 159.795 | 314.107 | 272.089 | 230.606 | 132.172 | 270.747 | 345.392 |
| Skewness | -0.126 | 0.043 | -0.392 | -0.901 | 0.083 | -0.727 | -0.417 | -0.211 |
| Kurtosis | -0.347 | 0.221 | 0.492 | 1.698 | 0.714 | 3.980 | 1.896 | -0.417 |

For the solutions activity a paired samples t-test showed that students scored significantly better on the final questions (M = 67.24% SD = 20.17%) compared to the initial questions (M = 60.27% SD = 20.40%) (t(183) = -4.119, $p < 0.001$).

For the fuel cells activity a paired samples t-test showed that students scored significantly better on the initial questions (M = 59.08% SD = 19.38%) compared to the final questions (M = 54.38% SD = 18.58%) (t(155) = 2.873, $p = 0.05$).

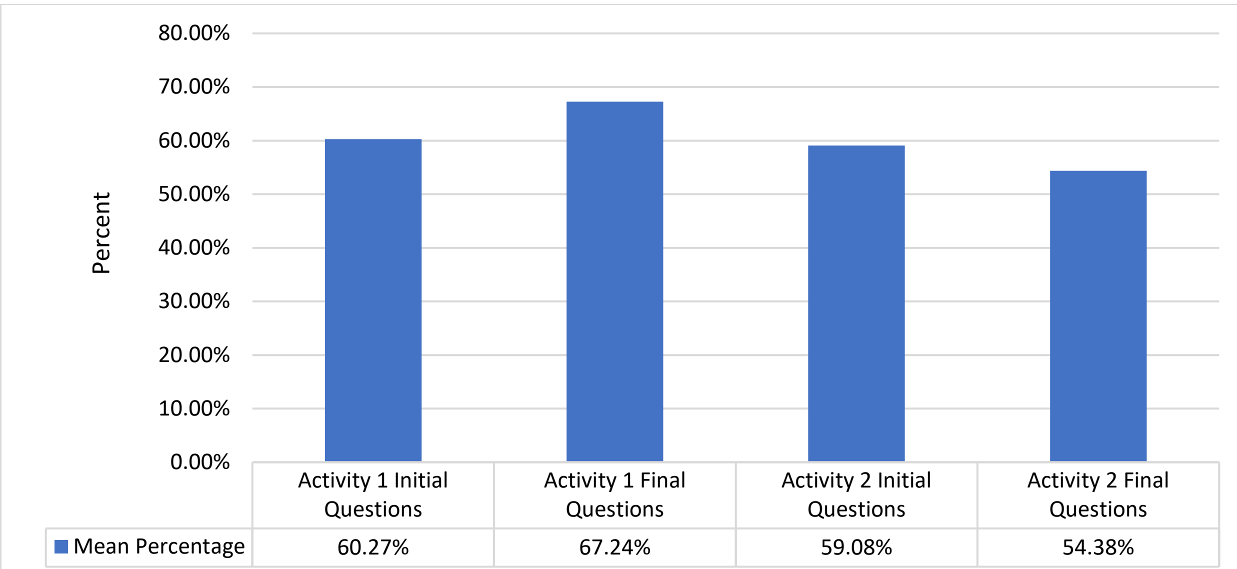



*Figure 3.3 Initial and final question means for the solutions activity and the fuel cells activity*

Pearson correlations were performed for each activity and semester compared to beginning of semester measures. For the fuel cells activity, students completed this activity towards the end of the semester so exam 3 was included in the correlation as a measure taken closer in time to when the fuel cells activity initial questions are completed by the students. Correlations by semester for the solutions activity is in **Table 3.10**. Correlations by semester for the fuel cells activity is in **Table 3.11**. Positive correlations between the solutions activity initial questions and beginning of semester measures support the scenario placement based on initial questions scores. Positive correlations between the fuel cells activity initial questions and beginning of semester measures support the scenario placement. The significant positive correlation with exam 3 supports the scenario placement. Exam 3 is a measure closer to the time-point when students complete activity 2 (fuel cells) initial questions and provides a better measure of student content knowledge than beginning of semester measures alone.

**Table 3.10 Correlations for the solutions activity initial questions and beginning of semester measures for semester 1 and 2**

| | | Initial questions semester 1 | Initial questions semester 2 |
|---|---|---|---|
| ACT COMP | Correlation | .258* | .325** |
| | *n* | *97* | *128* |
| ACT READ | Correlation | .239* | .287** |
| | *n* | *97* | *128* |
| ACT ENGL | Correlation | 0.173 | .274** |
| | *n* | *97* | *128* |
| ACT MATH | Correlation | .312** | .249** |
| | *n* | *97* | *128* |
| ACT SCIRE | Correlation | 0.147 | .257** |
| | *n* | *97* | *128* |
| Beginning of semester SLST | Correlation | .206* | .458** |
| | *n* | *120* | *127* |
| Beginning of semester SCI | Correlation | 0.012 | .325** |
| | *n* | *100* | *119* |
| Beginning of semester SLS | Correlation | 0.139 | .465** |
| | *n* | *96* | *99* |
| Placement Test | Correlation | 0.104 | .472** |
| | *n* | *124* | *150* |

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

**Table 3.11 Correlations for the fuel cells activity initial questions and beginning of semester measures for semester 1 and 2**

| | | Initial questions semester 1 | Initial questions semester 2 |
|---|---|---|---|
| ACT COMP | Correlation | .335** | .407** |
| | *n* | *111* | *115* |
| ACT READ | Correlation | .232* | .320** |
| | *n* | *111* | *115* |
| ACT ENGL | Correlation | .309** | .356** |
| | *n* | *111* | *115* |
| ACT MATH | Correlation | .291** | .362** |
| | *n* | *111* | *115* |
| ACT SCIRE | Correlation | .317** | .354** |
| | *n* | *111* | *115* |
| Beginning of semester SLST | Correlation | .328** | .601** |
| | *n* | *132* | *117* |
| Beginning of semester SCI | Correlation | 0.144 | .346** |
| | *n* | *102* | *103* |
| Beginning of semester SLS | Correlation | .304** | .476** |
| | *n* | *98* | *88* |
| Placement Test | Correlation | .189* | .388** |
| | *n* | *139* | *133* |
| Exam 3 | Correlation | .285** | .321** |
| | *n* | *139* | *134* |

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

## 3.4.2 Supplemental instruction and course results

When comparing supplemental instruction and course results the samples were those students who completed both supplemental instruction activities and those who completed one or no activities. For the two samples to start the semester at similar levels of understanding, the samples should not have a significant difference between beginning of semester measures such as ACT composite and sub-scores, placement exam and scale measures. Independent sample t-tests were performed for the SLST pre, SCI pre, SLS pre, placement test, and ACT composite score and sub-scores and none were significant at the 0.05 level. This indicates that there was not a significant difference between the mean for the students who completed the both supplemental instruction activities and those who did not at the beginning of the semester. The values for the independent t-tests are in **Table 3.12** and **Table 3.13**.

**Table 3.12 Independent samples t-tests for semester 1 (completed both activities minus completed 0 or 1 activity) (1 = finished 2 activities, 0 = finished 0 or 1 activities)**

| | | *n* | Mean | Std. Dev. | t | df | Sig. (2-tailed) |
|---|---|---|---|---|---|---|---|
| ACT COMP | 1 | *40* | 23.7 | 3.1 | -1.063 | 138 | 0.29 |
| | 0 | *100* | 24.41 | 3.7 | | | |
| ACT READ | 1 | *40* | 24.2 | 3.7 | -0.641 | 138 | 0.523 |
| | 0 | *100* | 24.76 | 5 | | | |
| ACT ENGL | 1 | *40* | 23.1 | 3.8 | -0.741 | 138 | 0.46 |
| | 0 | *100* | 23.71 | 4.6 | | | |
| ACT MATH | 1 | *40* | 22.5 | 3.9 | -2.161 | 138 | 0.032 |
| | 0 | *100* | 24.04 | 3.8 | | | |
| ACT SCIRE | 1 | *40* | 24.35 | 3.5 | -0.335 | 138 | 0.738 |
| | 0 | *100* | 24.59 | 4.1 | | | |
| SLST pre | 1 | *54* | 0.602 | 0.16 | -1.317 | 156 | 0.19 |
| | 0 | *104* | 0.633 | 0.12 | | | |
| SCI pre | 1 | *51* | 0.677 | 0.063 | -0.808 | 115 | 0.421 |
| | 0 | *66* | 0.686 | 0.06 | | | |
| SLS pre | 1 | *50* | 0.648 | 0.095 | -0.841 | 108 | 0.402 |
| | 0 | *60* | 0.662 | 0.074 | | | |
| Placement Test | 1 | *54* | 0.61 | 0.15 | 0.77 | 172 | 0.442 |
| | 0 | *120* | 0.591 | 0.15 | | | |

**Table 3.13 Independent samples t-tests for semester 2 (completed both activities minus completed 0 or 1) (1 = finished 2 activities, 0 = finished 0 or 1 activities)**

| | | *n* | Mean | Std. Dev. | t | df | Sig. (2-tailed) |
|---|---|---|---|---|---|---|---|
| ACT COMP | 1 | *56* | 23.48 | 3.7 | -0.145 | 173 | 0.885 |
| | 0 | *119* | 23.57 | 3.8 | | | |
| ACT READ | 1 | *56* | 22.98 | 5.2 | -1.653 | 173 | 0.1 |
| | 0 | *119* | 24.35 | 5.1 | | | |
| ACT ENGL | 1 | *56* | 23.05 | 5 | 0.131 | 173 | 0.896 |
| | 0 | *119* | 22.95 | 4.8 | | | |
| ACT MATH | 1 | *56* | 23.75 | 3.8 | 1.31 | 173 | 0.192 |
| | 0 | *119* | 22.92 | 3.9 | | | |
| ACT SCIRE | 1 | *56* | 23.86 | 3.7 | 0.467 | 173 | 0.641 |
| | 0 | *119* | 23.55 | 4.1 | | | |
| SLST pre | 1 | *55* | 0.613 | 0.16 | 0.046 | 167 | 0.964 |
| | 0 | *114* | 0.612 | 0.17 | | | |
| SCI pre | 1 | *52* | 0.69 | 0.072 | -0.1 | 151 | 0.92 |
| | 0 | *101* | 0.692 | 0.069 | | | |
| SLS pre | 1 | *46* | 0.665 | 0.096 | 0.015 | 124 | 0.988 |
| | 0 | *80* | 0.665 | 0.1 | | | |
| Placement Test | 1 | *61* | 0.621 | 0.15 | 0.66 | 200 | 0.51 |
| | 0 | *141* | 0.605 | 0.16 | | | |

The goal of the supplemental instruction activities is to support student learning in solutions and fuel cells. If the goal of the supplemental instruction activities has been met, then students who completed both supplemental instruction activities should score higher on the final exams or in the course than those who did not complete both activities. When investigated it was found that there was a significant difference between those students who completed two supplemental instruction activities compared to those who completed zero or one supplemental instruction activity. **Table 3.14** and **Table 3.15** contain the independent sample t-test information for semester 1 and semester 2 respectively.

**Table 3.14 Independent samples t-tests for semester 1 (both activities minus 0 or 1 activities) (1 = finished 2 activities, 0 = finished 0 or 1 activities)**

| | Paired Final Exam | | Conceptual Final Exam | | Course percent | |
|---|---|---|---|---|---|---|
| | 1 | 0 | 1 | 0 | 1 | 0 |
| *n* | *55* | *89* | *55* | *89* | *55* | *89* |
| Mean | 0.728 | 0.713 | 0.537 | 0.509 | 81.412 | 72.602 |
| Std. Deviation | 0.15 | 0.14 | 0.15 | 0.14 | 10.35 | 12.35 |
| t | 0.595 | | 1.118 | | 4.415 | |
| df | 142 | | 142 | | 142 | |
| Sig. (2-tailed) | 0.553 | | 0.266 | | 0.000** | |

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

**Table 3.15 Independent samples t-tests for semester 2 (2 activities minus 0 or 1 activities) (1 = finished 2 activities, 0 = finished 0 or 1 activities)**

| | Paired Final Exam | | Conceptual Final Exam | | Course percent | |
|---|---|---|---|---|---|---|
| | 1 | 0 | 1 | 0 | 1 | 0 |
| *n* | *62* | *127* | *62* | *127* | *62* | *127* |
| Mean | 0.723 | 0.695 | 0.543 | 0.505 | 88.569 | 80.850 |
| Std. Deviation | 0.15 | 0.16 | 0.14 | 0.16 | 7.7 | 12. |
| t | 1.146 | | 1.582 | | 4.642 | |
| df | 187 | | 187 | | 187 | |
| Sig. (2-tailed) | 0.253 | | 0.115 | | 0.000** | |

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

A goal of the supplemental instruction activities is to teach students chemistry. If the goal of the supplemental instruction activities has been met, then students who completed both supplemental instruction activities should score higher on the final than those who did not complete both activities. There was not a significant difference for the paired final or the conceptual final but there was a significant difference for the course grade.

## 3.5 Summary and Conclusions

Over 50% of students who completed the initial questions completed the activities; 66.19% who finished the initial questions for the solutions activity completed the final questions for the solutions activity and 56.52% of those who completed the initial questions for the fuel cells activity completed the final questions for the fuel cells activity. For the solutions activity, the majority of students (38.18%) placed into scenario 2 based on initial question score. For activity 2 (fuel cells) the majority of students (46.38%) placed into scenario 1 based on initial question score. For the solutions activity students performed significantly better on the final questions than the initial questions which supports the hypothesis that the activity supports student learning in solution. For the fuel cells activity students performed significantly better on the initial questions than the final questions.

For the fuel cell activity, a few reasons that students may have performed better on the initial questions compared to the final questions could be the content, and students not taking the final question seriously or wanting to be done (especially because the second activity is at the end of the semester and students may have more assignments due in other courses). Both the initial and the final questions are pulled from identical question pools based on subtopic, so a difference in the complexity of the questions does not exist.

There was not a significant difference in the mean for the final exam part 1 or part 2 between those students who completed both supplemental instruction activities and those who completed 0 or 1 activity. Students who completed both supplemental instruction activities had a significantly higher course percent than those students who did not complete both activities. Multiple measures make up the course score at a variety of time points throughout the semester so a significant difference in course percent could be attributed to more than just supplemental instruction activities.

# 3.6 Limitations

The supplemental instruction was designed to support student learning in the topics of solutions and fuel cells. Although scale is used as a theme in the supplemental instruction, student performance cannot convey how their scale ability changes as they move throughout the activity or as they complete both activities. The activities do not measure scale ability but chemistry knowledge.

A limitation of supplemental instruction is that more motivated students may be those completing the activities. The activities were low-stakes and as such students with lower motivation may have been less likely to interact with the activities.

A limitation during the analysis of how students utilized the activities may be combining the samples. Combining the semesters led to an increase in sample size and the results of the t-tests were not significant. However, there may be a difference between the semesters due to their different treatments. This may be a greater factor for the fuel cells activity which takes place at the end of the semester and after the treatment has taken place.

# 3.7 Implications for Instruction

Supplemental instruction can be used to help bridge the gap between what students are able to do on their own and what they are able to do with help. Supplemental instruction can be used with a framework, such as Johnstone's triangle, to improve content understanding. Supplemental instruction allows students additional instruction with a challenging topic. The adaptive learning model allows instruction to be targeted based on the amount of understanding a student already had about a topic. Any chemistry topic can be designed in this way and an instructor can use the information gathered by the supplemental instruction instrument to tailor their material for the students. The current supplemental instruction activities for both General Chemistry I and General Chemistry II are available via the scale website for an instructor to utilize them, or other materials, in their course.

# 3.8 Citations


Chieu, V. M. (2005) *Constructivist Learning: An Operational Approach for Designing Adaptive Learning Environments Supporting Cognitive Flexibility*. Thesis p 70-96.

Dry, M. (2017). Assessing the Utility of an Online Adaptive Tool in a Large Undergraduate Psychology Program. Report.

Eysenck, M. (2012). *Fundamentals of Cognition.* New York, NY: Psychology Press. p. 159

Gerlach, K.; Trate, J.; Blecking, A.; Geissinger, P.; Murphy, K. (2014). Investigation of Absolute and Relative Scaling Conceptions of Students in Introductory College Chemistry Courses. *J. Chem. Education*, *91*, 1526-1537.

Gerlach, K.; Trate, J.; Blecking, A.; Geissinger, P.; Murphy, K. (2014). Valid and Reliable Assessments to Measure Scale Literacy of Students in Introductory College Chemistry Courses. *Journal of Chemical Education*. *91*, 1538-1545.

Johnstone, A. H. (1993). The development of chemistry teaching: A changing response to changing demand. *Journal of Chemical Education, 70*(9), 701-705.

Kara, N. & Sevim, N. (2013). Adaptive learning systems: beyond teaching machines. *Contemporary Educational Technology*, *4*(2), 108-120.

Kerr, P. (2015) Adaptive Learning. *Technology for the Language Teacher,* 70(1), 88-93.

National Research Council; *Next Generation Science Standards: for states, by states*; Washington, D.C., National Academies Press: Washington, D.C., 2013.

Piaget, J., & Cook, M. T. (1952). *The origins of intelligence in children*. New York, NY: International University Press.

Siyepu, S. (2013). *The zone of proximal development in the learning of mathematics.* South African Journal of Education. *33*(2). p. 1-13.

Skinner, B. F. (1960). Teaching machines. *The Review of Economics and Statistics*, *42*, 189-191.

Stolurow, L. M. & Davis, D. (1965). Teaching machines and computer-based systems. In R. Glaser (Ed.), *Teaching machines and programmed learning II: Data and directions.* Washington, D.C.: National Education Association of the United States. 162-212.

Warford, M. K. (2011). *The zone of proximal teacher development.* Teaching and Teacher Education. *33*(2), 1-12.

# Chapter 4: Scale Conception of Students in Anatomy and Physiology I as measured through a one-on-one Scale Activity

## 4.1 Introduction

Previous research has found that preparatory and general chemistry students have a lower scale conception than chemistry experts[60,61]. The first step in determining if Anatomy and Physiology I is a good candidate for scale integration as a theme is to understand at what ability level current students in the course have with regard to scale. This chapter details the initial interviews with Anatomy and Physiology I students (novices in biological sciences) and their teaching's assistants (TA) (more experienced learners in biological sciences) examining their current conception of and ability with scale.

## 4.2 Background

Based on an original set of scale activities first published by Laubach, et al., Thomas R. Tretter and M. Gail Jones adapted an activity where a clothesline was stretched out across the classroom [62],[63]. The instructor placed 0 and 1 meter on the number line and students placed cards with values, both standard decimal, e.g. 2, and scientific notation, e.g. $10^2$ and $10^{-3}$. Students were also given object cards to place on the number line, e.g. atom and football field. Students were able to place the 2, 3, and $10^0$ relatively easily on the number line but struggled when the card contained a negative exponent such as $10^{-1}$. This prompted the instructor to lead a

[60] Gerlach 2014b: 1526-1537
[61] Trate 2017: 88-108
[62] Laubach, Royce, and Holzer 2000: 48-50
[63] Tretter and Jones 2003: 22-25

class discussion about conceptualizing size and creating benchmarks to help students identify the relative size of objects. The article ended with Tretter and Jones speaking to the importance of understanding logarithmic scales in biology and how this activity could be used to improve student scale conception[64].

Tretter and Jones continued to study scale. With Thomas Andre, Atsuko Negishi, and James Minogue they studied the understanding of scale of 5th, 7th, 9th, and 12th-grade students along with doctoral students[65]. Students were given the Scale of Objects Questionnaire (SOQ). The SOQ listed 6 objects, such as "length of a grain of white rice" and instructed students to select a size range. The size ranges were given on a 12-point scale ranging from "<1 nm" to ">1 billion meters". The SOQ was followed by a card sort activity. In the card sort activity participants were given 31 objects and instructed to sort them into piles according to size. The objects ranged in size from the subatomic to the galactic. Participants were more accurate with relative scaling (sorting objects) compared to absolute scaling (SOQ). Participants utilized landmarks, such as the size of a human, to establish scale with the more experienced students expressing the use of more landmarks than the novice students.

Tretter and Jones continued their research with James Minogue by studying scale conception of different expertise levels[66]. Students from grades 5, 7, 9, 12, and doctoral participated in written assessments and a card sort activity. Students from grades 5, 7, and 9 were classified as novices, students with in grade 12 were classified as experienced, and doctoral students were classified as experts. Students were given the Scale Anchoring Objects assessment (SAO) which consisted of two parts. Part A listed sizes in increasing order from "1 meter" to

64 Tretter and Jones 2003: 22-25
65 Tretter 2006a: 282-319
66 Tretter 2006b: 1061-1085

“1,000,000,000 meters (one billion meters)” and Part B listed sizes in decreasing order using body length, e.g. “equal to your body length” to “1/1,000,000,000 your body length (one billionth your body length)”. Next to the listed size in Part A and B was a space for students to write an object they identified with that size. A list of commonly identified objects was compiled by the researchers including “atom” and “ant”. Incorrect object listing was most often seen when students selected an object that was too large for a particular small length and when an object that was too small was selected for a large length. As expertise increased from novice to experienced to expert students the number of incorrect object listings decreased. After completing the SAO, students were interviewed with regards to their thinking about scale by asking how the student arrived at the object they wrote on the SAO. The greater experience a student had, the more specific strategies the students articulated, e.g. “In the chemistry book I taught from, atomic radii were listed in Angstroms” and Angstroms are “close to nanometer size”[67]. The specific strategies listed by the more experienced students separated into two categories: mathematical computations or object comparisons. Mathematical computations include use of the metric system while object comparison uses objects, such as comparing the object they are sorting to an atomic radii, to arrive at an answer. During the interview the more experienced a student was in scale, the more comfortable they reported being with the metric system. Experts expressed comfort in making mental jumps between large and small scales. Experienced students demonstrated a “transition to thinking like the experts”. Novice students were vague about their strategies or would use mathematical computations to estimate a size. There is a common “scale boundary” at the edge of human sight where students have difficulty overcoming the boundary to correctly place items. Experts can jump to a new scale, jump the

[67] Tretter 2006b: 1061-1085

scale boundary, and unitize within the new size and experts articulated the importance of experience in their understanding of scale.

Adapted from the interviews conducted by Thomas Tretter and Gail Jones[68], Karrie Gerlach et al. interviewed undergraduate chemistry students one-on-one while the students completed a card sorting activity and placed objects on a logarithmic number line[69]. Students in preparatory and general chemistry were classified as novices while chemistry graduate students were classified as experienced students. The interview consisted of four parts: bin creation and item sort (part I), item ordering within bins (part II), item ordering with measurements (part III), and item ordering on a number line (part IV). The interviews focused on absolute and relative scaling of objects by having participants first organize objects relative to other objects (relative scaling) and then placing the same objects on a logarithmic number line (absolute scaling). In part I students were instructed to create bins to sort objects by size. Once the bin labels were created students were given 20 object cards which only had an object name on them and instructed to sort the cards into the bin and within each bin by size. After part II the cards were collected and handed a second set of cards to sort into and within each bin. The second set of cards contained the same objects as the first set but also listed their size, in the most common unit with which the object is measured, such as an atom was listed as 100 pm. In part IV a logarithmic number line was placed in front of the student. Pieces of paper with the same objects and sizes listed were given to the student and the student was instructed to place the objects on the number line.

[68] Tretter 2006: 1061-1085
[69] Gerlach 2014b: 1526-1537

Results from this study showed that experienced students' conception of scale is more developed than that of the novices[70]. Participants relative scaling was strongest within 3 orders of magnitude (from $10^{-3}$ to $10^{3}$ meters) of adult height and adult height was often used as an anchor for determining sizes. Experienced students created more bins that would fall in the nonvisible range than novice students demonstrating the novice students narrow scale conception. The placement of the smaller objects, such as virus and bacterium, as similar in size supported the conclusion that the participants "perceive nonvisible, small objects as similar in size"[71]. The Anatomy and Physiology I scale activity interview protocol was adapted from Gerlach et al[72].

The interviews in chemistry led to the development of a chemistry class-wide laboratory scale activity[73]. The goal was to increase student scale conception and study scale conception with a larger sample. Parts I-III of the interview were used but students worked in pairs. Part IV of the activity was adapted into a worksheet that gave students practice working with a logarithmic number line. An absolute scaling activity was added that had students move from the size of a human to the size of an atom using the objects given in parts I-III. The results of the class-wide activity were consistent the results found in the chemistry interviews[74]. Students created one bin for all nonvisible objects, one bin for large objects, and multiple bins around their height demonstrating comfort with sizes surrounding adult height. Students struggled to correctly order virus and bacterium compared to each other as well as cruising height of a 747 jet

[70] Tretter 2006b: 1061-1085
[71] Gerlach 2014b: 1536-1537
[72] Gerlach 2014b: 1526-1530
[73] Trate 2017: 88-108
[74] Gerlach 2014b: 1536-1537

and the width of Wisconsin. Ordering accuracy improved when metric sizes were given with the objects.

Scale, as a cross-cutting concept, is important in chemistry and anatomy and physiology. Interviews with novice students of different disciplines have yielded similar results[75,76]. The hypothesis of this chapter is Anatomy and Physiology I students have a lower scale conception than experienced students and the demonstrated scale conception of Anatomy and Physiology I novice students is similar to novice chemistry students.

# 4.3 Methods

The scale activity interviews were conducted one-on-one in a semi-structured interview format in the last month of the semester of an Anatomy and Physiology I course (interview protocol (IRB approval # 14.404)). The 60-minute interviews were conducted and recorded following the protocol developed for the one-on-one interviews with chemistry students[77]. Notes were taken by the interviewer in real time and photographs were taken of Part IV of the activity. Two types of student were interviewed: novice students and experienced students[78]. Novice students were students currently taking Anatomy and Physiology I ($n$ = 22) and experienced students were the Anatomy and Physiology I and II TAs ($n$ = 10).

## 4.3.1 Adaptations of the activity

The interview protocol language was adapted from the interviews conducted with introductory chemistry students[79]. The original activity contained 20 object cards but was

[75] Tretter 2006b: *1061*-1085
[76] Gerlach 2014b: 1526-1537
[77] Gerlach 2014b: 1526-1537
[78] Gerlach 2014b: 1526-1537
[79] Gerlach 2014b: 1526-1537

reduced to 15 objects to reduce the amount of time for the interview and provide the students and the interviewers with a wide range of sizes with fewer similar size objects.

## 4.3.2 Exclusions

One experienced student was removed from the analysis. This was due to excessive errors, compared to the other experienced students, and student comments such as their brain "being fried" from writing their thesis but the compensation for the activities was worth it. The experienced student had 1.5× more errors in Part IV than the next highest experienced student's total magnitude errors. A box plot was created, and the experienced student was identified as an extreme outlier. The box plot is in **Figure 4.1**.

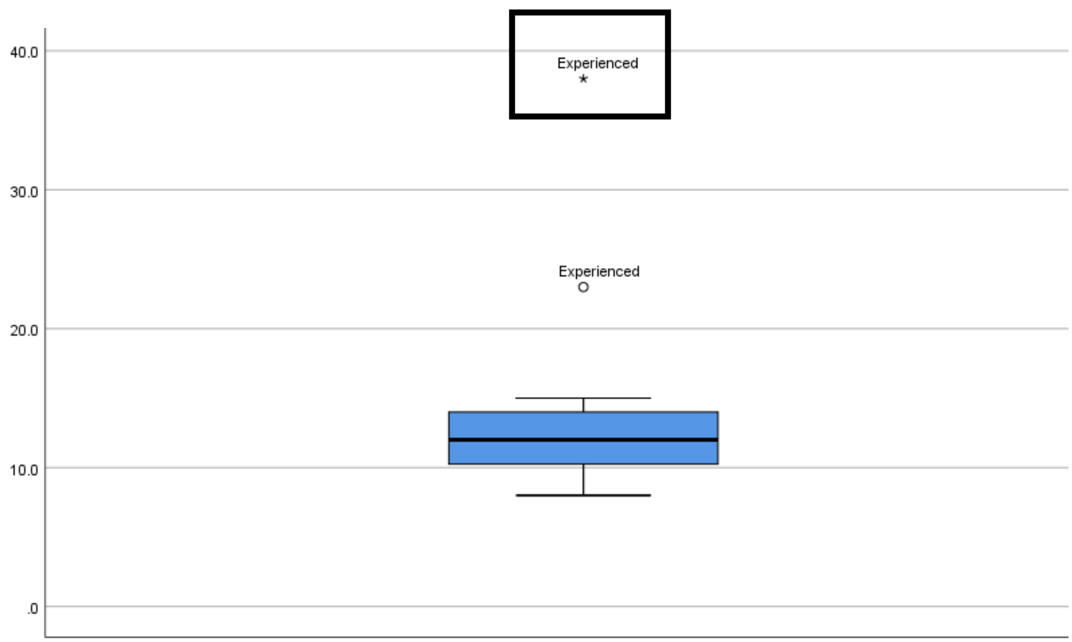



**Figure 4.1 Part IV Total number of errors box plot for experienced students**

### 4.3.3 Overview of activity

The scale activity interviews described in this chapter mirrored the activity used by Gerlach, et al. and consisted of four parts[80]:

- Part I: Bin creation and initial item sort
- Part II: Ordering objects within bins
- Part III: Ordering object with measurements within bins
- Part IV: Placing objects on a logarithmic number line.

The interview ended with follow-up open response questions.

#### 4.3.3.1 Overview of part I

Part I contributed to the investigation of relative scale conception. Students were instructed to make bins to sort objects by size. Students were given examples of bins that could be used to sort lengths of time such as “1 hour” or “the length of time to walk half a mile”. Time was used as the example to avoid influencing the students by giving examples of sizes. Bin creation criteria included no gaps between bins (no object could be placed between bins, one bin ends where the next begins), no overlaps (object placed in exactly one bin), and the end bins needed to be open-ended to include any potentially larger or smaller objects that could not be placed in a different bin. An example of bins a student may have made is shown in **Figure 4.2**. Students were not given any limit in their number of bins they could create but were told they would be sorting 15 objects. The students did not see the object cards until after their bins were created and checked for all of the requirements by the interviewer. Students were instructed that

[80] Gerlach 2014b: 1526-1537

they could change their bins at any point during the activity. If they chose to change bins, this was recorded along with their new bins.


*Figure 4.2 Example of student-created bins with no gaps, no overlaps, and open-ended bins*

After their bin creation, students were given 15 cards each with the name of a single specific object on them. **Table 4.1** lists the objects and object lengths, in the most common metric unit. Students were handed the cards sorted alphabetically and were instructed to sort the objects into the proper bin. If a student asked for clarification, for example, what type of cell, the interviewer provided the predetermined answer, a human red blood cell. An example of bins a student may have made with sorted objects is shown in **Figure 4.3**. During the interview, the interviewer recorded the bins the student created. If the student changed any bins, this was recorded along with into which bins the objects were sorted.

**Table 4.1 List of object names, abbreviations, and most commonly used measurements for the scale activity listed according to size**

| Object (from smallest to largest) | Abbreviation | Size |
|---|---|---|
| atomic nucleus | atom nuc | 10 fm |
| atom | atom | 100 pm |
| virus | virus | 100 nm |
| bacterium | bcm | 1 µm |
| cell | cell | 7 µm |
| hair width | hair | 100 µm |
| finger | finger | 8 cm |
| new pencil length | pencil | 21 cm |
| textbook | text | 28 cm |
| adult height | adult | 2 m |
| football field | field | 91 m |
| cruising altitude of 747 jet | jet | 11 km |
| width of Wisconsin | WI | 450 km |
| earth to moon | moon | 384 Mm |
| earth to sun | sun | 146 Tm |

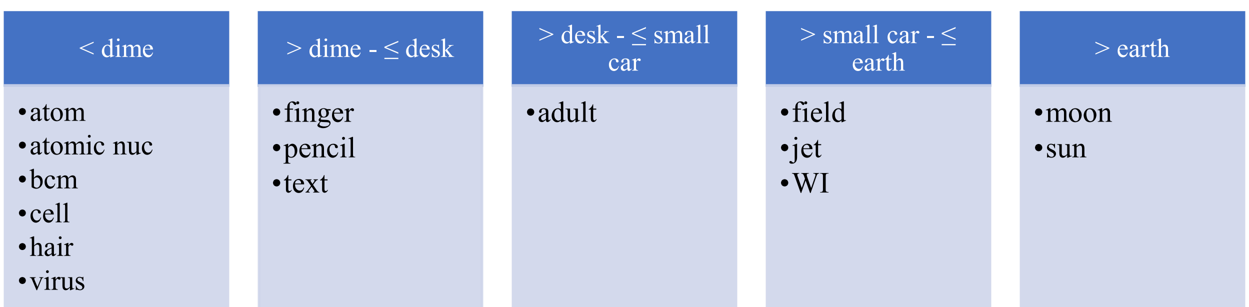



*Figure 4.3 Example of sorting cards into bins (objects only)*

### 4.3.3.2 Overview of part II

Part II contributed to the investigation of relative scale conception. Students were instructed to use the objects they had just sorted and order the objects from smallest to largest within each bin. The cards were the same as those given in Part I containing only the object name. Students were allowed to move cards between bins and reminded that they could change their bins at any time.  The interviewer recorded the bin labels, which objects were placed in each bin, and the ordering (smallest to largest) of the objects within each bin.  Once this information was recorded the cards were collected.

### 4.3.3.3 Overview of part III

Part III contributed to the investigation of relative scale conception.  Students were given 15 new object cards, sorted alphabetically, that listed the same objects *and* the size of the object, in the most common metric unit (**Table 4.1**).  Students were instructed to sort the cards into the bins and within each bin by size and again informed that at any point they could change their bins.  An example of how a student may have sorted the cards is shown in **Figure 4.4**.  The interviewer recorded the bin labels, the bin each object was placed in, and the order of the objects within each bin.  Once this information was recorded the objects and bins were collected.

| < dime | dime - desk | desk - small car | small car - earth | > earth |
|---|---|---|---|---|
| • atom 100 pm<br>• atomic nuc 10 fm<br>• bcm 1µm<br>• cell 7 µm<br>• hair 100 µm<br>• virus 100 nm | • finger 8 cm<br>• pencil 21 cm<br>• text 28 cm | • adult 2 m | • field 91 km<br>• jet 11 km<br>• WI 450 km | • moon 384 Mm<br>• sun 146 Tm |

*Figure 4.4 Example of sorting cards into bins (objects and sizes)*

### 4.3.3.4 Overview of part IV

Part IV contributed to the investigation of relative and absolute scale conception.  A logarithmic number line in scientific notation was placed in front of the student.  The number line ranged from $10^{-9}$ to $10^{9}$ and had no unit indicated.  The students were instructed to place the objects on the number line and define the unit they used.  Most students defined the unit after they had placed the objects on the number line. The objects for the number line were the same as the cards in Part III containing both the object name as well as the size listed in the most common unit. After the interview the number line was photographed. An example of how a student may have placed the objects on the number line is shown in **Figure 4.5**. After the interview the absolute placement of objects, boundaries of human sight, and current technology were recorded.

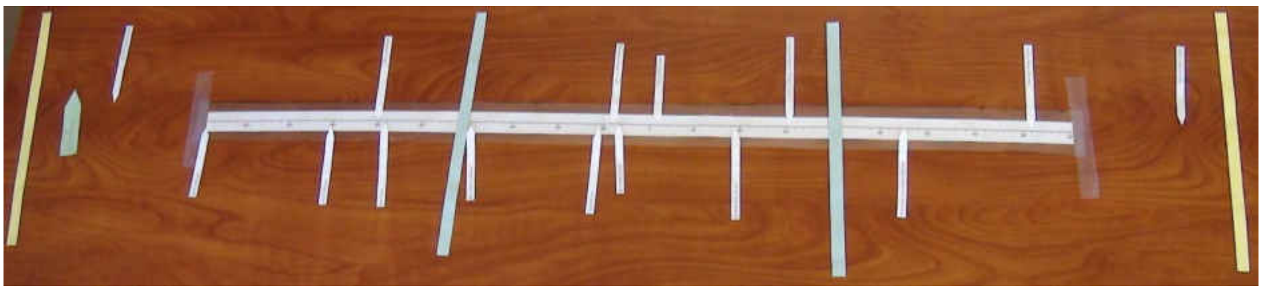


*Figure 4.5 Example of Part IV logarithmic number line; blue cardstock shows that the boundaries of human sight are 4 orders of magnitude in both directions.*

## 4.3.4 Data Analysis

The bins and placement of objects were analyzed. The bins and object relative object placement were analyzed for Parts I-III (relative scaling ability). Part IV measured absolute scale conception which allowed for analysis using the placement of the objects on the number line compared to where they should have been placed.

### 4.3.4.1 Part I

Part I was analyzed by determining the number and types of bins created. The bin names were recorded, and identified as either using of measurements, objects, or both as bin names. The number of nonvisible bins was determined by the following method. If the largest object, or measurement, used in the bin boundary was a nonvisible object or measurement (the threshold of human sight is $10^{-4}$ m) the bin was considered nonvisible. This guarantees that the entire bin would fall in the nonvisible region. For example, using the bins in **Figure 4.4** none of the bins are considered nonvisible because all of the bin labels have at least one visible object in the label. An additional example is a bin ranged cell-ant. The largest bin label boundary is ant which is a visible object and the bin would be considered visible but if the bin range was cell-macromolecule then the bin would fall in the nonvisible region.

The bin each object was placed in was analyzed. Using **Figure 4.4** as an example the objects atomic nucleus, atom, virus, bacterium, cell, and hair width are located in the smallest bin created (bin number 1). The bin boundaries presented the range of sizes the created bins could hold. The range of sizes the bins covered showed that even when students are told a range of sizes would be presented, the actual range of sizes the student was consciously aware of is narrower than what was presented for objects. Bins were analyzed based on which object was the largest that could fit within the bin boundary range. Examples are in **Table 4.2**.

**Table 4.2 Examples of bin boundaries and categorization based on object**

| Bin Boundary<br>(bin description in parentheses) | Categorization<br>(reasoning in parentheses) |
|---|---|
| globe<br>(bigger than a Rubik cube but smaller than the globe) | width of Wisconsin<br>(the width of Wisconsin is smaller than the globe) |
| garbage can<br>(bigger than a shoebox but smaller than a garbage can) | textbook<br>(a textbook is smaller than a garbage can) |
| snail<br>(larger than a molecule but smaller than a snail) | hair<br>(a hair is smaller than a snail) |

## 4.3.4.2 Part II and Part III

Part II and III were analyzed by recording the item ordering of the students. The item order was recorded and compared to the correct ordering of objects. Item ordering indicates students' relative size understanding when comparing objects to one another. An object that is supposed to be in ordering position 3 but has an average of 1 means that students, on average, placed the object first in the list (smaller than actual size). Another way this is recorded is that the object was placed "-2" meaning it should have been placed smaller, by two items.

## 4.3.4.3 Part IV

Part IV was analyzed based on orders of magnitude. The scoring method used to determine the errors of the placement of the object on the number line was determined based on being within ± 1 order of magnitude from the correct answer. For example, if adult height was placed at 1 km, the student would be scored as +2 (or 2 orders of magnitude too large). When an object fell outside of the range of the number line, greater than $10^9$ or less than $10^{-9}$, students were scored by the objects being placed outside of the number line (the correct answer for the object is recorded as $10^9$ or $10^{-9}$ depending on if the object was placed on the large or small end of the scale respectively) and the correct ordering of the objects. Students were scored based on individual item placement, the sum of their absolute errors, and their average amount of absolute errors.

# 4.4 Results

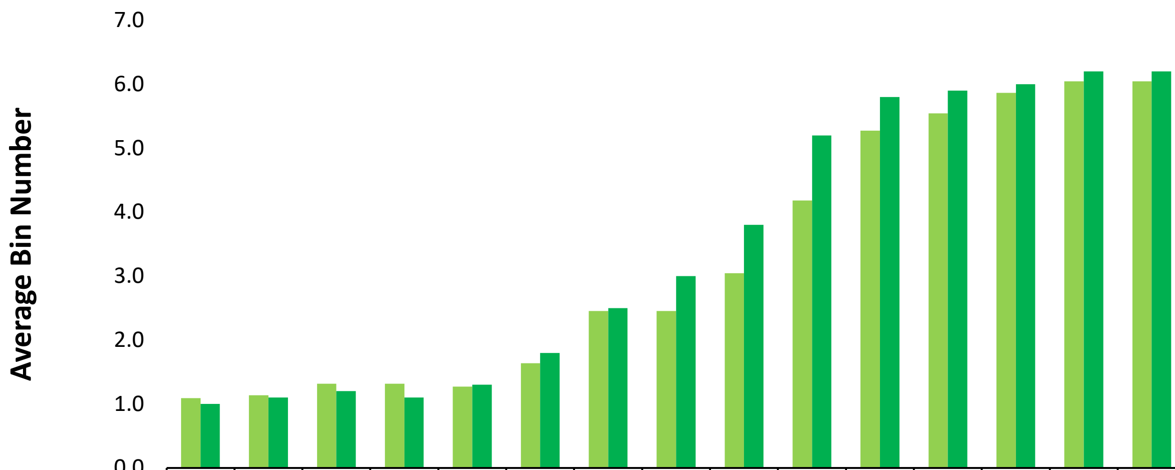
## 4.4.1 Part I: Bin creation

On average the novice students created 6.46 bins and experienced students created 6.30 bins. When creating their initial bins, 23% (5) of novice students and 20% (2) experienced students created at least one bin that could contain an object not visible to the naked eye. 68% of novice students and 40% of experienced students used objects to label bins; 32%, and 60%, respectively, used measurements, such as inches or millimeters. No student used objects and measurements as bin labels. One novice student and three experienced students changed bins during the interview. The novice student who changed bins went from using objects as bin labels to using metric system measurements and increased the number of bins they had from 4 to 8. **Table 4.3** describes the results of the student's bin creation with the number of nonvisible bins as well as those who created bins both greater and less than 3 orders of magnitude from 1 meter (boundaries of human sight). Beyond 3 orders of magnitude anchor points are created with new objects to provide a better sense of scale conception.

**Table 4.3 Nonvisible and orders of magnitude bin creation results for Anatomy and Physiology I novice and experienced students**

| | Novice Students $n = 22$ | Experienced Students $n = 10$ |
|---|---|---|
| Students who created at least 1 bin +/- 3 orders of magnitude from 1 meter | 16 (72%) | 10 (100%) |
| Students who created at least 2 bins +/- 3 orders of magnitude from 1 meter | 9 (40%) | 4 (40%) |
| Students who created 3 bins +/- 3 orders of magnitude from 1 meter | 8 (36%) | 2 (20%) |
| Students who created at least 1 nonvisible bin | 5 (23%) | 2 (20%) |
| Students who created at least 2 nonvisible bins | 3 (14%) | 1 (10%) |
| Students who created at least 3 nonvisible bins | 1 (4%) | 1 (10%) |

Most novice and experienced students created a single bin with bin boundaries that encompassed all items smaller than a hair width, a range of 10 orders of magnitude. Similar results were found for large orders of magnitude. Students created 1-2 bins for anything larger than 1 km, a range of 10 orders of magnitude. Students created 3-4 bins for the visible region, 7 orders of magnitude from finger to adult height. This shows that students are more comfortable in the size range they interact with daily, they have a greater number of bins in the region, but larger than adult height and smaller than finger students become less comfortable, with fewer bins and fewer size distinctions. **Figure 4.6** shows which bin, on average, each item fell in. For analysis of Part I the bins was numbered 1-smallest bin, to largest. For example, on average atomic nucleus was placed in bin 1 (1.1 average).

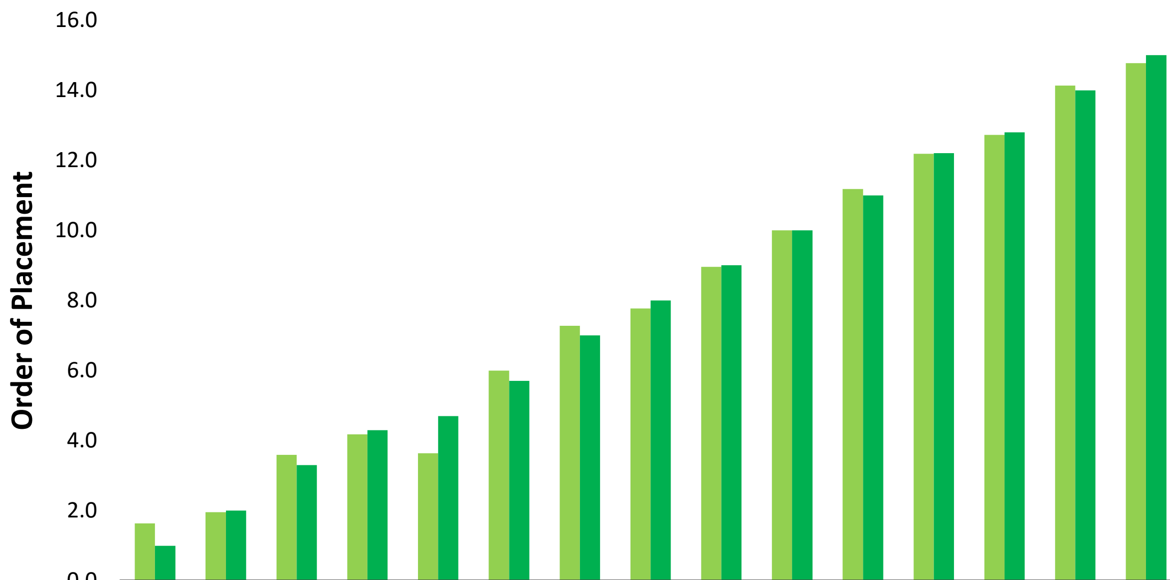



| | atom nuc | atom | virus | bcm | cell | hair | finger | pencil | text | adult | field | jet | WI | moon | sun |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Novices | 1.1 | 1.1 | 1.3 | 1.3 | 1.3 | 1.6 | 2.5 | 2.5 | 3.0 | 4.2 | 5.3 | 5.5 | 5.9 | 6.0 | 6.0 |
| Experienced | 1.0 | 1.1 | 1.2 | 1.1 | 1.3 | 1.8 | 2.5 | 3.0 | 3.8 | 5.2 | 5.8 | 5.9 | 6.0 | 6.2 | 6.2 |

*Figure 4.6 Part I: average bin number object was placed in, no sizes are given on object card; see abbreviation table for abbreviations*

## 4.4.2 Part II: Sorting objects without sizes

Students struggled to correctly place items in the nonvisible region as shown in **Figure 4.7**. A student who placed items correctly would have each object equal to exactly their placement order according to size, for example atomic nucleus would be 1.0 meaning atomic nucleus is the first item and the cruising altitude of a 747 jet would be the 12$^{th}$ object (a score of 12.0). Students struggled to correctly place virus, bacterium, and cell in the correct order. Virus, bacterium, and cell alternated between being placed at the 3$^{rd}$ and 4$^{th}$ object when those items are in fact 3$^{rd}$ (virus), 4$^{th}$ (bacterium), and 5$^{th}$ (cell). Students also found difficulty in placing cruising altitude of a 747 jet (12$^{th}$ object) and Wisconsin (13$^{th}$ object). Three experienced students changed bins during Part II.

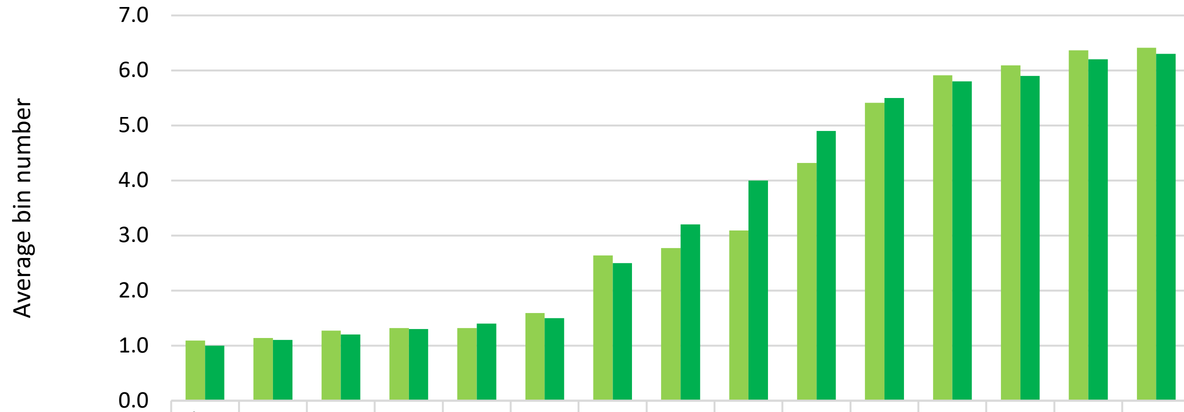



| | atom nuc (1) | atom (2) | virus (3) | bcm (4) | cell (5) | hair (6) | finger (7) | pencil (8) | text (9) | adult (10) | field (11) | jet (12) | WI (13) | moon (14) | sun (15) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Novice | 1.6 | 2.0 | 3.6 | 4.2 | 3.6 | 6.0 | 7.3 | 7.8 | 9.0 | 10.0 | 11.2 | 12.2 | 12.7 | 14.1 | 14.8 |
| Experienced | 1.0 | 2.0 | 3.3 | 4.3 | 4.7 | 5.7 | 7.0 | 8.0 | 9.0 | 10.0 | 11.0 | 12.2 | 12.8 | 14.0 | 15.0 |

*Figure 4.7 Part II: average object placement, no sizes given; see abbreviation table for abbreviations*

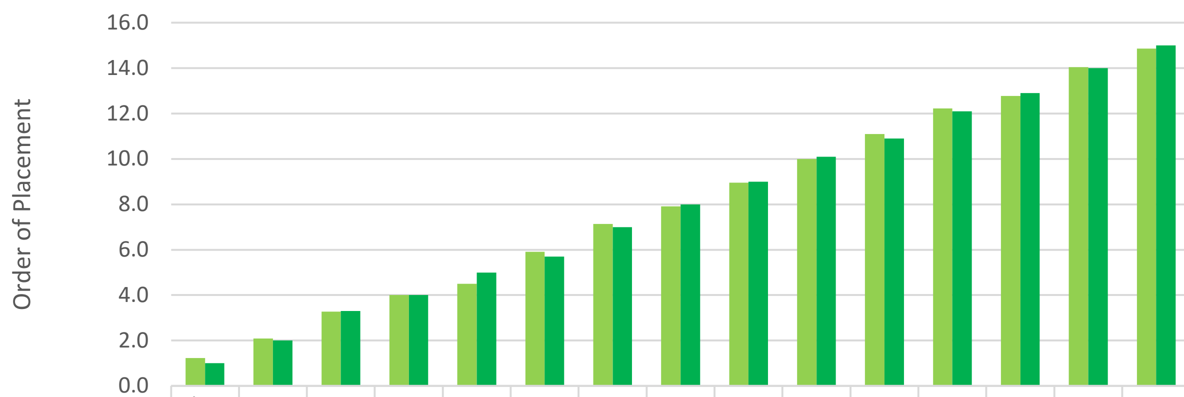
### 4.4.3 Part III: Sorting objects with sizes

The data collected for Part III was the student ordering of the objects with sizes. The average bin the object was placed in did not drastically change due to the objects being the same for both sets of cards and only one novice and two experienced students changed bins, see **Figure 4.8**. The average object placement of the cell changed from 3.6 to 4.5 for novice students and 4.7 to 5.0 for experienced students. When students had the sizes the order of objects was usually corrected as shown in **Figure 4.9**. The novice student changed bins during Part III.




| | atom nuc (1) | atom (2) | virus (3) | bcm (4) | cell (5) | hair (6) | finger (7) | pencil (8) | text (9) | adult (10) | field (11) | jet (12) | WI (13) | moon (14) | sun (15) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Novice | 1.1 | 1.1 | 1.3 | 1.3 | 1.3 | 1.6 | 2.6 | 2.8 | 3.1 | 4.3 | 5.4 | 5.9 | 6.1 | 6.4 | 6.4 |
| Experienced | 1.0 | 1.1 | 1.2 | 1.3 | 1.4 | 1.5 | 2.5 | 3.2 | 4.0 | 4.9 | 5.5 | 5.8 | 5.9 | 6.2 | 6.3 |

*Figure 4.8 Part III: average bin number object was placed in*



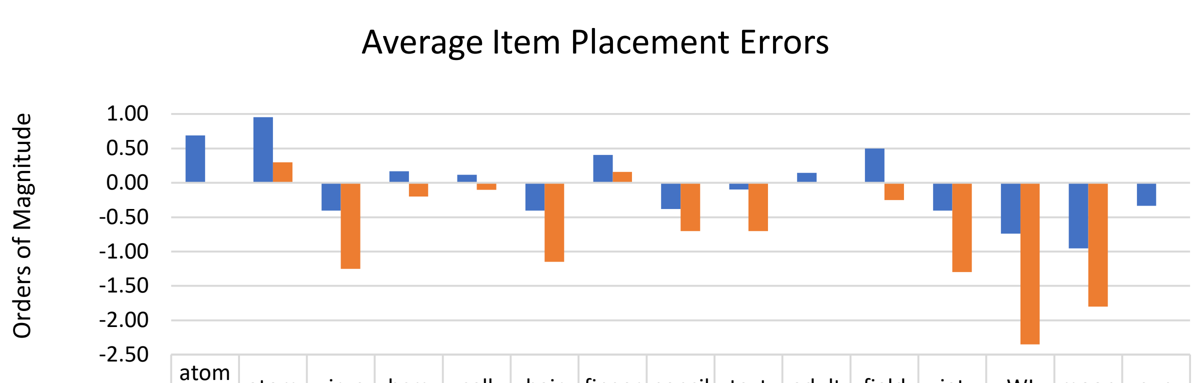

| | atom nuc (1) | atom (2) | virus (3) | bcm (4) | cell (5) | hair (6) | finger (7) | pencil (8) | text (9) | adult (10) | field (11) | jet (12) | WI (13) | moon (14) | sun (15) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Novice | 1.2 | 2.1 | 3.3 | 4.0 | 4.5 | 5.9 | 7.1 | 7.9 | 9.0 | 10.0 | 11.1 | 12.2 | 12.8 | 14.0 | 14.9 |
| Experienced | 1.0 | 2.0 | 3.3 | 4.0 | 5.0 | 5.7 | 7.0 | 8.0 | 9.0 | 10.1 | 10.9 | 12.1 | 12.9 | 14.0 | 15.0 |

*Figure 4.9 Part III: average object placement, sizes given*

## 4.4.4 Part IV: Logarithmic number line

In Part IV the placement of objects on a scientific logarithmic number line was measured. A student who placed all objects at the correct size would have a sum and average amount of errors of object placement of 0. The objects would be at exactly the correct order of magnitude (0 orders of magnitude away from correct answer).

Both novice and experienced students had errors in placing the objects on the number line. One specific example is a 747 jet was placed at a lower order of magnitude than its actual size (smaller). **Figure 4.10** has the average item placement errors with negative numbers being an object was placed smaller than its actual size and a positive number for objects placed larger than their actual size. Objects smaller than adult height were more often placed larger than their actual size. Both novice and experienced students placed cruising altitude of a 747 jet, width of Wisconsin, earth to moon, and earth to sun on the number line at a smaller position than their actual size. Experienced students had more difficulty placing larger objects than novice students which could be due to their domain-specific knowledge.




| | atom nuc | atom | virus | bcm | cell | hair | finger | pencil | text | adult | field | jet | WI | moon | sun |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Novice | 0.69 | 0.95 | -0.40 | 0.17 | 0.12 | -0.40 | 0.40 | -0.38 | -0.10 | 0.14 | 0.50 | -0.40 | -0.74 | -0.95 | -0.33 |
| Experienced | 0.00 | 0.30 | -1.25 | -0.20 | -0.10 | -1.15 | 0.16 | -0.70 | -0.70 | 0.00 | -0.25 | -1.30 | -2.35 | -1.80 | 0.00 |

*Figure 4.10 Part IV: average item placement errors*

When all the absolute value of the errors was totaled and the average was taken for novice and experienced students. The experienced students have fewer total errors than novice students. Experienced Anatomy and Physiology I students had 3 orders of magnitude less errors than novice Anatomy and Physiology I students. Chemistry novice students and anatomy novice students had the same number of errors, see **Figure 4.11**. **Table 4.4** lists descriptive statistics for the average amount of errors (where every 1 = 1 order of magnitude)

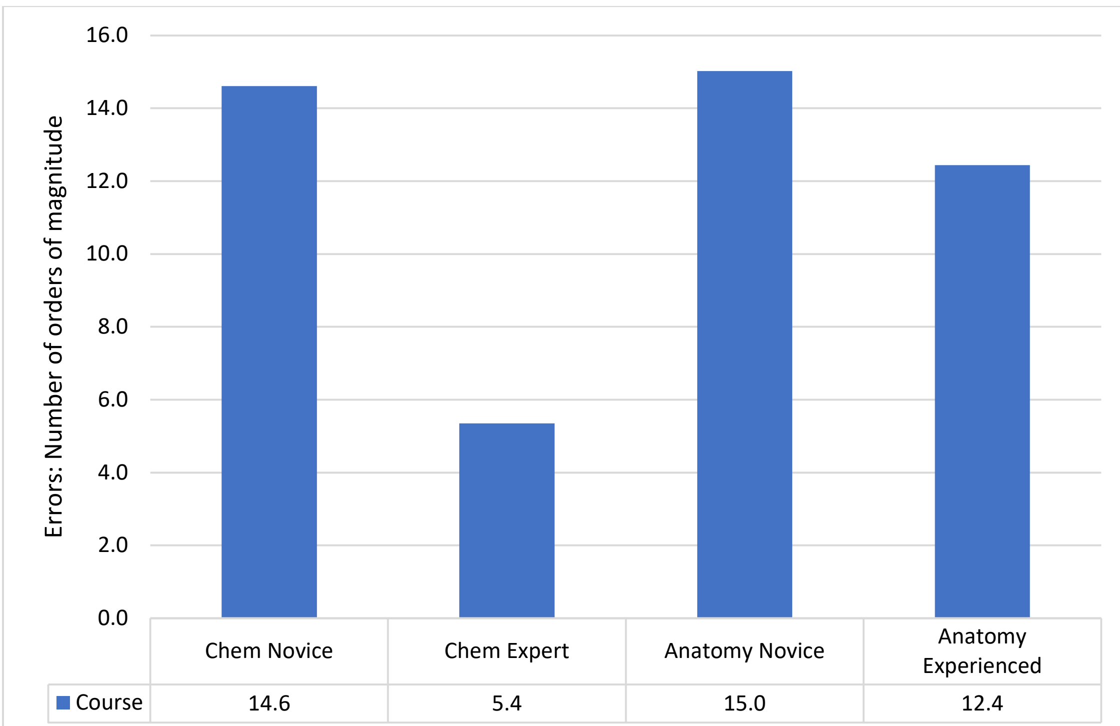



*Figure 4.41 Part IV: Average combined orders of magnitude errors by course*

**Table 4.4 Part IV: Number of orders of magnitude of errors by type**

| | Chem Novice | Chem Experienced | Anatomy and Physiology I Novice | Anatomy and Physiology I Experienced |
|---|---|---|---|---|
| Mean | 14.6 | 5.4 | 15.0 | 12.4 |
| Min | 9 | 1 | 0 | 8 |
| Max | 30.5 | 10.5 | 54.5 | 23 |
| Range | 21.5 | 9.5 | 54.5 | 15 |

# 4.5 Summary and Conclusions

Experienced students demonstrated greater scale ability than novice students. Experienced students were better able to order objects more precisely (relative scaling) and had less errors when putting objects on the number line (absolute scaling).

Novice and experienced students on average created 6.4 and 6.3 bins respectively. More novice students created bins labels using objects than experienced students which supports novice students being more comfortable sizing objects with respect to themselves. During bin creation 100% of experienced students and 72% of novice students created at least one bin greater or less than 3 orders of magnitude from 1 meter which supports the experienced students demonstrating a more developed scale ability than novice students.

Students appear more comfortable with the orders of magnitude surrounding their own size. This can be seen by the larger number of bins (3-4) for the visible region (7 orders of magnitude) while 1-2 bins were created for above 1 km and non-visible items, plus hair width.

Experienced students were more accurate in their object ordering (relative scaling) compared to novice students. Novice and experienced students improved their object ordering when given sizes along with the object name. Students struggled on both the small end as well as the large end of the scale as seen by the difficulty with placing virus, bacterium, and cell as well as cruising altitude of a 747 jet and the width of Wisconsin. For sizes smaller than 1 mm, students tended to place objects larger than their actual size, and objects larger than the average adult height were generally placed smaller than their actual size.

During the absolute scaling activity, novice students gave a variety of reasons for the various ordering of virus, bacterium, and cell. Some reasons students gave, not prompted by the interviewer, were along the lines of "cells make-up everything" with that logic leading to cells having to be the smallest of the three. The difficulty in placing the height of a 747 jet and Wisconsin state width for students could be the fact that they are approaching the edge of their scale knowledge at the height of a 747 jet, 11 km, which is 3 orders of magnitude greater than adult height.

The combined average orders of magnitude errors for both the chemistry novices and anatomy novices were equivalent. Novices in chemistry and anatomy and physiology have more errors than experienced students however Anatomy and Physiology I experienced students had more errors than chemistry experienced students. This could be related to the fact that the chemistry experienced students were all chemistry graduate students while the anatomy and physiology experienced students were not necessarily Anatomy and Physiology I graduate students. The experienced students had a better conception of scale as seen by their less amount of errors throughout the interviews. Experienced students may have had a more hands on experience that aided in their scale conception development that the novice students have yet to experience.

# 4.6 Limitations

One limitation is that the experts were not all Anatomy and Physiology or biological science graduate students. Many of these experienced students were graduate students from different domains, for example, anthropology, working as a laboratory TA for the Anatomy and Physiology I or II courses. Thus, the experienced students may not be representative of actual experienced anatomy and physiology students.

Another limitation is that students did not articulate their thought process throughout the activity. This prevented learning information about why students made decisions in the bin creation, card placement, or number line placement which may have yielded information as to their thought process dealing with scale. The goal of the activity was to understand students' current level of scale conception and not the student thought process regarding scale.

## 4.7 Implications for Instruction

Similar to chemistry novices, Anatomy and Physiology I students have demonstrated limited ability in scale and may benefit from scale instruction. Scale instruction can take place in many forms such as adapting the existing chemistry activity based on this activity to teach scaling as an Anatomy and Physiology I laboratory experiment. Any instructor looking to investigate student understanding of a topic should first conduct interviews to determine at what level students know the material or topic. The interviews may uncover specific areas of the topics that students specifically struggle with, for example novice students struggle with the smaller and larger ends of the scale.

## 4.8 Citations

Gerlach, K.; Trate, J.; Blecking, A.; Geissinger, P.; Murphy, K. (2014). Investigation of Absolute and Relative Scaling Conceptions of Students in Introductory College Chemistry Courses. *J. Chem. Educ. 91*, 1526-1537.

Laubach, C., Royce, C., and Holzer, M. (2000). Teaching to the Power of 10 *The Science Teacher* 66(7): 48-50.

Trate, J. (2017). *Integrating Scale-Themed Instruction Across the General Chemistry Curriculum and Selected In-Depth Studies* (Doctoral dissertation). University of Wisconsin-Milwaukee, Milwaukee, WI.

Tretter, T. R. and Jones, M. G. (2003). A Sense of Scale. *The Science Teacher* 70(1): 22-25.

Tretter, T. R.; Jones, M. G; Andre, T.; Negishi, A; and Minogue, J. (2006). Conceptual Boundaries and Distances: Students' and Experts' Conceptions of the Scale of Scientific Phenomena. *Journal of Research in Science Teaching*, *43*(3), 282-319.

Tretter, T. R.; Jones, M. G.; and Minogue, J. (2006). Accuracy of Scale Conceptions in Science: Mental Maneuverings across many Orders of Spatial Magnitude. *Journal of Research in Science Teaching, 43*(10), *1061*-1085.

# Chapter 5: Building a predictive model for an Anatomy and Physiology I course

## 5.1 Introduction

Scale is a cross-cutting concept that applies to biological sciences as well as chemistry. The interviews done with Anatomy and Physiology I novice and experienced students revealed that Anatomy and Physiology I novice students have a similar scale conception as chemistry novice students (see **Chapter 4**). The interviews also showed that the more experienced a student was at the course content, the more experienced they tended to be in scale (novice vs. experienced anatomy and physiology students). These results support Anatomy and Physiology I as a good candidate for the inclusion of scale as a theme.

Before scale can be integrated as a theme, a baseline must be established for students' scale conception and content knowledge. This allows for the effect of scale integration to be examined. The approach used was building a multiple regression model to predict an end of semester measure. Possible predictive measures include ACT composite and sub-scores, a university math placement exam and sub-scores, and scale measures (SLST, SCI, and SLS). The final measures available were course score, laboratory score, aggregate online quiz score, and aggregate take-home exam score. While these measures provided content measures no one final measure was medium or high stakes, content based, and taken solely at the end of the semester. This led to the creation of a cumulative final exam. This chapter details the development of a cumulative final exam and a multiple regression model predicting a student's cumulative final exam score and scale's role within the model in Anatomy and Physiology I.

## 5.2 Background

### 5.2.1 Test construction

A test is defined as a domain specific evaluation of a student's behavior[81]. A test plan must be developed before writing any items. A test plan includes defining the domain and range of questions, time limit, number of items, item format, and test format. The scope of the test should also be decided such as will the test cover an entire semester or just a few chapters. The scoring method of the test also needs to be established. A test can be scored by totaling up the items or sub-scores may exist depending on test construction.

Once these questions are answered items can be written according to the defined item format[82]. Items should cover the scope of the test, a range of difficulties, and be able to answer the purpose of the test as well as provide a measurement. If sub-sections occur, existing items can be aligned to the subsections, with integrity, or new items are written for each sub-section. Aligning with integrity means aligning a question to its actual sub-section and not to the sub-section one wishes the question would align. If item pools are developed the items should cover the same content and items should have similar complexity. Items should cover a range of complexities to categorize student learning.

### 5.2.2 Bloom's Taxonomy

One method for categorizing learning and understanding is using Bloom's Taxonomy. In 1956 Dr. Benjamin Bloom created a way to organize and classify orders of thinking and learning[83]. Bloom et al. separated educational activities into three domains: cognitive, affective,

[81] Standards for Educational and Psychological Testing 1999: 3
[82] Standards for Educational and Psychological Testing 1999: 37-42
[83] Bloom 1956: 1-15

and psychomotor. The cognitive domain deals with one's knowledge, the affective with feelings and attitudes, and psychomotor with physically completing a task. Focusing on the cognitive domain, Bloom et al. published "Bloom's Taxonomy" as a way of classifying levels of complexity and understanding of a topic, the same way biologists classify animals. Originally the six levels were knowledge, comprehension, application, analysis, synthesis, and evaluation. In 2001 the levels were renamed as listed in figure 5.1 starting with knowledge renamed as remember[84]. Remember is the lowest level of learning and create is the highest.

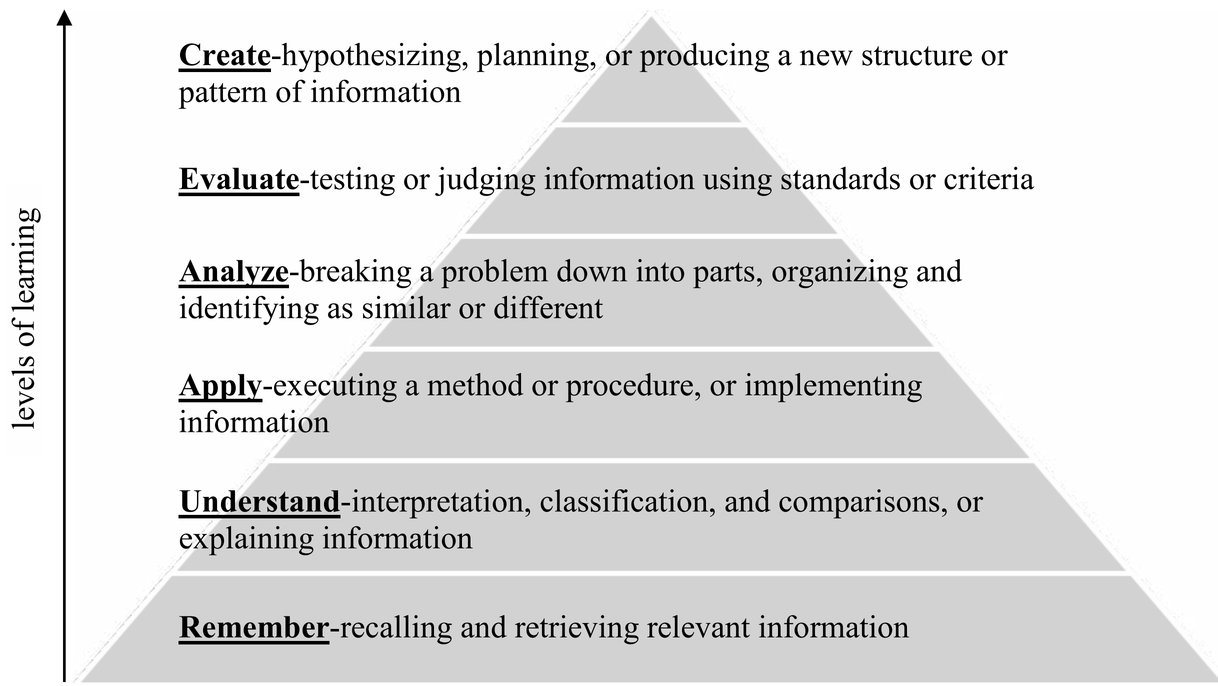



*Figure 5.5 Bloom's Taxonomy Levels*

**Remember**, the lowest level of complexity, refers to recalling and retrieving relevant information. An example is a question asking a student to recall a definition to a term. **Understand** is the next highest level of learning and deals with interpretation, classification, comparisons, or explaining information. An example of understand is when a student is asked to explain a definition of a term in their own words. **Apply** is executing a method or procedure, or

[84] Bloom 1956: 18

implementing information. An example of apply is students solving a math problem with a set sequence of steps. **Analyze** is breaking the problem down into parts, determining how parts are similar or different, or organizing parts. An example of analyze is asking students to identify a theme or predict an outcome. **Evaluate** is testing or judging the information by use of standards and specific criteria. An example of evaluate is asking a student to draw conclusions or modify a plan or experiment. **Create** is generating a hypothesis, designing, planning, or producing a new structure or pattern of information. An example of create is writing a hypothesis and designing a way to test it. Students level of learning and understanding starts with being able to remember the information and progresses up the pyramid as they learn to understand, apply, analyze, evaluate, and create.

The different levels allow instructors to assess students' learning of a topic at different levels of learning and thinking. The aligned questions aligned can generate a sub-score for each level. There are two methods to align questions to levels of learning and thinking. The first is that questions are generated at a particular level; the other is that existing items are aligned, with integrity, to a level. Aligning with integrity means aligning a question to its actual Bloom's Taxonomy level and not to the level the researcher, or instructor, wishes the question would align.

# 5.3 Methods

The regression model was built for student performance in an Anatomy and Physiology I course for the Fall 2017 and Spring 2018 combined semesters. When building a multiple regression model, the possible measures, both independent (predictive) and dependent (final), were identified. The possible predictive measures in this course were the ACT composite score and sub-scores, math placement exam scores and sub-scores, and beginning of semester scale measures (the SLST, SCI and scale literacy score (SLS). The possible final measures in this course were course score, laboratory score, aggregate online quiz score, aggregate take-home exam score, and a cumulative final exam score. Descriptive statistics of ACT composite, and sub-scores, math placement, and sub-scores, scale measures, and cumulative final exam are in **Chapter 2**.

## 5.3.1 Predictive measures

Before attending the university, students submit an ACT composite score and sub-scores with their application. Once accepted, students take a math placement exam which contains three categories: math basics, algebra, and trigonometry. The math placement exam items are different every year while the categories remained the same, until starting spring 2017. In spring 2017 students began taking a revised math placement exam which was sub-scored into three similar, but not identical, categories: math fundamentals (similar to math basics), advanced algebra (previously algebra), and trigonometry and analytic geometry (previously trigonometry alone). The combination of these sub-scores results in a nominal code that specifies which math course a student is eligible to take at the university. For some students it is possible to be accepted to the university without submitting an ACT composite score and sub-scores and/or a math placement and sub-scores.

At the beginning of their anatomy and physiology course, students complete the scale measures, the SCI and SLST, online. As discussed in the methods chapter, four questions were removed during analysis from the SLST because they threatened the validity of the measure, leading to an SLST adjusted score with the threats removed (SLST adj) (**Chapter 2**). The SLST consists of algorithmic questions (questions that required performance of a mathematical equation or process) and conceptual questions (questions requiring students to recall, understand, or apply information). During analysis the algorithmic and conceptual questions can be separated and scored to generate a SLST algorithmic score (SLST AGM) and a SLST conceptual score (SLST CON). Descriptive statistics for the SLST pre, SCI pre, SLS pre, SLST pre adjusted, and SLS pre adjusted can be found in **Chapter 2**. The descriptive statistics for the SLST algorithmic and conceptual are shown in **Table 5.1** for the students who completed beginning of semester measures (ACT composite and sub-scores, math placement and sub-scores, and scale measures and sub-scores) as well as completed the final exam.

**Table 5.1 Descriptive statistics for algorithmic and conceptual scale literacy skills test sub-scores**

| | SLST pre AGM | SLST pre AGM adj | SLST pre CON | SLST pre CON adj |
|---|---|---|---|---|
| *n* | *184* | *184* | *184* | *184* |
| Minimum | 0.000 | 0.000 | 0.156 | 0.172 |
| Maximum | 1.000 | 1.000 | 0.813 | 0.793 |
| Mean | 0.408 | 0.394 | 0.449 | 0.444 |
| Median | 0.385 | 0.333 | 0.438 | 0.414 |
| Mode | 0.231 | 0.333 | 0.438 | .379[a] |
| Std. Deviation | 0.21 | 0.21 | 0.14 | 0.14 |
| Variance | 0.043 | 0.044 | 0.020 | 0.020 |
| Skewness | 0.584 | 0.617 | 0.410 | 0.350 |
| Kurtosis | -0.232 | -0.225 | -0.263 | -0.322 |

[a]Multiple modes exist

## 5.3.2 Final measures

The course score is a weighted average of the various assignments the students completed over the course of the semester. The breakdown is as follows: 20% in-class activities, 20% online assessments, 40% laboratory component, and 20% from take home exams. The in-class activities were attendance, participation and worksheets. The online assessments were the online quizzes and an online activity, and the take home exams consisted of three take home exam scores contributing to the category percent. The laboratory component consisted of weekly lab worksheets, participation, and a midterm and end of semester laboratory practical.

The online assessments category consisted of online quizzes and an online activity. Quizzes were taken online after a student completed the digital textbook chapter reading assignment, approximately two quizzes a week. The questions chosen for the quiz were selected by the online program in an adaptive learning format. As a student completed the assigned reading, they were prompted to answer questions about the reading. The questions presented to the student during the quiz portion of the assignment were influenced by the student's performance on questions within the text while they completed the reading portion of the assignment. For each quiz every student was asked the same number of questions, ranging from 4 to 45 depending on the quiz, but the topics and difficulty of the questions varied depending on how a student performed while answering the in-text questions. Students were allowed three attempts while the quiz was open and at the end of the semester if a student had yet to pass a quiz, they were given a fourth attempt. To pass a quiz, a student had to get a score of 85% or higher. These quizzes were considered low stakes because they were graded as pass or fail. The total quiz scores and an online activity contributed 20% to a student's total course score.

A student's total laboratory score contributed to 40% of a student's total class score. This consisted of performing the experiments, worksheets, and two laboratory practicals. Each TA taught two laboratory periods and students attended lab once a week. The professor would give the TAs an outline of the topic(s) to be covered, and the teaching's assistants would either use an experiment that was already written and available to them via an instructor section of the course website or would write their own. Each of the TAs would prepare their own pre-lab talk for the students.  Students may receive a worksheet or other material which may be already written or developed by the TA.  There was also variation if the TAs required students to remain in lab for the entire 3-hour period or if, once the student's experiment was complete, they could leave. Two laboratory practicals were given by the TAs during the last week of labs, one week before the last week of classes.  The laboratory practical was written by the TAs to take the entire lab period and were based on an outline provided by the instructor.

The Anatomy and Physiology I course consisted of three take-home exams.  These exams were short answer consisting of 4 to 9 questions drawn from a question bank and provided to the students via the course website.  Students were given the same number of questions, but the questions were randomized.  The students submitted their final answers by uploading a document with the test questions and their responses.  Students had approximately one month to complete and upload their answers to each exam.  The exams were graded by the professor based on a rubric, and exams contributed 20% of a student's total class score.  The exams were distributed in September, October, and November and collected in October, November, and December, for fall, and distributed in February, March, April and collected March, April, and May for spring.

Each of the current measures had at least one reason why they would not be an ideal candidate for the dependent variable of a multiple regression model.  The laboratory practical had variability depending on which TA a student had for lab.  This means not every student took the exact same lab practical and, based on the teaching methods of the individual TAs, students may have gotten different laboratory experiences.  Grading of the assignments and practicals was not consistent because the students didn't perform identical practicals.

The online quizzes were taken over the course of the semesters and ranged from 4 to 45 items.  These quizzes also were low stakes as students were given full credit if they scored over

85%.  The take home exams were greater stakes than the quizzes and consisted of students answering similar questions from a question bank.  A downside to the take home exams was that the students were allowed a month to complete the exam and students would have been able to use outside resources, such as the textbook, internet, or each other (despite being instructed to work alone) to answer the exam questions.

The course score is a weighted average of the various scores students have received over the course of the semester.  The potential final measure that had variability in their scoring contribute to the course score as well.  Predicting the course score would include predicting measures that have variability in how they are scored.  The course score takes into account all measures students have completed throughout the entire duration of the semester.  Predicting this measure would mean that the predicted measure is influenced by scores taken at an early point in the semester before students have been instructed in all the topics the course covers.

The existing measures did not provide a measure at the end of the semester that was content based, high stakes, and controlled for exposure time to the student.  This led to the decision to construct a cumulative final exam for the course.  The course professors took the parameters and developed an online cumulative final exam based on content and Bloom's Taxonomy level.  Students were able to complete the two hour, 97-item exam during the last week of the semester in the McGraw Hill Connect online system.  Once the exam was opened, students had to complete the entire exam within the time limit.  The questions were divided into question banks based on Bloom's Taxonomy level and course content chapter.  The items were pulled from multiple question pools.  The number of questions in each question pool is listed in **Table 5.2**.  Question order for each student was the same.  Each student was given two remember questions from chapter 1, then two understand question from chapter 1, etc. The full list is in **Table 5.3**.

**Table 5.2 List of the number of items that are in each question pool by chapter and Bloom's Taxonomy level**

| | Remember | Understand | Apply | Analyze | Evaluate |
|---|---|---|---|---|---|
| Chapter 1 | 13 | 10 | 13 | | |
| Chapter 2 | 18 | 24 | 11 | 3 | |
| Chapter 3 | 23 | 19 | 10 | 6 | |
| Chapter 4 | 27 | 24 | 8 | | |
| Chapter 5 | 18 | 17 | 11 | 4 | 1 |
| Chapter 6 | 19 | 13 | 10 | 5 | |
| Chapter 7 | 22 | 11 | 8 | | |
| Chapter 8 | 26 | 13 | 9 | 1 | |
| Chapter 9 | 21 | 30 | 7 | 4 | |
| Chapter 10 | 15 | 10 | 3 | | |
| Chapter 11 | 21 | 18 | 4 | 1 | |
| Chapter 12 | 18 | 14 | 5 | 2 | |
| Chapter 13 | 24 | 8 | 2 | | |
| Chapter 14 | 19 | 20 | 4 | | |
| Chapter 15 | 38 | 21 | 8 | 2 | |
| Chapter 16 | 26 | 21 | 4 | 10 | |

**Table 5.3 Question order students received by chapter and Bloom's Taxonomy level for the cumulative final exam**

| Chapter-Start | Level | Chapter | Level | Chapter | Level | Chapter | Level |
|---|---|---|---|---|---|---|---|
| 1 | remember | 5 | analyze | 8 | remember | 12 | analyze |
| 1 | remember | 5 | apply | 9 | analyze | 13 | remember |
| 1 | understand | 5 | understand | 9 | apply | 13 | remember |
| 1 | understand | 5 | understand | 9 | remember | 13 | understand |
| 1 | apply | 5 | remember | 9 | remember | 13 | understand |
| 1 | apply | 5 | remember | 9 | understand | 13 | apply |
| 2 | analyze | 6 | analyze | 9 | understand | 13 | apply |
| 2 | apply | 6 | apply | 10 | remember | 14 | remember |
| 2 | understand | 6 | understand | 10 | remember | 14 | remember |
| 2 | understand | 6 | understand | 10 | understand | 14 | understand |
| 2 | remember | 6 | remember | 10 | understand | 14 | understand |
| 2 | remember | 6 | remember | 10 | apply | 14 | apply |
| 3 | analyze | 5 | evaluate | 10 | apply | 14 | apply |
| 3 | apply | 7 | apply | 11 | remember | 15 | analyze |
| 3 | understand | 7 | apply | 11 | remember | 15 | apply |
| 3 | understand | 7 | understand | 11 | understand | 15 | understand |
| 3 | remember | 7 | understand | 11 | understand | 15 | understand |
| 3 | remember | 7 | remember | 11 | apply | 15 | remember |
| 4 | apply | 7 | remember | 11 | analyze | 15 | remember |
| 4 | apply | 8 | analyze | 12 | remember | 16 | remember |
| 4 | understand | 8 | apply | 12 | remember | 16 | remember |
| 4 | understand | 8 | understand | 12 | understand | 16 | understand |
| 4 | remember | 8 | understand | 12 | understand | 16 | understand |
| 4 | remember | 8 | remember | 12 | apply | 16 | apply |
| | | | | | | 16 | analyze |
| | | | | | | | End |

The 97 items the students took were not evenly distributed among Bloom's Taxonomy level. Instead students had the largest number of remember and understand questions with less questions in each category as levels of learning and understanding increased. The Bloom's Taxonomy level and chapter association, for each question was determined by the authors of the online system. Table 5.4 shows the number of questions students answered at each complexity level and gives an example question. The number of questions each student was given from each section is listed in **Table 5.4**.

**Table 5.4 Number of items pulled from each chapter and Bloom Taxonomy level for each student**

| | Remember | Understand | Apply | Analyze | Evaluate |
|---|---|---|---|---|---|
| Chapter 1 | 2 | 2 | 2 | | |
| Chapter 2 | 2 | 2 | 1 | 1 | |
| Chapter 3 | 2 | 2 | 1 | 1 | |
| Chapter 4 | 2 | 2 | 2 | | |
| Chapter 5 | 2 | 2 | 1 | 1 | 1 |
| Chapter 6 | 2 | 2 | 1 | 1 | |
| Chapter 7 | 2 | 2 | 2 | | |
| Chapter 8 | 2 | 2 | 1 | 1 | |
| Chapter 9 | 2 | 2 | 1 | 1 | |
| Chapter 10 | 2 | 2 | 2 | | |
| Chapter 11 | 2 | 2 | 1 | 1 | |
| Chapter 12 | 2 | 2 | 1 | 1 | |
| Chapter 13 | 2 | 2 | 2 | | |
| Chapter 14 | 2 | 2 | 2 | | |
| Chapter 15 | 2 | 2 | 1 | 1 | |
| Chapter 16 | 2 | 2 | 1 | 1 | |
| Total | 32 | 32 | 22 | 10 | 1 |
| Example | Which of the following sets of directional terms are most appropriately referred to as opposites? (Ch. 1) | Inorganic chemists study substances _______ carbon, while organic chemists study substances _______ carbon. (Ch. 2) | A swollen, painful area of the skin that is also hot and red are symptoms that accompany what process? (Ch. 4) | Label the Types of Ossification to the Bone (Ch. 6) | One type of experimental contraceptive device is a skin patch that contains a chemical absorbed through the skin. Which of the following substances would most likely be the type of chemical involved? (Ch. 5) |

Sub-scores were calculated for the final exam based on content or learning or thinking level. The cumulative final counted as an exam score and contributed to the 20% of the class score along with the take home exams. **Figure 5.2** shows the predictive and final measures available. The cumulative final exam score is taken at the end of the semester, is medium stakes, content based, and timed. The final exam is considered medium stakes due to the percentage the exam contributes to a student's overall course score (5%).

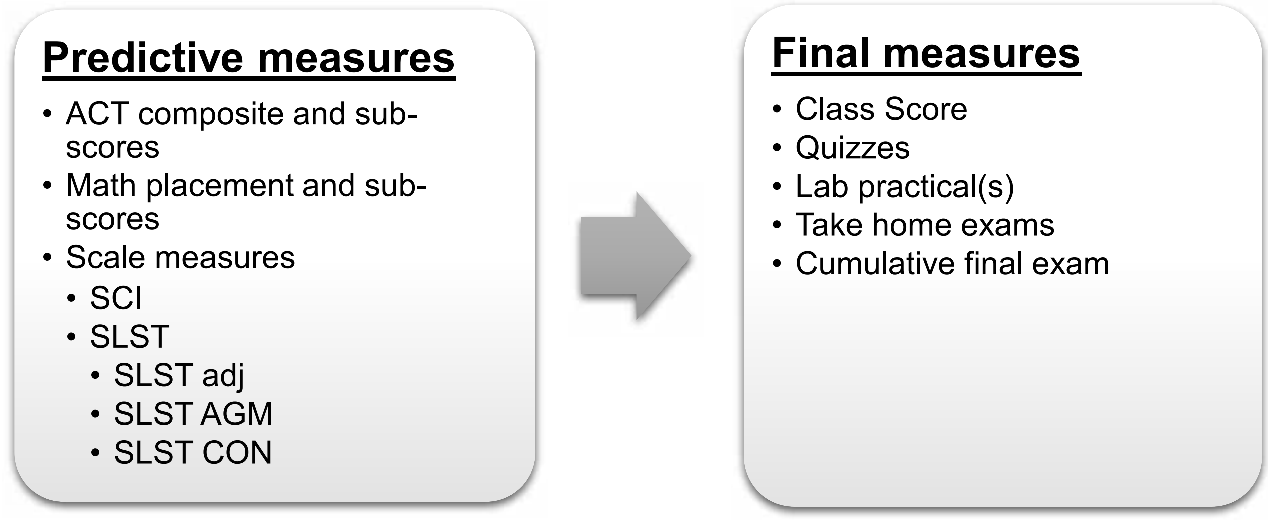



*Figure 5.6 List of predictive and final measures in Anatomy and Physiology I*

### 5.3.3 Multiple regression

Multiple regression is a statistical technique used to predict a dependent variable from two or more predictive, independent, variables. As with all statistical techniques there are assumptions that must be met to properly interpret the results. For multiple regression the assumptions are[85]:

- Independent and dependent variables have a linear relationship
- No multicollinearity (highly correlated independent variables)
- Residuals should be normally distributed
- Homogeneity of variance

Multiple regression is a form of linear regression modeling with two or more variables[86]. The generic equation for a line is Y = MX + B and the equation for multiple regression follows a

[85] Gravetter and Wallnau 2015: 557-581
[86] Berry and Stanley 1985: 9-18

similar format with each variable (X) multiplied by a coefficient (β) plus a constant (α) to determine the dependent variable (Y).

$$Y = \alpha + \beta_1 X_{1j} + \beta_2 X_{2j} + \cdots + \beta_k X_{kj}$$

This format for the equation is why the independent and dependent variables must have a linear relationship. If the relationship is anything but linear then the equation of the line would no longer be accurate.

For the independent variables to each contribute to the calculation of the dependent variable the independent variables should be distinct from one another. The independent variables should have no multicollinearity. Multicollinearity means that two independent variables, in a multiple regression model, have a linear relationship with each other such that one could be predicted or generated from the other. An example would be a measure that directly contributes to a total score. The measure and the score would have multicollinearity. Pearson's correlation can be examined as a potential indicator as multicollinearity as well as knowledge of how measures are collected.

Homogeneity of variance, or homoscedasticity, means that the error in the model, residuals, across the whole range of data is minimized for that model. The distance between all points and the regression line equation is minimized and uniform. Ways to check for this is by determining the sum, average, and normality of the residuals. Residuals are calculated by taking the actual observed dependent value minus the predicted dependent value. The sum, and average, of the residuals should be zero because this means that the regression line is optimized for all data points. The residuals should be normally distributed if the model is optimized. The Kolmogorov-Smirnov test is used to determine the normality of the residuals[87].

When a multiple regression model is computed an $R^2$ value is generated which is the fraction of the variability in the dependent variable that is accounted for by the equation. For example, an $R^2$ value of 0.63 means that 63% of the variance in the dependent variable is explained by the model.

When deciding the best model to use there are a variety of factors such as explaining a good amount of the variance in the dependent measure but also use of predictors, which come out as significant in the model, that make sense to predict the dependent variable. The "best" regression equation will be a balance of these factors. The "best" model will explain a certain amount of the dependent variable. The dependent variable that the regression model predicts should be high stakes for the students, timed, based on the content learned in the course, and at the end of the semester. The dependent variable should be able to show different levels of understanding of course concepts and separate the students by performance level.

# 5.4 Results

The predictive variables can be scale measures (SLST, SCI, and SLS), ACT measures (ACT composite, ACT reading, ACT science, and ACT math), and math placement measures (math basics, algebra, and trigonometry). The first factor for determining potential multiple regression predictors is theory followed by use of the Pearson correlations. The Pearson correlations can tell the direction, and strength of association between two variables. A portion

[87] Osborne and Waters 2002: 1

of the correlation table is shown in **Table 5.5** and **5.6** and a complete correlation table is in **Appendix C**.

**Table 5.5 Pearson correlation (*n* = 184)**

| | Final Exam score | Remember sub-score | Understand sub-score | Apply sub-score | Analyze sub-score |
|---|---|---|---|---|---|
| ACT COMP | .529** | .357** | .525** | .508** | .464** |
| ACT READ | .445** | .273* | .414** | .419** | .366** |
| ACT ENGL | .473** | .295* | .450** | .398** | .365** |
| ACT MATH | .424** | .327** | .473** | .517** | .435** |
| ACT SCIRE | .414** | .350** | .494** | .411** | .462** |
| SLST_Pre | .407** | .438** | .461** | .387** | .436** |
| SCI_Pre | .268** | 0.159 | .347** | .391** | .323** |
| SLS Pre | .424** | .405** | .480** | .434** | .453** |
| SLST_pre_adj | .400** | .433** | .451** | .376** | .433** |
| SLS Pre adj | .419** | .401** | .473** | .426** | .451** |
| SLST pre AGM | .310** | .394** | .388** | .357** | .405** |
| SLST pre AGM adj | .301** | .368** | .366** | .353** | .402** |
| SLST pre CON | .406** | .400** | .436** | .347** | .389** |
| SLST pre CON adj | .390** | .370** | .415** | .322** | .368** |

***Values are significant at the 0.01 level*

Math placement scores are reported differently than ACT scores and sub-scores. ACT scores are provided by the students whereas math placement scores are provided by the university system. The process for storing the math placement sub score values is not consistent across institutions in the system (also varying depending on the location at which the student tested as students can take the placement test at any institution in the system) so some students only had a single number for the math placement score while others also had the sub-scores available. The sample size for those who had a math placement score as well as submitted an

ACT score *and* beginning of semester scale measures is small (*n* = 80) due to the fact that not every student had all of the measures. If math placement was not used for sample limitation, and not in the model, the sample size more than doubles (*n* = 184), and sample size is an important consideration when building the model. If math placement is used in the model, the sample size is low and care must be taken to ensure the sample is representative. If math placement is not a predictor in the model, the model must be examined for both sample sizes to see if the model holds for both sample sizes. The correlation table containing math placement sub-scores, is shown in **Table 5.6**.

**TABLE 5.6 Pearson correlation with math placement sub-scores (*n* = 80)**

| | ALG | TRG | MBSC |
|---|---|---|---|
| ACT COMP | .659** | .633** | .618** |
| ACT READ | .397** | .414** | .425** |
| ACT ENGL | .540** | .492** | .510** |
| ACT MATH | .764** | .765** | .680** |
| ACT SCIRE | .604** | .546** | .529** |
| SLST pre | .466** | .477** | .453** |
| SCI pre | .410** | .365** | .293** |
| SLS pre | .505** | .501** | .460** |
| SLST pre adjusted | .467** | .494** | .465** |
| SLS pre adjusted | .506** | .515** | .471** |
| SLST pre algorithmic | .407** | .404** | .372** |
| SLST pre algorithmic adjusted | .393** | .385** | .342** |
| SLST pre conceptual | .434** | .453** | .437** |
| SLST pre conceptual adjusted | .426** | .474** | .460** |
| Final Exam score | .529** | .482** | .552** |
| Remember sub-score | .538** | .615** | .550** |
| Understand sub-score | .530** | .675** | .488** |
| Apply sub-score | .527** | .638** | .541** |
| Analyze sub-score | 0.361 | .444* | .378* |

**Values are significant at the 0.05 level*
***Values are significant at the 0.01 level*

The predictive variables are separated into scale measures, ACT measures, and math placement measures. After looking at theory, the highest correlating measures in each category provide a starting place for determining a regression model. The Pearson correlation coefficient indicates how two variables are linearly related to one another. A large correlation coefficient between two variables indicates a relationship between the variables and would be a good initial indicator, along with a theoretical basis for inclusion of the predictors, as to what would be a significant predictor in a multiple regression model.

The first reported model was predicting the cumulative final exam using the top predictor in each category: ACT composite ($\beta = 0.458$, $p < 0.001$), math basics ($\beta = 0.205$, $p = 0.056$), and scale literacy score pre adjusted ($\beta = 0.136$, $p = 0.196$). The results of the regression analysis indicated that one predictor was significant, ACT composite, and explained 48.8% of the variance in final exam performance ($R^2 = 0.488$, $F(3,76) = 24.14$, $p < 0.000$). This model is not ideal as only one predictor, ACT composite, was significant. The sample size is also lower, and this is because of including the math placement sub score as a predictor. Running the model without math placement as a condition increases the sample size and ACT composite and scale literacy score pre adjusted are the predictors. The results of the regression analysis for ACT composite ($\beta = 0.427$, $p < 0.001$) and scale literacy score pre adjusted ($\beta = 0.184$, $p = 0.014$) explains 27.0% of the variance in final exam performance ($R^2 = 0.270$, $F(2,181) = 39.36$, $p < 0.001$). Comparing the models for the two sample sizes, both models have ACT composite as a significant predictor however the larger sample size explained less of the variance and found Scale Literacy Score pre adjusted to be a significant predictor.

ACT composite, math basics, and Scale Literacy Skills Test pre adjusted were run as predictors in a model predicting final exam performance. Scale Literacy Skills Test pre adjusted was chosen due to the fact that Scale Literacy Score is an average of the SLST and SCI and of those two components SLST has a higher correlation coefficient to the final exam than SCI. The result was one significant predictor, ACT composite (β = 0.443, $p < 0.001$), with math basics (β = 0.194, $p = 0.068$), and Scale Literacy Skills Test pre adjusted (β = 0.182, $p = 0.076$). The model explained 49.8% of the variance in final exam percentage ($R^2 = 0.498$, F(3,76) = 25.12, $p < 0.001$). Once again there is one model that explains a greater percentage of the variance but only has one significant predictor which is not a math placement sub score. If the model is run with the larger sample size the results were ACT composite (β = 0.438, $p < 0.001$) and Scale Literacy Skills Test pre adjusted (β = 0.176, $p = 0.016$) explaining 30.2% of the variance in final exam percentage ($R^2 = 0.302$, F(2,181) = 39.22, $p < 0.001$). The scale measures can continue to be broken down. Anatomy and physiology contains less math then chemistry so the SLST can be broken down into conceptual and algorithmic questions.

ACT composite, MBSC, and SLST pre conceptual questions adjusted were used as predictors in a model predicting final exam percentage. The result was two significant predictors, ACT composite (β = 0.423, $p < 0.001$) and SLST pre conceptual questions adjusted (β = 0.182, $p = 0.040$), with math basics (β = 0.194, $p = 0.066$). The model explained 50.5% of the variance in final exam percentage ($R^2 = 0.505$, F(3,76) = 25.85, $p < 0.001$). Math placement is not a significant predictor so the model was run with the larger sample size. The result was two significant predictors, ACT composite (β = 0.446, $p < 0.001$) and SLST pre conceptual questions adjusted (β = 0.158, $p = 0.032$). The model explained 29.8% of the variance in the final exam percentage (R2 = 0.298, F(2,181) = 38.33, $p < 0.001$).

Using the model with ACT composite and SLST pre conceptual questions adjusted, with a sample size of 184, the assumptions were checked. The linear relationship between independent and dependent variables as well as the lack of multicollinearity can be shown by the Pearson correlations in **Table 5.6**. While most of the correlations are significant, predictors were categorized by what they measured and whether sub-scores existed. Only one predictor from a category could be used at a time. For example, ACT sub-scores directly contribute to the ACT composite score and so an ACT sub score and ACT composite would not be used in a regression model simultaneously. Therefore, ACT composite and ACT sub-scores were grouped as a category. The normality of residuals was checked with the Kolmogorov-Smirnov test (D(184) = 0.054, $p$ = 0.200) and was found to not be significant meaning the data is normally distributed. This is also reflected in the Q-Q plot, **Figure 5.4**, and histogram, **Figure 5.5**. The sum and average of the residuals for a model should be close to or exactly zero. The sum of the residuals is $5.97\times10^{-13}$ and the average of the residuals is $2.48\times10^{-15}$ which are extremely close to zero.

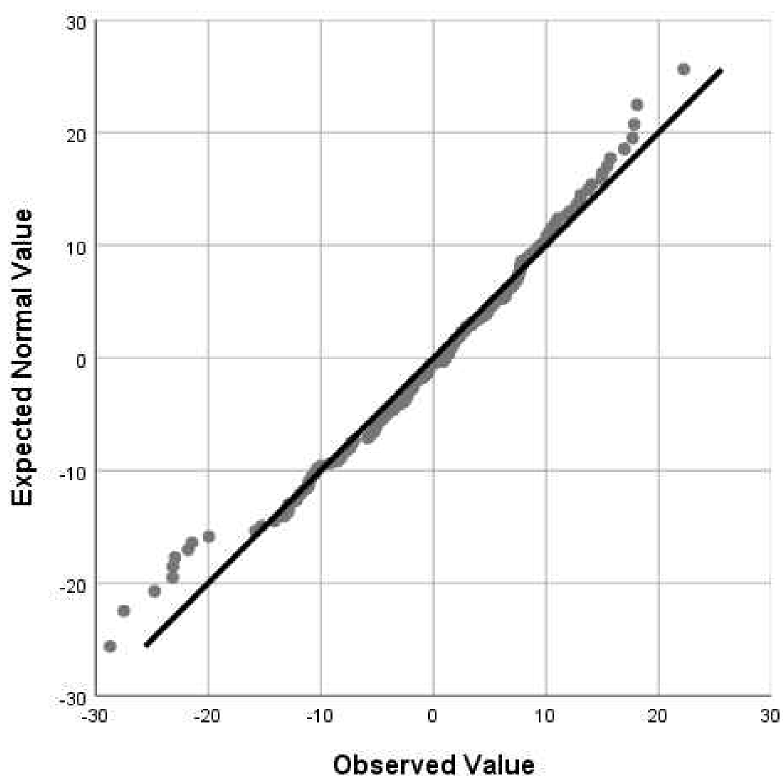



*Figure 5.7 Q-Q plot for the multiple regression residuals*

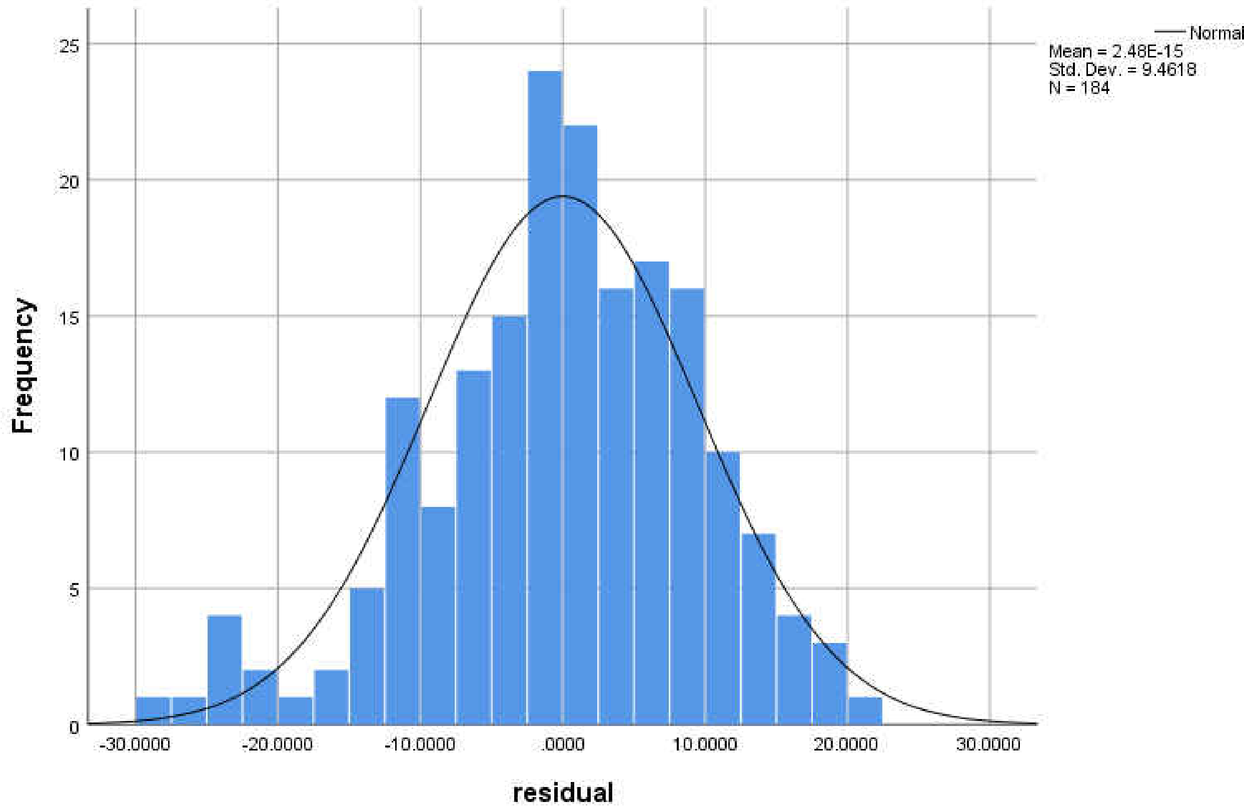



*Figure 5.8 Histogram for multiple regression residuals*

# 5.5 Summary and Conclusions

Anatomy and Physiology I has no prerequisites, therefore the potential predictive measures were those required to enter the university and those imbedded in the coursework at the beginning of the semester. These predictive measures included ACT composite and sub-scores, math placement and sub-scores, and scale measures (SLST, SCI, SLS). Similar to results in chemistry, ACT composite and a scale measure were found to be significant predictors in the model[88].

[88] Trate 2017: 17-33

None of the existing measures provided a content based measure at the end of the semester taken by all students. A cumulative final exam was built based on Bloom's taxonomy and book chapter. The cumulative final exam provided a way to measure students' learning of the course material at the end of the semester. Using Bloom's Taxonomy to establish learning levels allowed another layer of analysis. Sub-scores were generated for each Bloom's level and provided information not only at the chapter content covered but the depth of a student's learning within each chapter.

The best model is the one that explains the greatest amount of variance while also containing logical predictors and meeting all assumptions. The best multiple regression model generated for Anatomy and Physiology I predicted 50.5% (for $n$ = 80) and 29.8% (for $n$ = 184) percent of the variance in cumulative final exam score using ACT composite score and SLST pre conceptual questions adjusted as predictors. These measures are significantly correlating at the beginning of the semester to final exam performance and have shown to be significant predictors in other courses in addition to Anatomy and Physiology I[89]. There is a difference depending on the sample which may be attributed to the diversity of the course in the majors that take the course and the fact that the course has no pre-requisites. The model provides a way to measure differences across semesters such as integrating scale into the course. A scale measure appearing as a significant predictor of final exam performance supports the case for integrating scale as a theme into the Anatomy and Physiology I course.

[89] Trate 2017: 17-33

## 5.6 Limitations

Anatomy and Physiology I had a limited number of predictive and final measures available. Because of this, there is the possibility that a better multiple regression model exists for this sample using different measures not currently used in the course. Another limitation is that the heterogeneous population influenced by being a course commonly taken by freshmen coming from a diverse range of high schools and the course having no prerequisites. Because of the high diversity in students' final model may be the best model possible with the sample and measures available.

With the math placement exam changing yearly, and with the sub-score redistribution, this measure was not a good candidate for use as a predictive factor. The way the math placement scores are stored limited the sample size. Math placement may be a significant predictor but previous math placement sections are unavailable for future students in the course. If the math placement score was more stable across semesters, or years, this may lead to a better predictive model.

## 5.7 Implications for Instruction

Course performance in Anatomy and Physiology I can be predicted using scale and standard predictive measures. If an instructor is looking to build a multiple regression model for their course, they should look at the current available predictors to determine if the measures are representative of a student at the beginning of the semester. Another predictive measure an instructor could utilize is the SLST and SCI, after checking reliability and validity for their population, as another predictive measure. The instructor also needs to determine what final measure(s) would make a good dependent variable that is at the end of the semester, high stakes, and content based.  If an instructor prefers they may build a final exam.  One method is by implementing a variety of question levels based off of Bloom's Taxonomy as well as course content.

# 5.8 Citations


Berry, W.D. and Stanley, F. (1985). Multiple Regression in Practice (Series: Quantitative Applications in the Social Sciences). Thousand Oaks, CA. SAGE Publications, INC.

Bloom, B.S. (Ed.). Engelhart, M.D., Furst, E.J., Hill, W.H., Krathwohl, D.R. (1956). Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain. New York: David McKay Co Inc.

Gravetter, F. J. and Wallnau, L. B. (2015). Statistics for the Behavioral Sciences. Boston, MA: Cengage Learning. Ch. 6, p. 557-581.

Osborne, J. W. and Waters, E. (2002). Four Assumptions of Multiple Regression that Researchers Should Always Test. *Practical Assessment, Research & Evaluation*. *8*(2), 1-5.

Standards for Educational and Psychological Testing. (1999). Washington, D.C., American Educational Research Association. p. 3.

Trate, J. (2017). *Integrating Scale-Themed Instruction across the General Chemistry Curriculum and Selected In-Depth Studies* (Doctoral dissertation). University of Wisconsin-Milwaukee, Milwaukee, WI.

# Appendix A: Anatomy and Physiology I Student Majors

**Table A.1 Anatomy and Physiology I student majors**

| Major | Percent | Major | Percent |
|---|---|---|---|
| Accounting | 0.27 | Health Care Admin (Int) | 2.16 |
| Anthropology (Int) | 0.14 | Health Sciences (Int) | 3.24 |
| Architectural Studies | 0.14 | History (Int) | 0.14 |
| Art History (Int) | 0.14 | Information Sci and Technology | 0.68 |
| Biochemistry | 1.22 | Inter Arts (Int) | 0.14 |
| Biological Sci | 6.62 | Kinesiology (Int) | 10 |
| Biomedical Engineering | 3.92 | Management Info Systems (Int) | 0.14 |
| Biomedical Sciences (Int) | 16.21 | Mathematics (Int) | 0.27 |
| Business (Int) | 1.49 | Mechanical Engineering (Int) | 0.27 |
| Chemistry (Int) | 0.94 | Microbiology | 0.54 |
| Classics | 0.14 | Music Education | 0.27 |
| Comm Sci & Dis (Int) | 3.38 | No major listed | 3.65 |
| Committee Interdis (Int) | 0.14 | Nursing (Int) | 19.46 |
| Communication (Int) | 0.14 | Nursing Collaborative | 0.14 |
| Computer Science | 0.94 | Nutrition (Int) | 3.24 |
| Conservation Sci (Int) | 0.14 | Occupational Studies (Int) | 1.76 |
| Criminal Justice | 0.4 | Philosophy (Int) | 0.14 |
| Education | 1.49 | Political Science (Int) | 0.4 |
| Electrical Engineering (Int) | 0.14 | Psychology | 1.89 |
| English (Int) | 0.27 | Social Work | 0.27 |
| Exceptional Education (Int) | 0.4 | Sociology (Int) | 0.27 |
| Film | 0.14 | Spanish | 0.14 |
| Finance (Int) | 0.14 | Undecided (Int) | 11.89 |
| | | Total | 100 |

# Appendix B:  Scale Assessments

- B.1:  Anatomy and Physiology I Scale Literacy Skills Test
    - B.1.1:  Scale Literacy Skills Test pre-administration item statistics
    - B.1.2:  Scale Literacy Skills Test post-administration item statistics

- B.2:  Anatomy and Physiology I Scale Concept Inventory
    - B.2.1:  Scale Concept Inventory pre-administration item statistics
    - B.2.2:  Scale Concept Inventory post-administration item statistics

- B.3:  Anatomy and Physiology I Laboratory Survey Items
    - B.3.1:  Laboratory Survey items
    - B.3.2:  Laboratory Survey pre-administration item statistics
    - B.3.3:  Laboratory Survey post-administration item statistics

# B.1 Anatomy and Physiology I Scale Literacy Skills Test

## B.1.1 Scale Literacy Skills Test pre-administration item statistics

**Table B.1 Anatomy and Physiology I Scale Literacy Skills Test pre-administration item statistics (*n* = 833)**

| Item | Key | DIF | Discr | %A | %B | %C | %D | Attr A | Attr B | Attr C | Attr D |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | C | 0.813 | 0.282 | 5.6 | 10.2 | 81.3 | 2.9 | -0.09 | -0.10 | 0.28 | -0.09 |
| 2 | B | 0.617 | 0.310 | 14.9 | 61.7 | 14.0 | 9.4 | 0.03 | 0.31 | -0.21 | -0.13 |
| 3 | A | 0.385 | 0.285 | 38.5 | 44.1 | 6.1 | 11.3 | 0.29 | -0.07 | -0.08 | -0.13 |
| 4 | A | 0.222 | 0.354 | 22.2 | 64.0 | 8.5 | 5.3 | 0.35 | -0.12 | -0.16 | -0.07 |
| 5 | A | 0.140 | 0.253 | 14.0 | 12.0 | 3.5 | 70.5 | 0.25 | 0.03 | 0.02 | -0.29 |
| 6 | A | 0.336 | 0.469 | 33.6 | 51.0 | 4.8 | 10.6 | 0.47 | -0.26 | -0.04 | -0.17 |
| 7 | D | 0.376 | 0.498 | 3.4 | 20.0 | 39.0 | 37.6 | -0.06 | -0.12 | -0.31 | 0.50 |
| 8 | D | 0.364 | 0.552 | 37.0 | 22.7 | 4.0 | 36.4 | -0.18 | -0.29 | -0.07 | 0.55 |
| 9 | D | 0.264 | 0.321 | 38.4 | 31.3 | 3.8 | 26.4 | 0.01 | -0.27 | -0.05 | 0.32 |
| 10 | C | 0.367 | 0.426 | 41.9 | 8.4 | 36.7 | 13.0 | -0.33 | -0.13 | 0.43 | 0.04 |
| 11 | A | 0.532 | 0.552 | 53.2 | 11.8 | 27.0 | 8.0 | 0.55 | -0.21 | -0.22 | -0.12 |
| 12 | B | 0.588 | 0.523 | 2.4 | 58.8 | 24.2 | 14.5 | -0.06 | 0.52 | -0.29 | -0.17 |
| 13 | B | 0.678 | 0.310 | 15.5 | 67.8 | 8.2 | 8.5 | -0.11 | 0.31 | -0.10 | -0.10 |
| 14 | B | 0.309 | 0.274 | 39.6 | 30.9 | 9.1 | 20.4 | 0.09 | 0.27 | -0.14 | -0.23 |
| 15 | D | 0.357 | 0.404 | 20.2 | 17.0 | 27.1 | 35.7 | -0.08 | -0.23 | -0.09 | 0.40 |
| 16 | C | 0.389 | 0.296 | 13.7 | 28.5 | 38.9 | 19.0 | -0.14 | -0.22 | 0.30 | 0.07 |
| 17 | B | 0.236 | 0.173 | 30.5 | 23.6 | 41.5 | 4.3 | -0.30 | 0.17 | 0.16 | -0.03 |
| 18 | B | 0.235 | 0.191 | 7.9 | 23.5 | 31.9 | 36.6 | -0.03 | 0.19 | 0.01 | -0.17 |
| 19 | A | 0.323 | 0.271 | 32.3 | 19.1 | 38.8 | 9.8 | 0.27 | -0.06 | -0.22 | 0.02 |
| 20 | B | 0.468 | 0.361 | 24.6 | 46.8 | 24.6 | 4.0 | -0.12 | 0.36 | -0.18 | -0.06 |
| 21 | D | 0.382 | 0.325 | 19.3 | 13.9 | 28.6 | 38.2 | 0.01 | -0.16 | -0.17 | 0.32 |
| 22 | C | 0.459 | 0.509 | 14.5 | 26.3 | 45.9 | 13.3 | -0.17 | -0.22 | 0.51 | -0.12 |
| 23 | B | 0.318 | 0.256 | 34.3 | 31.8 | 15.1 | 18.7 | 0.04 | 0.26 | -0.15 | -0.14 |
| 24 | A | 0.836 | 0.264 | 83.6 | 4.4 | 5.0 | 7.0 | 0.26 | -0.09 | -0.11 | -0.06 |
| 25 | C | 0.148 | -0.040 | 8.5 | 68.1 | 14.8 | 8.6 | -0.07 | 0.16 | -0.04 | -0.04 |
| 26 | B | 0.070 | 0.007 | 10.3 | 7.0 | 57.7 | 25.0 | -0.08 | 0.01 | 0.13 | -0.05 |
| 27 | D | 0.574 | 0.375 | 3.4 | 19.7 | 19.6 | 57.4 | -0.07 | -0.14 | -0.16 | 0.38 |
| 28 | C | 0.395 | 0.502 | 46.1 | 9.1 | 39.5 | 5.3 | -0.26 | -0.14 | 0.50 | -0.09 |
| 29 | C | 0.382 | 0.444 | 24.8 | 32.3 | 38.2 | 4.7 | -0.27 | -0.16 | 0.44 | -0.01 |
| 30 | C | 0.411 | 0.458 | 12.8 | 31.6 | 41.1 | 14.5 | -0.18 | -0.33 | 0.46 | 0.05 |
| 31 | B | 0.618 | 0.361 | 8.3 | 61.8 | 26.8 | 3.1 | -0.12 | 0.36 | -0.22 | -0.03 |
| 32 | A | 0.499 | 0.170 | 49.9 | 40.7 | 6.8 | 2.5 | 0.17 | -0.04 | -0.11 | -0.03 |
| 33 | C | 0.334 | 0.314 | 2.8 | 27.3 | 33.4 | 36.6 | -0.06 | -0.22 | 0.31 | -0.03 |
| 34 | C | 0.205 | 0.116 | 14.0 | 31.1 | 20.5 | 34.3 | -0.10 | -0.12 | 0.12 | 0.11 |
| 35 | A | 0.204 | -0.018 | 20.4 | 23.5 | 24.4 | 31.7 | -0.02 | 0.13 | -0.02 | -0.08 |
| 36 | C | 0.400 | 0.350 | 9.2 | 40.1 | 40.0 | 10.7 | -0.10 | -0.28 | 0.35 | 0.04 |
| 37 | C | 0.806 | 0.365 | 4.9 | 11.5 | 80.6 | 3.0 | -0.08 | -0.25 | 0.36 | -0.02 |
| 38 | B | 0.779 | 0.347 | 3.7 | 77.9 | 8.0 | 10.3 | -0.07 | 0.35 | -0.18 | -0.10 |
| 39 | A | 0.598 | 0.542 | 59.8 | 17.3 | 14.8 | 8.2 | 0.54 | -0.22 | -0.22 | -0.09 |
| 40 | B | 0.521 | 0.321 | 8.9 | 52.1 | 27.4 | 11.6 | -0.13 | 0.32 | -0.07 | -0.11 |
| 41 | B | 0.729 | 0.440 | 1.7 | 72.9 | 4.7 | 20.8 | -0.05 | 0.44 | -0.10 | -0.29 |
| 42 | A | 0.533 | 0.289 | 53.3 | 19.3 | 19.8 | 7.6 | 0.29 | -0.16 | -0.06 | -0.06 |
| 43 | A | 0.359 | 0.202 | 35.9 | 47.3 | 10.6 | 6.2 | 0.20 | -0.01 | -0.11 | -0.08 |
| 44 | A | 0.629 | 0.264 | 62.9 | 7.7 | 17.6 | 11.8 | 0.26 | -0.12 | -0.11 | -0.04 |
| 45 | A | 0.349 | 0.310 | 34.9 | 13.4 | 29.8 | 21.8 | 0.31 | -0.16 | -0.09 | -0.05 |

## B.1.2: Scale Literacy Skills Test post-administration item statistics

**Table B.2 Anatomy and Physiology I Scale Literacy Skills Test post-administration item statistics (*n* = 783)**

| Item | Key | DIF | Discr | %A | %B | %C | %D | Attr A | Attr B | Attr C | Attr D |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | C | 0.789 | 0.350 | 6.8 | 10.3 | 78.9 | 4.0 | -0.13 | -0.17 | 0.35 | -0.05 |
| 2 | B | 0.637 | 0.195 | 14.3 | 63.7 | 15.5 | 6.5 | 0.08 | 0.19 | -0.19 | -0.08 |
| 3 | A | 0.405 | 0.245 | 40.5 | 41.9 | 7.5 | 10.1 | 0.25 | -0.05 | -0.07 | -0.12 |
| 4 | A | 0.267 | 0.408 | 26.7 | 59.8 | 8.9 | 4.6 | 0.41 | -0.17 | -0.16 | -0.09 |
| 5 | A | 0.178 | 0.303 | 17.8 | 11.7 | 6.3 | 64.2 | 0.30 | 0.02 | -0.03 | -0.29 |
| 6 | A | 0.386 | 0.462 | 38.6 | 47.0 | 5.5 | 8.9 | 0.46 | -0.26 | -0.08 | -0.13 |
| 7 | D | 0.390 | 0.520 | 3.3 | 22.2 | 35.5 | 39.0 | -0.08 | -0.18 | -0.26 | 0.52 |
| 8 | D | 0.360 | 0.516 | 33.5 | 25.2 | 5.4 | 36.0 | -0.08 | -0.35 | -0.08 | 0.52 |
| 9 | D | 0.248 | 0.357 | 39.5 | 28.2 | 7.5 | 24.8 | 0.04 | -0.27 | -0.13 | 0.36 |
| 10 | C | 0.392 | 0.285 | 39.3 | 9.6 | 39.2 | 11.9 | -0.13 | -0.17 | 0.29 | 0.01 |
| 11 | A | 0.496 | 0.484 | 49.6 | 14.2 | 27.3 | 8.9 | 0.48 | -0.22 | -0.18 | -0.08 |
| 12 | B | 0.596 | 0.502 | 5.4 | 59.6 | 24.4 | 10.6 | -0.13 | 0.50 | -0.24 | -0.13 |
| 13 | B | 0.682 | 0.339 | 15.7 | 68.2 | 9.8 | 6.3 | -0.13 | 0.34 | -0.11 | -0.09 |
| 14 | B | 0.369 | 0.199 | 36.3 | 36.9 | 11.7 | 15.1 | 0.15 | 0.20 | -0.19 | -0.16 |
| 15 | D | 0.312 | 0.379 | 20.1 | 20.7 | 28.1 | 31.2 | -0.04 | -0.26 | -0.09 | 0.38 |
| 16 | C | 0.374 | 0.235 | 12.1 | 32.6 | 37.4 | 17.9 | -0.06 | -0.25 | 0.23 | 0.08 |
| 17 | B | 0.240 | 0.011 | 29.2 | 24.0 | 40.1 | 6.6 | -0.13 | 0.01 | 0.12 | 0.00 |
| 18 | B | 0.267 | 0.101 | 9.2 | 26.7 | 33.0 | 31.2 | -0.02 | 0.10 | 0.01 | -0.09 |
| 19 | A | 0.275 | 0.310 | 27.5 | 23.6 | 37.9 | 11.0 | 0.31 | -0.11 | -0.15 | -0.05 |
| 20 | B | 0.489 | 0.383 | 25.3 | 48.9 | 21.8 | 4.0 | -0.14 | 0.38 | -0.16 | -0.08 |
| 21 | D | 0.358 | 0.339 | 19.9 | 16.1 | 28.2 | 35.8 | -0.07 | -0.13 | -0.14 | 0.34 |
| 22 | C | 0.496 | 0.469 | 12.5 | 25.8 | 49.6 | 12.1 | -0.08 | -0.21 | 0.47 | -0.18 |
| 23 | B | 0.359 | 0.209 | 28.7 | 35.9 | 16.7 | 18.6 | 0.06 | 0.21 | -0.16 | -0.12 |
| 24 | A | 0.820 | 0.274 | 82.0 | 6.9 | 4.5 | 6.6 | 0.27 | -0.13 | -0.10 | -0.04 |
| 25 | C | 0.160 | -0.090 | 11.7 | 65.4 | 16.0 | 6.9 | -0.06 | 0.18 | -0.09 | -0.02 |
| 26 | B | 0.349 | 0.101 | 10.3 | 34.9 | 31.2 | 23.6 | -0.14 | 0.10 | 0.04 | 0.00 |
| 27 | D | 0.524 | 0.300 | 5.5 | 23.0 | 19.2 | 52.4 | -0.09 | -0.10 | -0.11 | 0.30 |
| 28 | C | 0.411 | 0.451 | 39.0 | 13.3 | 41.1 | 6.6 | -0.10 | -0.24 | 0.45 | -0.11 |
| 29 | C | 0.434 | 0.397 | 17.9 | 31.5 | 43.4 | 7.2 | -0.23 | -0.09 | 0.40 | -0.07 |
| 30 | C | 0.418 | 0.321 | 13.8 | 31.8 | 41.8 | 12.6 | -0.18 | -0.21 | 0.32 | 0.08 |
| 31 | B | 0.594 | 0.375 | 6.1 | 59.4 | 29.8 | 4.7 | -0.09 | 0.38 | -0.23 | -0.05 |
| 32 | A | 0.475 | 0.220 | 47.5 | 40.1 | 8.9 | 3.4 | 0.22 | -0.04 | -0.13 | -0.05 |
| 33 | C | 0.392 | 0.227 | 5.4 | 24.6 | 39.2 | 30.8 | -0.12 | -0.20 | 0.23 | 0.10 |
| 34 | C | 0.199 | 0.018 | 12.5 | 39.5 | 19.9 | 28.1 | -0.08 | -0.19 | 0.02 | 0.26 |
| 35 | A | 0.213 | 0.000 | 21.3 | 24.9 | 26.7 | 27.1 | 0.00 | 0.03 | 0.05 | -0.07 |
| 36 | C | 0.434 | 0.310 | 8.3 | 38.8 | 43.4 | 9.5 | -0.10 | -0.24 | 0.31 | 0.04 |
| 37 | C | 0.778 | 0.386 | 5.2 | 13.5 | 77.8 | 3.4 | -0.10 | -0.25 | 0.39 | -0.03 |
| 38 | B | 0.746 | 0.394 | 5.9 | 74.6 | 9.8 | 9.7 | -0.12 | 0.39 | -0.19 | -0.08 |
| 39 | A | 0.593 | 0.527 | 59.3 | 19.4 | 15.1 | 6.3 | 0.53 | -0.23 | -0.24 | -0.05 |
| 40 | B | 0.538 | 0.321 | 9.8 | 53.8 | 28.2 | 8.2 | -0.10 | 0.32 | -0.17 | -0.06 |
| 41 | B | 0.686 | 0.480 | 3.2 | 68.6 | 9.1 | 19.2 | -0.07 | 0.48 | -0.19 | -0.21 |
| 42 | A | 0.524 | 0.347 | 52.4 | 18.0 | 19.2 | 10.5 | 0.35 | -0.18 | -0.11 | -0.05 |
| 43 | A | 0.418 | 0.379 | 41.8 | 40.2 | 12.8 | 5.2 | 0.38 | -0.12 | -0.19 | -0.08 |
| 44 | A | 0.622 | 0.444 | 62.2 | 9.5 | 17.5 | 10.9 | 0.44 | -0.14 | -0.18 | -0.13 |
| 45 | A | 0.310 | 0.242 | 31.0 | 16.9 | 30.9 | 21.2 | 0.24 | -0.22 | 0.02 | -0.03 |

# B.2 Anatomy and Physiology I Scale Concept Inventory

## B.2.1 Scale Concept Inventory pre-administration item statistics

**Table B.3 Anatomy and Physiology I Scale Concept Inventory pre-administration item statistics (*n* = 590)**

| Item | Key | %A | %B | %C | %D | %E | Positive(%) | Negative(%) |
|---|---|---|---|---|---|---|---|---|
| 1 | + | 20.8 | 40.5 | 5.3 | 24.9 | 8.5 | 61.4 | 33.4 |
| 2 | - | 1.7 | 12.7 | 24.6 | 44.2 | 16.8 | 14.4 | 61.0 |
| 3 | - | 14.9 | 29.0 | 13.2 | 34.2 | 8.6 | 43.9 | 42.9 |
| 4 | + | 6.8 | 32.4 | 25.4 | 27.8 | 7.6 | 39.2 | 35.4 |
| 5 | - | 2.7 | 16.8 | 11.5 | 45.8 | 23.2 | 19.5 | 69.0 |
| 6 | + | 15.1 | 34.1 | 12.7 | 31.4 | 6.8 | 49.2 | 38.1 |
| 7 | + | 8.1 | 27.3 | 39.7 | 18.8 | 6.1 | 35.4 | 24.9 |
| 8 | + | 23.1 | 38.1 | 18.0 | 14.6 | 6.3 | 61.2 | 20.8 |
| 9 | | 7.5 | 32.5 | 29.2 | 25.9 | 4.9 | 40.0 | 30.8 |
| 10 | + | 8.0 | 32.4 | 41.7 | 15.6 | 2.4 | 40.3 | 18.0 |
| 11 | + | 10.0 | 34.7 | 26.4 | 24.9 | 3.9 | 44.7 | 28.8 |
| 12 | + | 9.8 | 40.8 | 12.0 | 29.2 | 8.1 | 50.7 | 37.3 |
| 13 | + | 8.0 | 33.4 | 42.9 | 13.6 | 2.2 | 41.4 | 15.8 |
| 14 | - | 3.4 | 19.5 | 16.1 | 41.7 | 19.3 | 22.9 | 61.0 |
| 15 | + | 20.8 | 50.8 | 9.2 | 15.9 | 3.2 | 71.7 | 19.2 |
| 16 | + | 7.6 | 40.0 | 27.8 | 22.0 | 2.5 | 47.6 | 24.6 |
| 17 | - | 19.0 | 37.3 | 19.0 | 20.3 | 4.4 | 56.3 | 24.7 |
| 18 | | 6.1 | 28.8 | 20.2 | 34.9 | 10.0 | 34.9 | 44.9 |
| 19 | - | 4.4 | 19.5 | 36.8 | 30.0 | 9.3 | 23.9 | 39.3 |
| 20 | + | 10.7 | 40.0 | 30.0 | 18.0 | 1.4 | 50.7 | 19.3 |
| 21 | + | 9.3 | 38.1 | 16.8 | 31.2 | 4.6 | 47.5 | 35.8 |
| 22 | + | 42.5 | 48.3 | 4.7 | 3.7 | 0.7 | 90.8 | 4.4 |
| 23 | - | 2.9 | 15.8 | 21.2 | 45.9 | 14.2 | 18.6 | 60.2 |
| 24 | | 4.1 | 23.7 | 23.2 | 36.3 | 12.7 | 27.8 | 49.0 |
| 25 | + | 11.0 | 36.9 | 19.2 | 28.1 | 4.7 | 48.0 | 32.9 |
| 26 | + | 6.3 | 26.9 | 27.6 | 33.7 | 5.4 | 33.2 | 39.2 |
| 27 | **V** | 48.1 | 51.9 | 0.0 | 0.0 | 0.0 | 100.0 | 0.0 |
| 28 | + | 15.6 | 38.3 | 28.1 | 16.6 | 1.4 | 53.9 | 18.0 |
| 29 | + | 5.3 | 17.6 | 18.1 | 49.8 | 9.2 | 22.9 | 59.0 |
| 30 | + | 3.2 | 15.4 | 37.1 | 32.2 | 12.0 | 18.6 | 44.2 |
| 31 | - | 4.4 | 21.5 | 23.9 | 37.8 | 12.4 | 25.9 | 50.2 |
| 32 | - | 12.7 | 56.9 | 15.9 | 11.4 | 3.1 | 69.7 | 14.4 |
| 33 | + | 10.2 | 49.3 | 18.3 | 18.6 | 3.6 | 59.5 | 22.2 |
| 34 | - | 8.3 | 36.6 | 41.7 | 11.0 | 2.4 | 44.9 | 13.4 |
| 35 | - | 10.8 | 52.4 | 19.8 | 13.2 | 3.7 | 63.2 | 16.9 |
| 36 | + | 11.9 | 51.5 | 21.5 | 13.2 | 1.9 | 63.4 | 15.1 |
| 37 | - | 3.7 | 20.2 | 14.2 | 49.2 | 12.7 | 23.9 | 61.9 |
| 38 | + | 5.8 | 27.8 | 26.8 | 34.7 | 4.9 | 33.6 | 39.7 |
| 39 | + | 7.8 | 31.4 | 42.4 | 16.6 | 1.9 | 39.2 | 18.5 |
| 40 | - | 12.2 | 48.5 | 18.3 | 15.9 | 5.1 | 60.7 | 21.0 |

## B.2.2 Scale Concept Inventory post-administration item statistics

**Table B.4 Anatomy and Physiology I Scale Concept Inventory post-administration item statistics (*n* = 395)**

| Item | Key | %A | %B | %C | %D | %E | Positive(%) | Negative(%) |
|---|---|---|---|---|---|---|---|---|
| 1 | + | 23.3 | 37.5 | 5.3 | 25.1 | 8.9 | 60.8 | 33.9 |
| 2 | - | 2.5 | 11.1 | 20.5 | 43.0 | 22.8 | 13.7 | 65.8 |
| 3 | - | 14.4 | 31.6 | 15.4 | 29.6 | 8.9 | 46.1 | 38.5 |
| 4 | + | 10.4 | 33.7 | 24.8 | 24.3 | 6.8 | 44.1 | 31.1 |
| 5 | - | 1.0 | 13.7 | 11.9 | 43.3 | 30.1 | 14.7 | 73.4 |
| 6 | + | 18.0 | 31.9 | 13.7 | 28.6 | 7.8 | 49.9 | 36.5 |
| 7 | + | 10.9 | 30.4 | 32.2 | 20.3 | 6.3 | 41.3 | 26.6 |
| 8 | + | 30.4 | 37.0 | 19.7 | 9.6 | 3.3 | 67.3 | 12.9 |
| 9 | | 7.6 | 28.9 | 27.6 | 28.9 | 7.1 | 36.5 | 35.9 |
| 10 | + | 10.6 | 35.9 | 33.4 | 17.0 | 3.0 | 46.6 | 20.0 |
| 11 | + | 10.4 | 32.4 | 23.5 | 27.3 | 6.3 | 42.8 | 33.7 |
| 12 | + | 11.9 | 33.4 | 11.4 | 34.4 | 8.9 | 45.3 | 43.3 |
| 13 | + | 9.6 | 36.7 | 34.2 | 17.5 | 2.0 | 46.3 | 19.5 |
| 14 | - | 4.1 | 18.5 | 14.7 | 39.7 | 23.0 | 22.5 | 62.8 |
| 15 | + | 21.0 | 51.1 | 9.4 | 14.2 | 4.3 | 72.2 | 18.5 |
| 16 | + | 8.9 | 37.7 | 23.0 | 25.3 | 5.1 | 46.6 | 30.4 |
| 17 | - | 24.3 | 40.8 | 15.7 | 14.4 | 4.8 | 65.1 | 19.2 |
| 18 | | 5.8 | 26.3 | 19.7 | 37.2 | 10.9 | 32.2 | 48.1 |
| 19 | - | 5.3 | 23.3 | 28.6 | 30.6 | 12.2 | 28.6 | 42.8 |
| 20 | + | 12.9 | 45.1 | 22.5 | 17.0 | 2.5 | 58.0 | 19.5 |
| 21 | + | 12.9 | 37.2 | 19.0 | 28.6 | 2.3 | 50.1 | 30.9 |
| 22 | + | 44.1 | 47.3 | 5.1 | 3.0 | 0.5 | 91.4 | 3.5 |
| 23 | - | 2.5 | 18.2 | 24.8 | 36.5 | 18.0 | 20.8 | 54.4 |
| 24 | | 5.8 | 27.1 | 15.7 | 38.0 | 13.4 | 32.9 | 51.4 |
| 25 | + | 10.6 | 37.0 | 20.5 | 26.3 | 5.6 | 47.6 | 31.9 |
| 26 | + | 5.8 | 30.4 | 27.8 | 31.1 | 4.8 | 36.2 | 35.9 |
| 27 | **V** | 52.2 | 47.8 | 0.0 | 0.0 | 0.0 | 100.0 | 0.0 |
| 28 | + | 13.9 | 38.5 | 25.3 | 18.5 | 3.8 | 52.4 | 22.3 |
| 29 | + | 6.6 | 19.2 | 18.7 | 43.5 | 11.9 | 25.8 | 55.4 |
| 30 | + | 5.3 | 19.0 | 35.7 | 29.9 | 10.1 | 24.3 | 40.0 |
| 31 | - | 5.6 | 21.5 | 18.2 | 39.5 | 15.2 | 27.1 | 54.7 |
| 32 | - | 12.7 | 52.4 | 19.5 | 12.7 | 2.8 | 65.1 | 15.4 |
| 33 | + | 11.1 | 45.1 | 22.0 | 18.7 | 3.0 | 56.2 | 21.8 |
| 34 | - | 10.9 | 35.2 | 35.2 | 14.7 | 4.1 | 46.1 | 18.7 |
| 35 | - | 10.4 | 46.8 | 21.0 | 16.2 | 5.6 | 57.2 | 21.8 |
| 36 | + | 12.9 | 46.1 | 27.6 | 11.9 | 1.5 | 59.0 | 13.4 |
| 37 | - | 4.1 | 17.2 | 14.7 | 50.9 | 13.2 | 21.3 | 64.1 |
| 38 | + | 5.6 | 27.3 | 24.6 | 36.7 | 5.8 | 32.9 | 42.5 |
| 39 | + | 8.6 | 37.7 | 32.2 | 19.2 | 2.3 | 46.3 | 21.5 |
| 40 | - | 15.4 | 45.8 | 16.5 | 18.2 | 4.1 | 61.3 | 22.3 |

# B.3: Anatomy and Physiology I Laboratory Survey Items

## B.3.1: Laboratory Survey items

**Table B.5 Anatomy and Physiology I Laboratory Survey Items[a]**

| Objective items |
| --- |
| • If the balance reads 0.1053 grams, you should record 0.11 grams in your notebook. |
| • Using objects to make estimations of length is as accurate as using a ruler. |
| • Precision of a measurement can be estimated by calculating a standard deviation. |
| • A reasonable reduction in mass for your experiment is 130%. |
| • Using the highest magnification will give you the best view of a sample. |
| • A percent difference calculation reveals an error of ~50%, this tells you that your experimental value is off by a factor of 2 from the accepted value. |
| • If an original solution contains 15% glucose and the amount of glucose is doubled, the new solution would be 30% glucose. |
| • The same magnification can be used to view most cells under a microscope. |
| • Increasing the number of measurements decreases the amount of error associated with that measurement. |
| • Percent error or percent difference calculations tell you the degree to which your experimental value differs from your second experimental value. |
| • Adding more water while making a solution would result in a higher calculated concentration. |
| • Microscopes are used to view features that are not visible to the naked eye. |
| **Subjective items** |
| • I expect the lab will help reinforce the concepts taught in lecture. |
| • I expect to understand things better on the cellular level because of lab. |
| • I don't think I will learn anything in lab. |
| • Lab will be a helpful component to this course for demonstrating important concepts. |
| • The laboratory activities will help me learn lecture concepts that are unable to be demonstrated in a classroom setting. |
| • I expect my understanding of the microscopic nature of matter will be increased by the laboratory activities. |
| • I don't expect the laboratory activities to match well with the lecture topics. |
| **Verification item** |
| • Of lab and lecture, lab gives the greatest opportunity to collect experimental data. |

[a]Items for "pre" survey are shown. "Post" items are identical except for the addition of past tense language – for example "I don't think I will learn anything in lab." was changed to "I didn't learn anything in lab."

## B.3.2: Laboratory Survey pre-administration item statistics

**Table B.6 Anatomy and Physiology I Laboratory Survey pre-administration item statistics (*n* = 914)**

| Item | Key | %A | %B | %C | %D | %E | Omit | Positive (%) | Negative (%) |
|---|---|---|---|---|---|---|---|---|---|
| 1 | + | 65.3 | 33.6 | 1.0 | 0.1 | 0.0 | 0.00 | 98.9 | 0.1 |
| 2 | + | 50.8 | 45.5 | 3.4 | 0.2 | 0.0 | 0.11 | 96.3 | 0.2 |
| 3 | - | 6.9 | 25.4 | 20.9 | 34.0 | 12.6 | 0.22 | 32.3 | 46.6 |
| 4 | - | 1.3 | 6.3 | 10.3 | 40.0 | 41.6 | 0.44 | 7.7 | 81.6 |
| 5 | + | 5.8 | 35.1 | 48.0 | 7.2 | 3.1 | 0.77 | 40.9 | 10.3 |
| 6 | - | 0.1 | 0.2 | 2.4 | 25.2 | 71.7 | 0.44 | 0.3 | 96.8 |
| 7 | - | 0.1 | 3.6 | 55.1 | 26.4 | 14.1 | 0.66 | 3.7 | 40.5 |
| 8 | - | 3.0 | 11.7 | 20.8 | 49.8 | 13.1 | 1.64 | 14.7 | 62.9 |
| 9 | + | 1.4 | 18.2 | 62.0 | 15.5 | 2.2 | 0.66 | 19.6 | 17.7 |
| 10 | + | 51.0 | 43.8 | 3.8 | 1.4 | 0.0 | 0.00 | 94.7 | 1.4 |
| 11 | - | 15.6 | 46.5 | 18.3 | 15.8 | 3.7 | 0.11 | 62.1 | 19.5 |
| 12 | - | 1.4 | 17.9 | 19.0 | 48.9 | 12.6 | 0.11 | 19.4 | 61.5 |
| 13 | + | 49.0 | 48.6 | 1.6 | 0.7 | 0.0 | 0.11 | 97.6 | 0.7 |
| 14 | + | 17.1 | 39.5 | 26.1 | 15.1 | 2.0 | 0.22 | 56.6 | 17.1 |
| 15 | **V** | 47.0 | 53.0 | 0.0 | 0.0 | 0.0 | 0.00 | 100.0 | 0.0 |
| 16 | + | 7.2 | 46.2 | 33.8 | 10.5 | 1.8 | 0.55 | 53.4 | 12.3 |
| 17 | - | 1.5 | 8.2 | 18.7 | 48.5 | 23.0 | 0.11 | 9.7 | 71.4 |
| 18 | + | 37.2 | 56.9 | 4.3 | 0.8 | 0.3 | 0.55 | 94.1 | 1.1 |
| 19 | - | 0.4 | 3.2 | 10.3 | 54.4 | 31.5 | 0.22 | 3.6 | 85.9 |
| 20 | + | 59.2 | 36.5 | 3.0 | 1.0 | 0.1 | 0.22 | 95.7 | 1.1 |

## B.3.3: Laboratory Survey post-administration item statistics

**Table B.7 Anatomy and Physiology I Laboratory Survey post-administration item statistics (*n* = 768)**

| Item | Key | %A | %B | %C | %D | %E | Omit | Positive (%) | Negative (%) |
|---|---|---|---|---|---|---|---|---|---|
| 1 | + | 32.0 | 46.5 | 9.1 | 7.7 | 4.4 | 0.26 | 78.5 | 12.1 |
| 2 | + | 29.0 | 50.5 | 12.6 | 6.1 | 1.3 | 0.39 | 79.6 | 7.4 |
| 3 | - | 9.4 | 27.6 | 18.1 | 26.2 | 18.5 | 0.26 | 37.0 | 44.7 |
| 4 | - | 3.1 | 10.3 | 14.6 | 34.5 | 37.5 | 0.00 | 13.4 | 72.0 |
| 5 | + | 7.4 | 38.2 | 36.7 | 12.8 | 4.7 | 0.26 | 45.6 | 17.4 |
| 6 | - | 1.0 | 2.7 | 9.6 | 34.4 | 52.0 | 0.26 | 3.8 | 86.3 |
| 7 | - | 5.2 | 7.0 | 38.4 | 29.7 | 19.1 | 0.52 | 12.2 | 48.8 |
| 8 | - | 5.1 | 13.9 | 25.0 | 39.3 | 15.0 | 1.69 | 19.0 | 54.3 |
| 9 | + | 2.7 | 20.2 | 51.6 | 20.7 | 3.6 | 1.17 | 22.9 | 24.3 |
| 10 | + | 31.6 | 48.2 | 10.8 | 6.9 | 2.3 | 0.13 | 79.8 | 9.2 |
| 11 | - | 15.6 | 44.8 | 18.4 | 16.9 | 4.0 | 0.26 | 60.4 | 21.0 |
| 12 | - | 5.7 | 21.0 | 14.1 | 45.6 | 13.7 | 0.00 | 26.7 | 59.2 |
| 13 | + | 31.4 | 55.2 | 8.5 | 4.3 | 0.5 | 0.13 | 86.6 | 4.8 |
| 14 | + | 19.4 | 38.8 | 21.0 | 16.1 | 4.7 | 0.00 | 58.2 | 20.8 |
| 15 | **V** | 44.9 | 55.1 | 0.0 | 0.0 | 0.0 | 0.00 | 100.0 | 0.0 |
| 16 | + | 7.9 | 42.2 | 31.9 | 13.2 | 4.4 | 0.39 | 50.1 | 17.6 |
| 17 | - | 7.2 | 11.1 | 12.5 | 41.3 | 28.0 | 0.00 | 18.2 | 69.3 |
| 18 | + | 24.2 | 54.4 | 16.1 | 4.6 | 0.4 | 0.26 | 78.6 | 4.9 |
| 19 | - | 8.7 | 19.9 | 19.8 | 37.2 | 14.2 | 0.13 | 28.6 | 51.4 |
| 20 | + | 53.4 | 38.3 | 3.5 | 3.1 | 1.7 | 0.00 | 91.7 | 4.8 |

# Appendix C: General Chemistry II scale-themed supplemental instruction activities

- C.1: Activity 1 (solutions)
- C.2: Activity 2 (fuel cells)

# C.1: Activity 1 (solutions)

**Initial Questions**:

**Table C.1 Activity 1 (solutions) initial questions**

| Sections | Number of Questions in Pool | Number of Questions Pulled |
|---|---|---|
| Heating and Cooling Curves | 5 | 1 |
| Intermolecular Forces | 8 | 1 |
| Phase Change | 5 | 1 |
| Phase Diagram | 5 | 1 |
| Solution Amounts | 8 | 1 |
| Intermolecular Forces in Solutions | 7 | 1 |
| Vapor Pressure Lowering | 5 | 1 |
| Boiling Point Elevation | 6 | 1 |
| Phase Diagrams of Solutions | 5 | 1 |
| Total | 54 questions | 9 questions |

Scoring (1 point for each answered correctly):

>75% (7-9 questions) – Scenario 3

50-75% (5-6 questions) – Scenario 2

<50% (0-4 questions) – Scenario 1

---

**Scenario 1**:

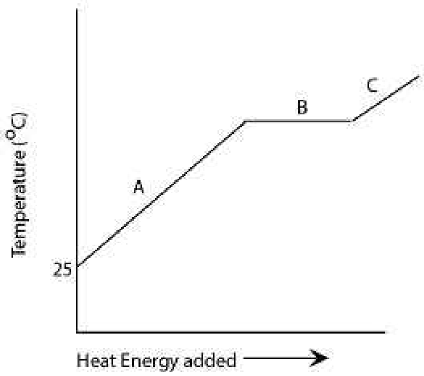
**Introduction: You go into your kitchen planning to make rice. You find your roommate left a measuring cup of a clear, colorless liquid (unknown liquid) right next to your measuring cup of water. You decide to boil both (in separate pots) to observe if there are differences.**


**Unknown liquid** **Water**

**Please use the hints provided as they are designed to help you with answering the questions. Any time you see a definition, you will find the definition and related information in the hint. Good luck!**

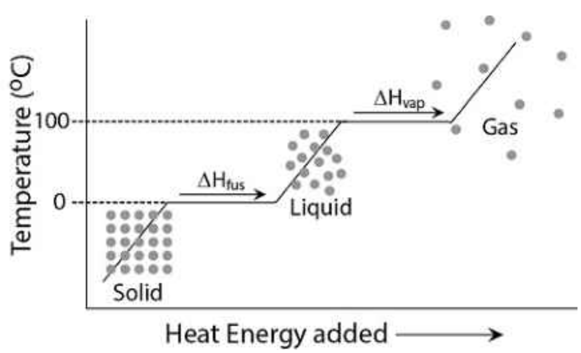
**You slowly heat both liquids while plotting temperature of the liquid over time and generate a heating curve for each substance. You notice these graphs look very similar to ones you've seen in your chemistry class and remember that you can get a lot of information about a substance from a plot such as this.**

1. Identify where boiling is occurring on the heating curve generated for water.

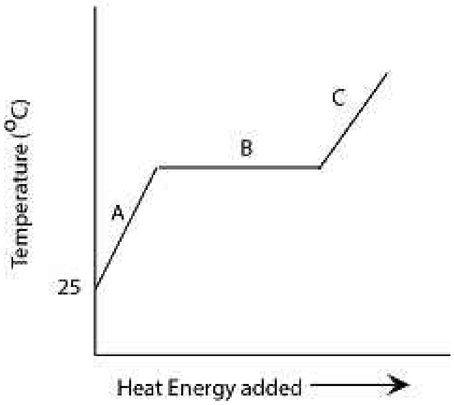



A.
B.
C.

**Hint:** A heating/cooling curve is generated by plotting the temperature of a substance over time. If a constant heat source is used, the amount of time that passes is equivalent to the amount of heat added to the substance. During a phase change, the temperature of the substance does not change as all of the added heat energy is being used to overcome the forces holding the particles together. Once the phase change is complete, the temperature of the substance will once again rise.

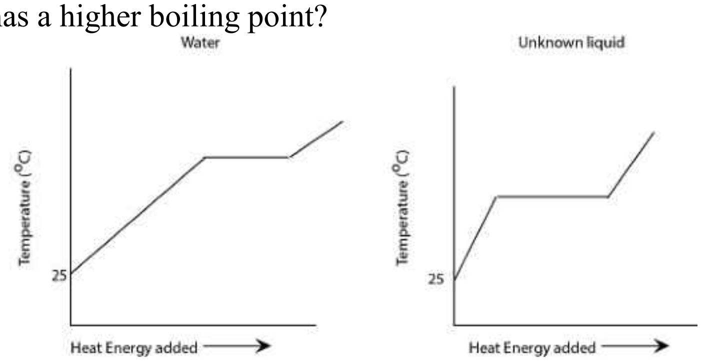



2. Identify where boiling is occurring on the heating curve generated for the unknown liquid.




A.
B.
C.

3. Which substance has a higher boiling point?

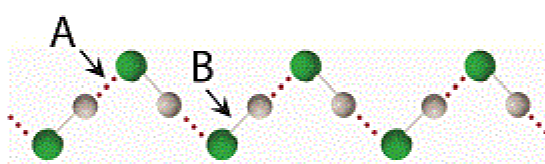



a. Unknown liquid

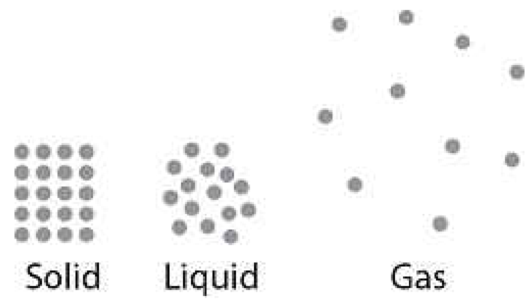
b. Water

**Now that you know the unknown liquid has a lower boiling point than water, you start to think about what particle level properties both of these liquids exhibit and how those properties relate to their relative boiling points.**

4. Using this particle level diagram, which letter designates what is overcome to boil a substance?




**Hint:** Recall that the difference between a solid, liquid, and gas is only that the distance between the particles in each phase of matter has increased. To answer this question, think about the factors that determine the state of matter of a given substance at room temperature. [Image: Hint phases of matter.jpg]




5. Is the strength of intermolecular forces of water equal to that of the unknown liquid?
   A. Yes

   B. No

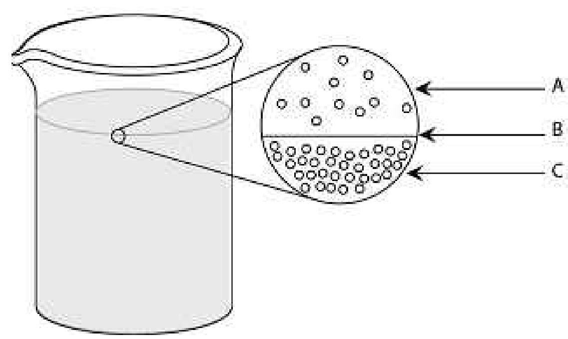
6. Explain why your previous answer is correct.

7. Which substance has stronger intermolecular forces: water or the unknown liquid?

8. Explain why your previous answer is correct.

**Since water has a higher boiling point than the unknown liquid, you are certain that means water has stronger intermolecular forces than the unknown liquid. You also remember from chemistry class that all liquids have vapor pressure, but start to wonder how intermolecular forces affect the quantity of vapor particles that exist above your two liquids.**

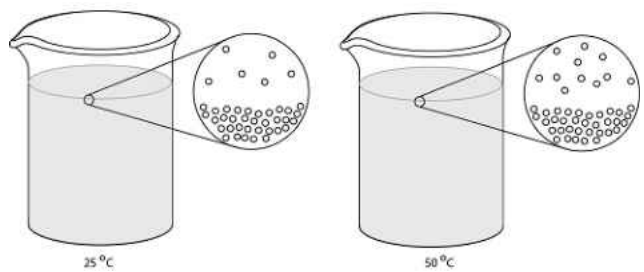
9. On the diagram, select the letter corresponding to where vapor pressure is measured.




A.
B.
C.

**Hint:** The vapor pressure of a liquid is the equilibrium pressure of a vapor above (and by extension exerted on) its liquid. At the surface of every liquid a dynamic equilibrium of condensation and evaporation is established. As individual molecules gain sufficient kinetic energy they escape the liquid phase into the vapor phase. Molecules in the vapor phase rejoin the liquid phase as they interact with the surface of the liquid and lose sufficient amounts of kinetic energy. Factors that affect the rate of each of these processes can include: temperature, pressure, and intermolecular forces.

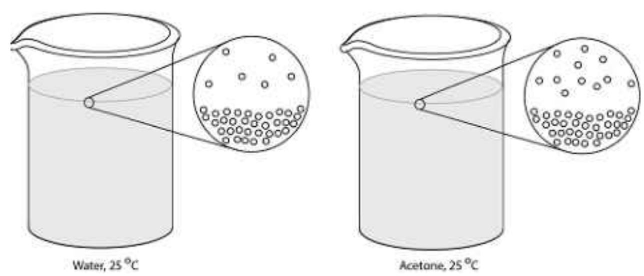
10. The figure shows the same liquid on the particle level at different temperatures. Based on the figure, as the temperature of a liquid increases, the vapor pressure:




A. Increases

B. Decreases

11. The figure shows different liquids on the particle level at the same temperature. Based on the figure, as intermolecular forces of pure substances increase, the vapor pressure:




A. Increases

B. Decreases

12. Using your answers to numbers 10 and 11 explain the relationship between temperature, vapor pressure, and intermolecular forces.

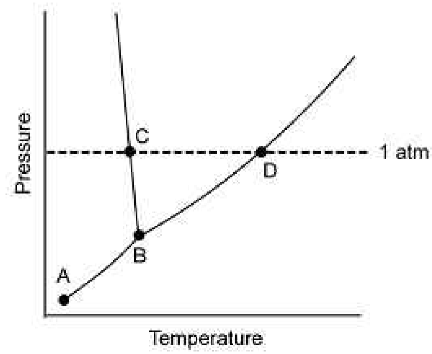
**Knowing now that water has a lower vapor pressure than the unknown liquid, you want to understand how having a lower vapor pressure means water requires more energy (i.e. a higher temperature) than the unknown liquid to boil.**

13. What must be true of the vapor pressure and the external pressure before a liquid will boil?

14. Explain your answer to number 13. Make sure to include why this must happen before boiling can be observed.

**Hint:** As liquid molecules gain sufficient kinetic energy to overcome the forces holding them in the liquid phase, the molecules enter the vapor phase. When a liquid is heated to near its boiling point, this process does not only happen at the surface of the liquid and water vapor "bubbles" can be observed forming along the bottom of the pan. If the external pressure is greater than the vapor pressure these bubbles cannot rise to the surface and boiling is not observed.

15. Which letter on the phase diagram corresponds to the normal boiling point of a liquid?



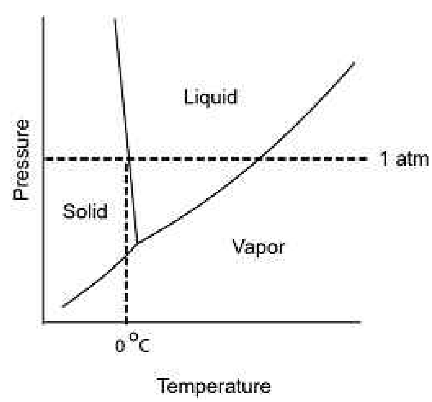

A. A

B. B

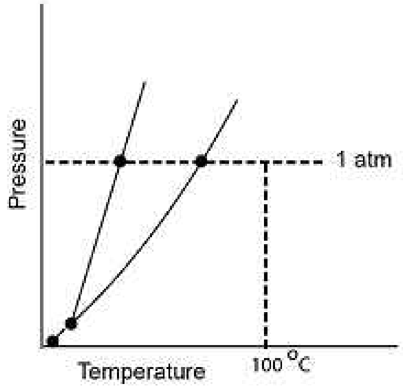
C. C

D. D

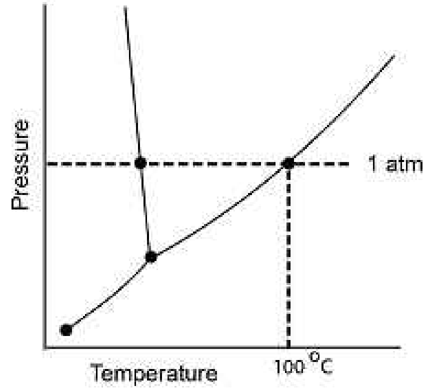
**Hint:** A **phase diagram** summarizes the conditions under which a substance exists as a solid, liquid, or gas. The diagram is divided into three regions, each of which represents a pure phase, and the line separating any two regions indicates conditions under which these two phases can exist in equilibrium. The **normal boiling point** of a liquid is the boiling point when the external pressure is 1 atm. Similarly, as illustrated in the diagram, the **normal freezing/melting point** is located on the solid-liquid boundary at 1 atm of pressure.



16. Which phase diagram corresponds to the unknown liquid?

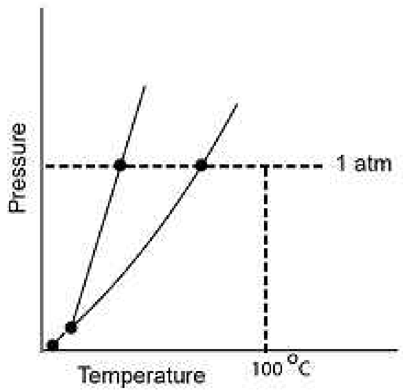



Phase Diagram 1

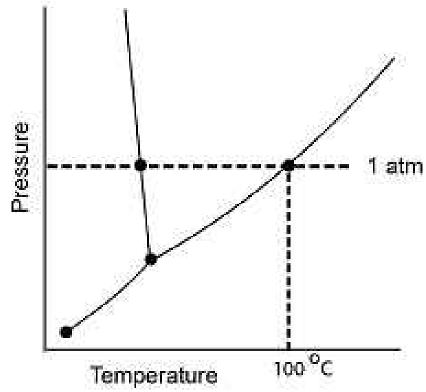



Phase Diagram 2

17. Which phase diagram corresponds to water?




Phase Diagram 1




Phase Diagram 2

18. Which liquid are you going to use to make your rice?


A. Unknown liquid

B. Water

Scoring (1 point for each answered correctly):

≥60% (11-18 questions) – Scenario 1 Questions

<60% (0-10 questions) – repeat with a note to check the hints provided

---

**<u>Scenario 1 Questions</u>**:

**Table C.2 Activity 1 (solutions) scenario 1 questions**

| Section | Number of Questions in Pool | Number of Questions Pulled |
|---|---|---|
| Heating and Cooling Curves | 5 | 1 |
| IMF | 8 | 2 |
| Phase Change | 5 | 1 |
| Phase Diagram | 5 | 1 |
| Total | 23 questions | 5 |

Scoring (1 point for each answered correctly):

100% (5 questions) – Scenario 3

≥50-99% (3-4 questions) – Scenario 2

<50% (0-2 questions) – repeat

---

**<u>Scenario 2</u>**:

**<u>Introduction</u>: You are planning to make rice using a recipe that calls for a 2:1 ratio of water to rice. You measure out 2 cups of water and pour it in the pot. As you add a teaspoon of salt to the water and start the heat, you think about the ways solutions are different than pure substances, like water.**

**Please use the hints provided as they are designed to help you with answering the questions. Any time you see a definition you will find the definition and any other relevant information in the hint. Good luck!**

1. If a teaspoon of salt weights 5 g and a metric cup is equal to 250 mL, what is the molar concentration of the salt solution in the pot? <u>Report your answer to 5 significant figures.</u> MW NaCl = 58.44 g/mol

2. If the molality (*m*) of the solution is actually 0.16974 *m*, what is the density of the solution (in g•mL$^{-1}$)?

The molar concentration of a solution ($M$) is the number of moles of solute in the volume of the solution (in L):

$$\text{molar concentration } (M) = \frac{\text{mol solute}}{\text{volume solution (L)}}$$

The molal concentration ($m$) of a solution is the number of moles of solute in the mass of the solvent (in kg):

$$\text{molal concentration } (m) = \frac{\text{mol solute}}{\text{mass solvent (kg)}}$$

To use this, let's consider a 1.000 M sodium hydroxide solution that also has a molal concentration of 1.028 *m*. Using molal concentration, we can then determine the density of the solution:

First: We know that 40.00 g of sodium hydroxide was used to make 1.000 L of solution. We can use that to calculate the mass of water:

$$40.00 \text{ g solute}\left(\frac{\text{mol solute}}{40.00 \text{ g solute}}\right)\left(\frac{\text{kg solvent}}{1.028 \text{ mol solute}}\right) = 0.9724 \text{ kg solvent}$$

Second: Now that we have the mass of water and sodium hydroxide, the sum is the mass of the solution:

$$40.00 \text{ g solute} + 972.4 \text{ g solvent} = 1012.4 \text{ g solution}$$

Lastly: Density is the mass of the solution in the volume of solution:

**Hint**:

$$\text{Density} = \frac{1012.4 \text{ g solution}}{1000.0 \text{ mL solution}} = 1.0124 \text{ g}\cdot\text{mL}^{-1}$$

3. You go to the fridge looking for something to drink while you are cooking and see your roommate's container of juice. Select all of the possible concentration units for the container of juice you found.

a. ppm
b. %v/v
c. g
d. g/mol
e. mL
f. g/mL

**After adding the salt you notice that your new solution doesn't appear to look any differently than it did before you added the salt. You can no longer see grains of salt in your pot of water so you know that on the symbolic and particle levels your solution would have to be represented differently to show what has happened. You've been studying for an upcoming chemistry exam and decide to test yourself first on symbolic representations.**

4. What is the best symbolic representation for your salt solution?
   a. $Na^+$(*aq*) and $Cl^-$(*aq*)
   b. Na(*s*) and $Cl_2$(*g*)
   c. Na(*aq*) and $Cl_2$(*aq*)
   d. NaCl(*l*)

**Hint:** Symbolic representations use chemical equations and symbols to give you important information about a substance or a reaction. For example, the symbolic representation for ice is $H_2O$(*s*). If you are given only this information, you could model this substance on both the macroscopic and particle levels.

To correctly represent a solution symbolically, you must first determine if the solute is an **electro lyte** or **nonelectrolyte**. Strong electrolytes completely dissociate in solution, breaking apart to form cations and anions. Nonelectrolytes do not dissociate when dissolved in solution (will not form ions - there will not be ions in solution). An example symbolic representation for both an electrolyte solution (KOH) and a nonelectrolyte solution ($C_6H_{12}O_6$) are shown below.

Potassium hydroxide is a strong electrolyte for which an aqueous solution could be represented as:

KOH(*aq*), but would **more correctly** be represented as $K^+$(*aq*) and $OH^-$(*aq*).

Glucose is a nonelectrolyte for which an aqueous solution could **ONLY** be represented as:

$C_6H_{12}O_6$(*aq*)

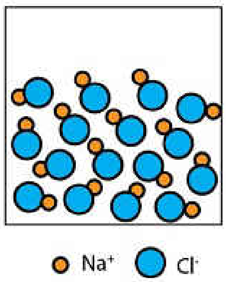
5. Methanol ($CH_3OH$)(*l*) is also soluble in water. What is the best symbolic representation of an aqueous solution of methanol?
   a. $CH_4$(*aq*) and $H_2O$(*l*)
   b. $CH_3^+$(*aq*) and $OH^-$(*aq*)
   c. $CH_3$(*aq*) and OH(*aq*)
   d. $CH_3OH$(*aq*)

**Hint:** Copy of hint from #4

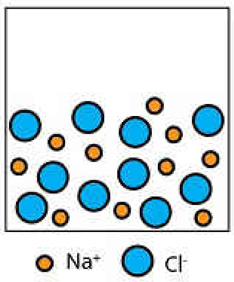
**Feeling confident you understand how to represent solutions symbolically, you decide to test yourself on representing solutions on the particle level.**

6. Which particle level diagram corresponds to the pure salt before it is added to the pot of water?

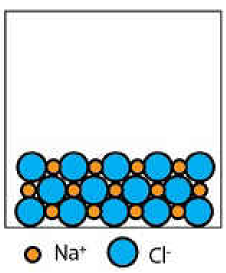
A.



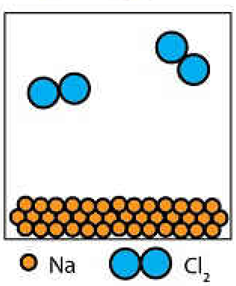

B.



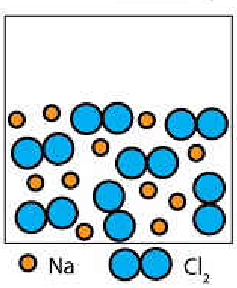

C.



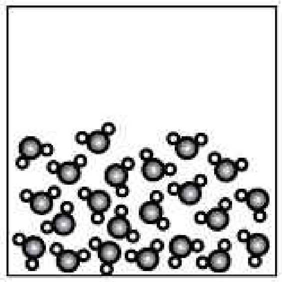

D.




E.

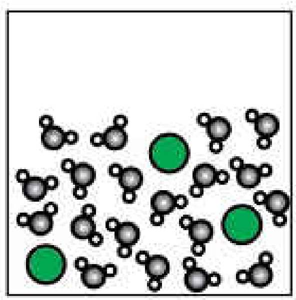


**Hint**: Particle-level diagrams show the arrangement of atoms and molecules within a substance and how each particle interacts with one another. Particle level diagrams include very specific information that demonstrates why substances behave the way that they do on the macroscopic level. This connection between structure and function is key in understanding not just what different substances do but why they do it. An example particle level diagram for a sample of $H_2O(l)$ is shown below.



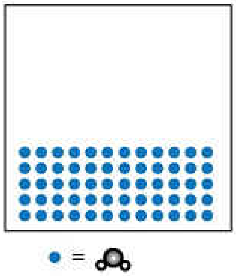
$H_2O$ (l)

If sugar was dissolved in the above sample of water to make a solution, the particle level diagram reflects this change by showing water molecules hydrating each sugar molecule. The arrangement of water molecules in an aqueous solution is specific to the type of intermolecular forces present in the solute but water molecules may or may not be included in some particle level diagrams of solutions depending on what aspect of the solution is being emphasized.

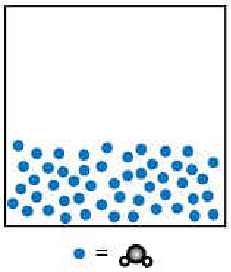


$C_6H_{12}O_6$ (aq)

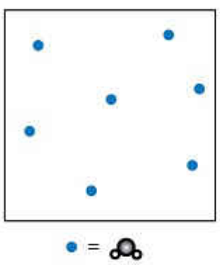
7. Which diagram corresponds to liquid water?



A.

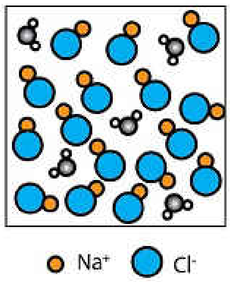


B.

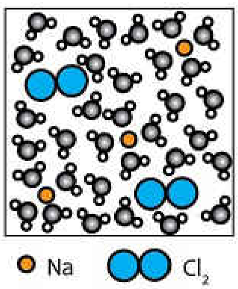


C.

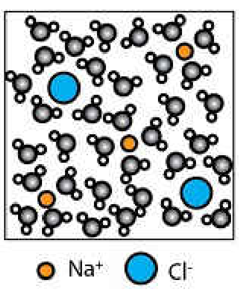
8. Which diagram corresponds to your salt solution?



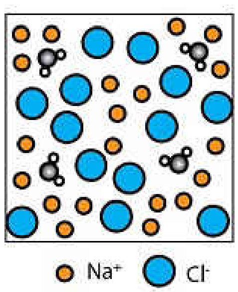

A.




B.




C.




D.

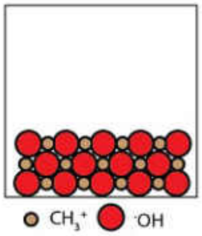
9. What happened to the distance between the sodium ions and the chloride ions from the solid to the solution?

A. Increase

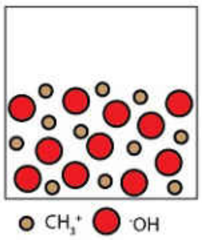
B. Decrease

10. Which diagram corresponds to pure methanol ($CH_3OH$)? (The normal boiling point of methanol is 64.70°C.)

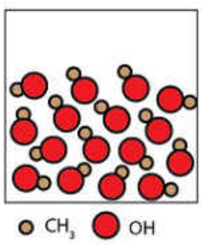
A.



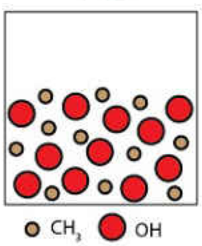

B.



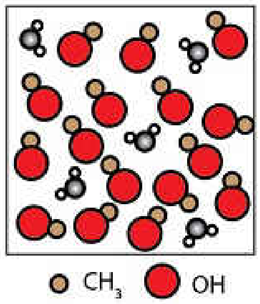

C.



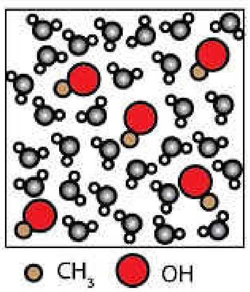

D.



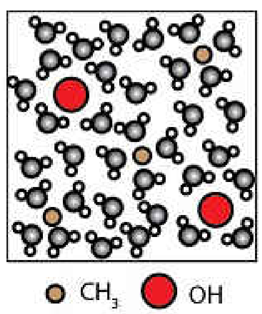
11. Which diagram corresponds to methanol in solution?



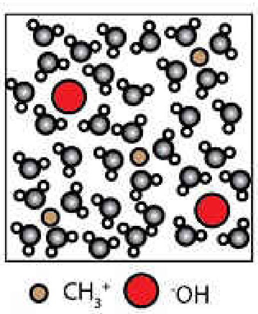

A.



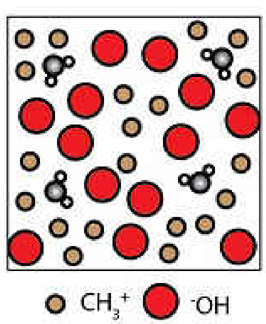

B.




C.




D.




E.

12. What happened to the distance between the methanol molecules from the pure liquid to the solution?

A. Increase
B. Decrease

**Based on your particle level drawings you can see that a solution is much different than a pure substance and start to think about how those differences affect the properties of a solution.**

13. In addition to dispersion forces, what are the intermolecular forces present in your salt solution?

**Hint:**

To assign the intermolecular forces between two substances:

1. Determine the structure of each substance
2. Determine the intermolecular forces of each substance

To determine the IMF that exist **between** the substances →

Compare the IMF:

a. **Dispersion forces** (or induced dipole-induced dipole forces) exist between *all substances* in a solution.
b. A *nonpolar* and a *polar* substance would have **induced dipole-dipole forces**
c. A *polar* and a *polar* substance would have **dipole-dipole forces**
d. A *ionic* and a *nonpolar* substance would have **ion-induced dipole forces**
e. A *ionic* and a *polar* substance would have **ion-dipole forces**

14. Qualitatively explain the forces present in a salt solution.
**Hint:** In your explanation, be sure to include how the charges present on the ions interact with the dipoles present in water.

15. In addition to dispersion forces, what are the intermolecular forces present in your methanol solution?
**Hint:** same as question 13

16. Will methanol hydrogen bond with water?

A. Yes
B. No

**Hint:** Recall that hydrogen bonding is a special type of dipole-dipole interaction between the hydrogen atom in a polar bond, such as H-N, O-H, or F-H, and an electronegative O, N, or F atom. An important extension of this definition would add that hydrogen bonding occurs in both pure substances and solutions. Take for example a solution in which both the solute ($NH_3$) and the solvent ($H_2O$) are capable of hydrogen bonding. When the two are mixed, the hydrogen bonding that occurs between different $NH_3$ molecules would be classified as "solute-solute interactions", the hydrogen bonding that occurs between different $H_2O$ molecules would be classified as "solvent-solvent interactions", and the hydrogen bonding that occurs between $NH_3$ molecules and $H_2O$ molecules would be classified as "solvent-solute interactions".

17. Qualitatively explain the forces present in a methanol solution.
**Hint:** In your explanation, be sure to include how the dipoles present in methanol interact with the dipoles present in water.

18. Based on the explanations you gave in numbers 14 and 17, do you think adding salt will make any difference in the time it takes to cook your rice?

Scoring (1 point for each answered correctly):

≥60% (11-18 questions) – Scenario 2 Questions

<60% (0-10 questions) – repeat with a note to check the hints provided

---

<u>**Scenario 2 Questions**</u>:

**Table C.3 Activity 1 (solutions) scenario 2 questions**

| Section | Number of Questions in Pool | Number of Questions Pulled |
|---|---|---|
| Heating and Cooling Curves | 8 | 2 |
| IMF | 7 | 3 |
| Total | 15 questions | 5 |

Scoring (1 point for each answered correctly):

≥50-100% (3-5 questions) – Scenario 3

<50% (0-2 questions) – repeat

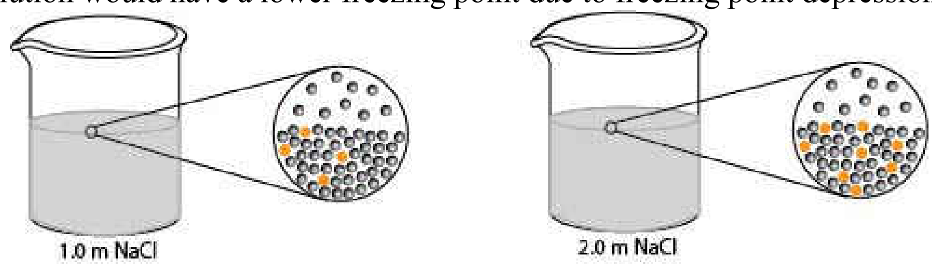
---

**Scenario 3**:

**Introduction: You are making rice using a boiling salt water solution. You relate this back to the chapter on freezing point depression that you just finished reading for your chemistry class. In lecture you learned that the freezing point of a solution is lower than the freezing point of the pure solvent used to make the solution. You remember that this is called freezing point depression and that it belongs to a group of phenomenon that are independent of the identity of the solute but are dependent on the quantity of the solute in solution. You know that boiling point elevation and vapor pressure lowering also belong to this group and you start thinking about how you might be observing the effects of these properties as you cook.**

**Please use the hints provided as they are designed to help you with answering the questions. Any time you see a definition you will find the definition and any other relevant information in the hint. Good luck!**

1. Which solution would have a lower freezing point due to freezing point depression?




a. 1.0 *m* NaCl

b. 2.0 *m* NaCl

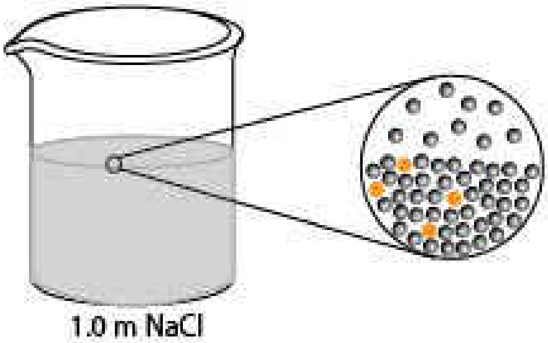
**Hint**: Freezing-point depression is the observed lowering of the freezing point of a pure substance when it is combined with a solute to make a solution. Freezing-point depression belongs to a special class of properties called colligative properties that only depend on the quantity of solute in solution, not on the identity of the solute.

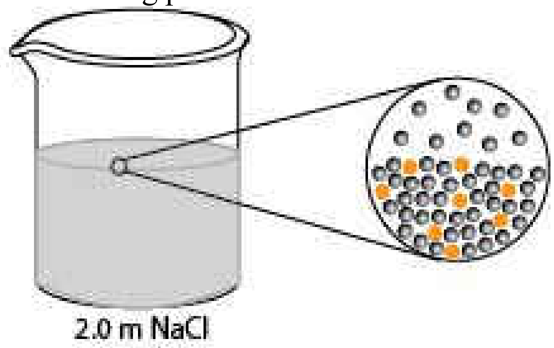
The equation for freezing-point depression is:

$$\Delta T_f = K_f m$$

where $\Delta T_f$ is the change in freezing temperature of the solution from the pure solvent, $K_f$ is the molal freezing-point depression constant (units of °C/$m$), and $m$ is the molality of the solution.

2. Which solution would have a higher boiling point due to boiling point elevation?




a. 1.0 *m* NaCl

b. 2.0 *m* NaCl

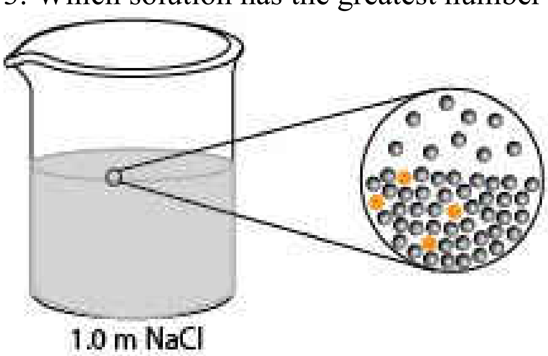
**Hint**: Boiling point elevation is the observed increase of the boiling point of a pure substance when it is combined with a solute to make a solution. Boiling point elevation belongs to a special class of properties called colligative properties that only depend on the quantity of solute in solution, not on the identity of the solute.

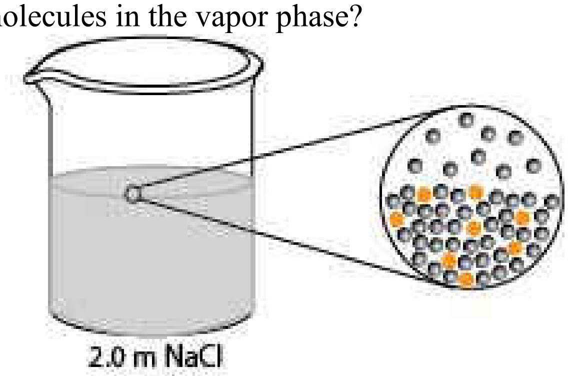
The equation for boiling point elevation is:

$$\Delta T_b = K_b m$$

where $\Delta T_b$ is the change in boiling temperature of the solution from the pure solvent, $K_b$ is the molal boiling point elevation constant (units of °C/*m*), and *m* is the molality of the solution.

3. Which solution has the greatest number of water molecules in the vapor phase?



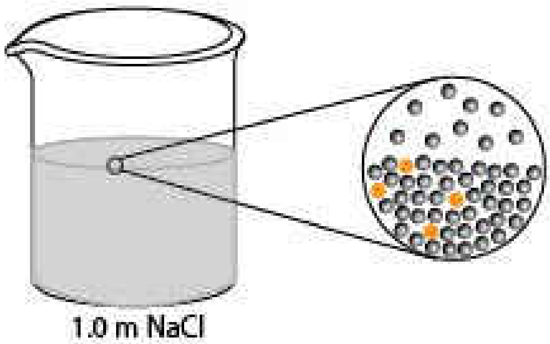

a. 1.0 *m* NaCl

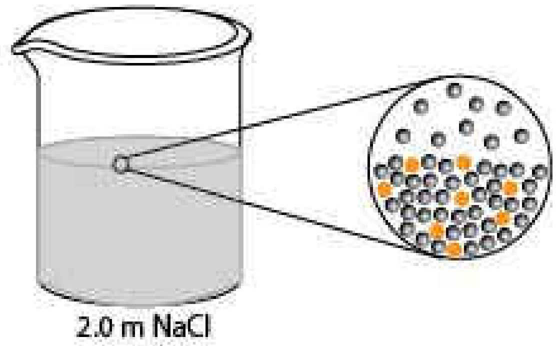
b. 2.0 *m* NaCl

c. Both solutions have an equal number of water molecules in the vapor phase.

4. Explain why your answer to number 3 is correct.

5. Which solution has the highest vapor pressure?



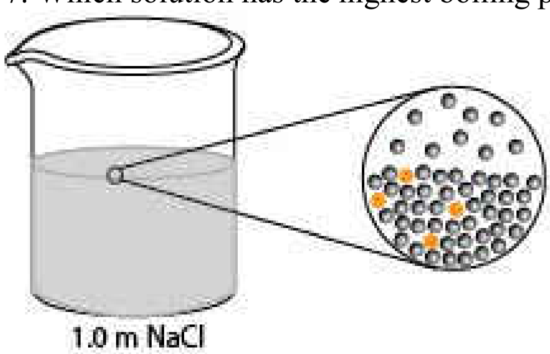

a. 1.0 *m* NaCl

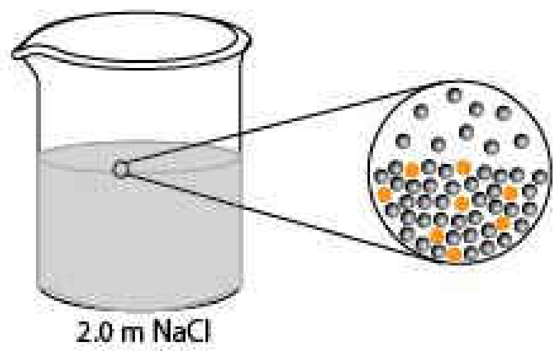
b. 2.0 *m* NaCl

c. Both solutions have the same vapor pressure.

6. Explain why your answer to number 5 is correct.

7. Which solution has the highest boiling point?




a. 1.0 *m* NaCl

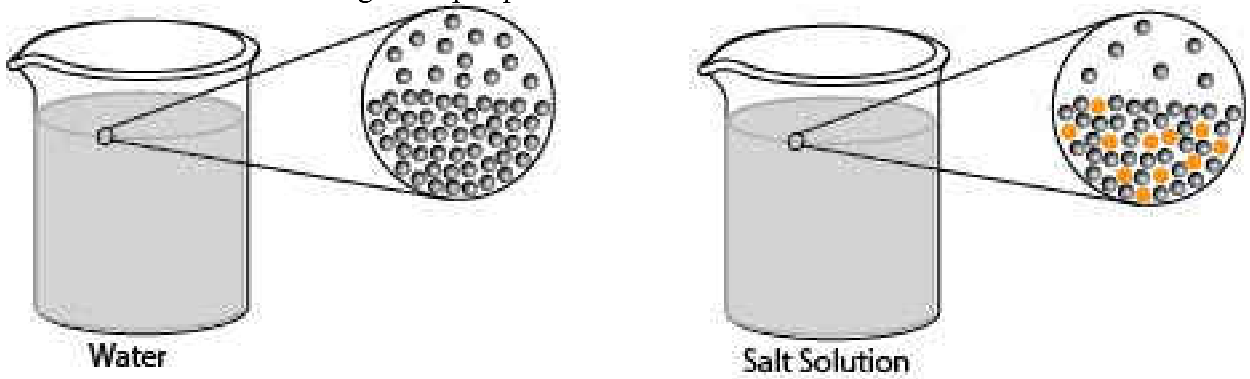
b. 2.0 *m* NaCl

c. Both solutions have the same boiling point.

8. Explain why your answer to number 7 is correct.

**Recall that vapor pressure is due to the particles in the vapor state above the liquid.**

9. Which substance has a higher vapor pressure?




a. The salt solution has a higher vapor pressure.

b. Water has a higher vapor pressure.

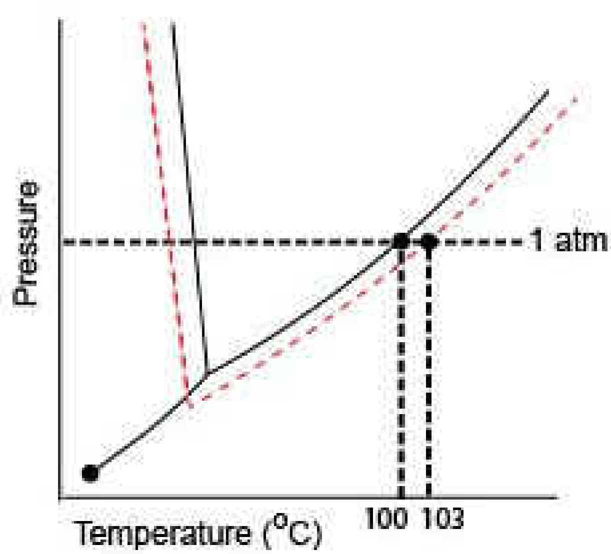
c. Water and a salt solution have equal vapor pressures.

10. Explain why your answer to number 9 is correct.

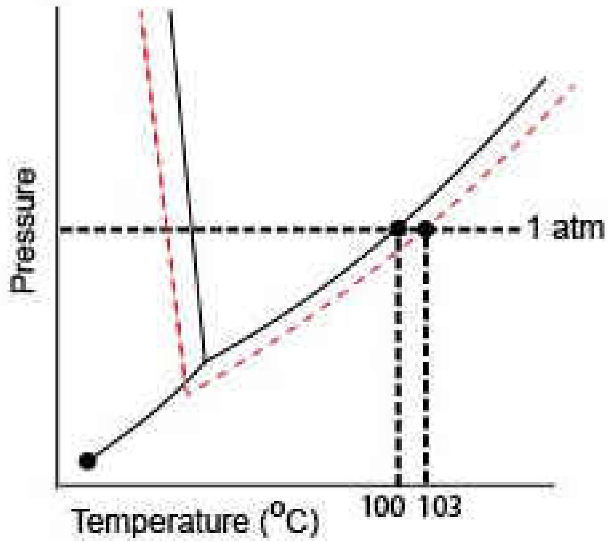
**Vapor pressure lowering and boiling point elevation are two examples of colligative properties. Because both are related to how much solute is present in a solution recall how vapor pressure relates to boiling point.**

11. What happens when the vapor pressure equals the external pressure? __________

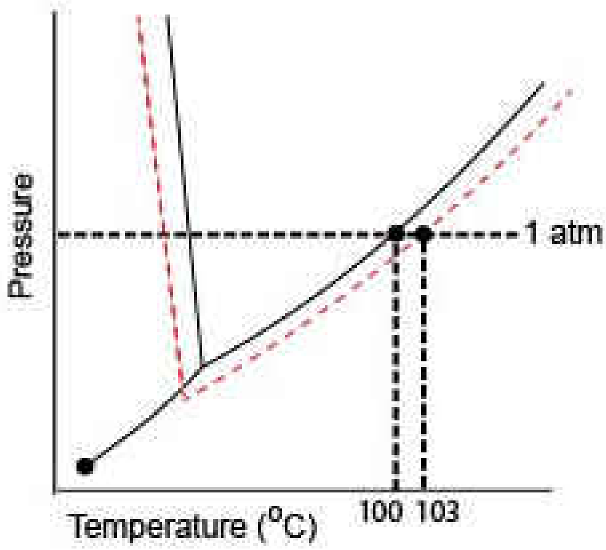
12. Using the phase diagram, what is the normal boiling point for the solvent (in °C)?
_____________ °C



13. Using the phase diagram, what is the boiling point for the solution (in °C)? ___________ °C




14. Using the phase diagram, what is the change in temperature ($\Delta T_b$) for the solution (in °C)? ____________ °C




15. Did the addition of salt to the pot affect the cooking time of the rice?
   a. Yes
   b. No

16. Based on your answers to this activity, why do you think salt is added to water when cooking rice?

Scoring (1 point for each answered correctly):

≥60% (10-16 questions) – Scenario 3 Questions

<60% (0-9 questions) – repeat with a note to check the hints provided

---

**Scenario 3 Questions**:

**Table C.4 Activity 1 (solutions) scenario 3 questions**

| Section | Number of Questions in Pool | Number of Questions Pulled |
|---|---|---|
| Vapor Pressure Lowering | 5 | 2 |
| Boiling Point Elevation | 6 | 2 |
| Phase Diagrams of Solutions | 5 | 1 |
| Total | 16 questions | 5 |

Scoring (1 point for each answered correctly):

≥60-100% (4-5 questions) – Final questions

<60% (0-3 questions) – repeat

---

**Final Questions**:

**Table C.5 Activity 1 (solutions) final questions**

| Sections | Number of Questions in Pool | Number of Questions Pulled |
|---|---|---|
| Heating and Cooling Curves | 5 | 1 |
| Intermolecular Forces | 8 | 1 |
| Phase Change | 5 | 1 |
| Phase Diagram | 5 | 1 |
| Solution Amounts | 8 | 1 |
| Intermolecular Forces in Solutions | 7 | 1 |
| Vapor Pressure Lowering | 5 | 1 |
| Boiling Point Elevation | 6 | 1 |
| Phase Diagrams of Solutions | 5 | 1 |
| Total | 54 questions | 9 questions |

---

# C.2: Activity 2 (fuel cells)

**<u>Initial Questions</u>**:

**Table C.6 Activity 2 (fuel cells) initial questions**

| Sections | Number of Questions in Pool | Number of Questions Pulled |
|---|---|---|
| Galvanic Cells | 10 | 1 |
| Cell Potential | 8 | 1 |
| System/Surroundings | 8 | 1 |
| Macroscopic-Gases | 15 | 1 |
| Symbolic-Gases | 5 | 1 |
| Particulate-Gases | 5 | 1 |
| Symbolic-Reaction | 15 | 1 |
| Nernst Equation | 9 | 1 |
| Spontaneity | 15 | 1 |
| Energy Diagrams | 12 | 1 |
| Reaction Mechanisms | 4 | 1 |
| Total | 106 questions | 11 questions |

Scoring (1 point for each answered correctly):

>75% (8-11 questions) – Scenario 3

50-75% (6-7 questions) – Scenario 2

<50% (0-5 questions) – Scenario 1

---

<u>Scenario 1</u>:
**<u>Introduction</u>: You have been chosen to test drive a hydrogen fuel cell car (referred to as fuel cell vehicle or FCV). You may have heard that these cars are more efficient and better for the environment than a car that uses gasoline as the fuel (referred to as a standard vehicle or SV). As you walk to the new car you start to think about how this car is different than your car.**

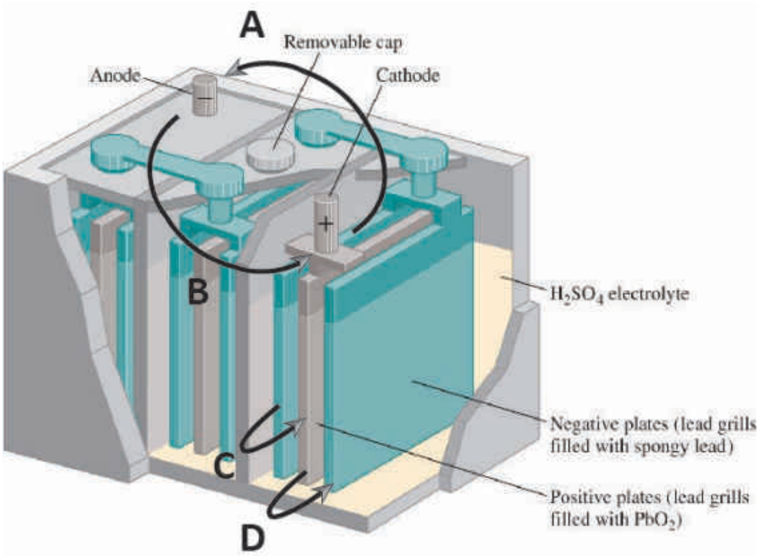
**Please use the hints provided as they are designed to help you with answering the questions. Good luck!**

**You go outside on a cold winter day to drive to school. You have recently been chosen to test drive a hydrogen fuel cell car. As you start your hydrogen fuel cell car you wonder if the temperature will affect how the car warms up compared to a summer day.**

**Today is a particularly cold day and you know that sometimes a standard vehicle, after sitting overnight in cold temperatures, may not start. While this occurrence is the result of several factors, the greatest concern is that the battery has failed. You know that the fuel**

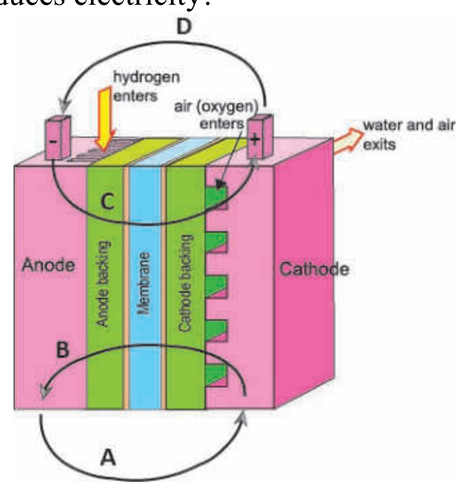
**cell in a FCV is analogous to the battery in a SV in that it produces energy, but you start to wonder what makes the FCV different and if the FCV will act similarly in cold temperatures.**

1. First, you think about the battery in your SV. Which part of this image shows where a SV battery produces electricity?




   a. A
   b. B
   c. C
   d. D

2. Thinking about how the SV battery is different from a FCV, which part of this image shows how a FCV produces electricity?




   a. A
   b. B
   c. C
   d. D

**The FCV uses fuel cells to produce electricity. In theory, a single hydrogen fuel cell can produce 1.23 V of electricity, but, in reality, the output is closer to only 0.7 V of electricity.**

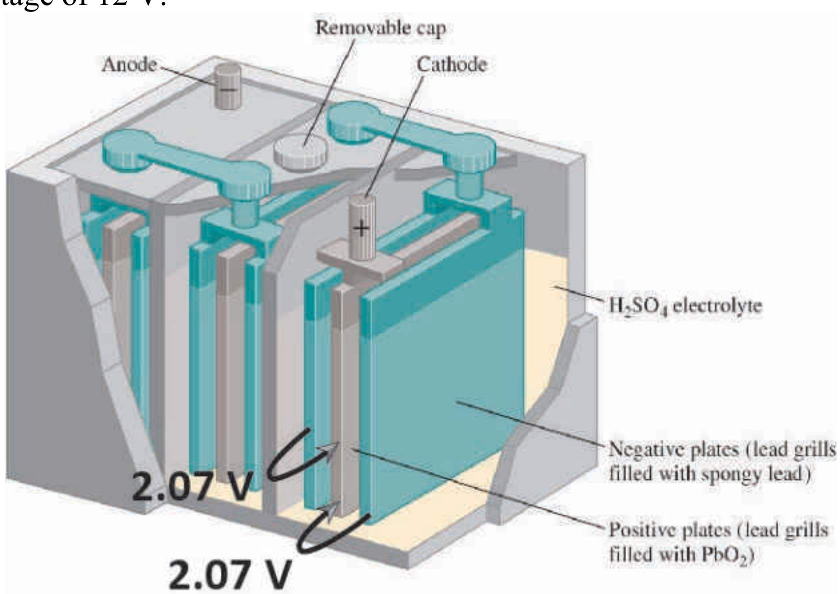
3. Which object(s) could be powered by 0.7 V of electricity? Select all that apply.
    a. A small flashlight
    b. A laptop
    c. A cellphone
    d. A house

4. Do you think 0.7 V is enough to power a car? Yes/No

5. Explain your answer.

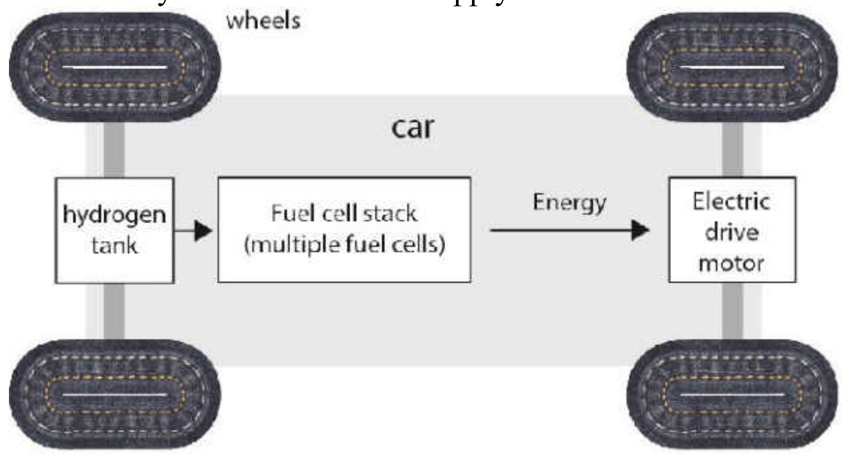
6. How could you increase the total voltage produced in order to be able to run a FCV?
    a. Increase the amount of platinum catalyst in a fuel cell
    b. Increase the surface area of the fuel cell
    c. Increase the number of fuel cells
    d. Increase the size of a fuel cell

**Hint**: Lead storage batteries are commonly used in cars. Most vehicles contain six identical cells connected together. Each cell has a voltage of 2 V so connecting six identical cells together gives a total voltage of 12 V.




7. Hydrogen fuel cells can be tiny so a FCV would be able to hold many cells. If the average size of the fuel cell in the FCV is 200 µM how many fuel cells do you need to have an output voltage of 200 V?

8. The amount of electricity produced by the fuel cells is dependent on both the temperature and the pressure of the system. Below is a schematic of the FCV. Identify the components of the system. Select all that apply.

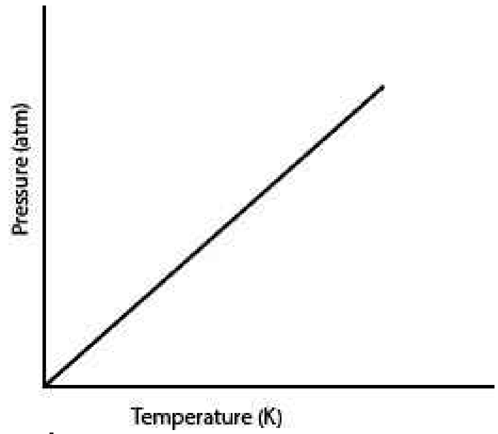



   a. Hydrogen tank
   b. Fuel cell stack
   c. Electric drive motor
   d. Wheels
   e. Car

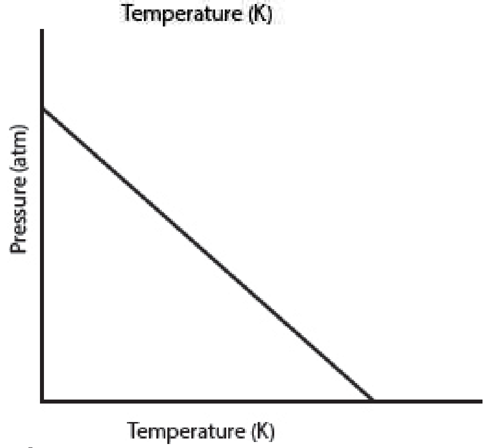
9. Let's focus on the system components. What chemical reaction is occurring in the fuel cell to produce electricity? The FCV has a maximum temperature rating of 125ºC (257ºF).
   a. Hydrogen → water vapor
   b. Hydrogen → liquid water
   c. Hydrogen + oxygen → water vapor
   d. Hydrogen + oxygen → liquid water

10. As you are driving to school you notice that the fuel gauge on the FCV is showing the FCV is low on fuel. What does this mean?
    a. You are running low on hydrogen gas.
    b. You are running low on oxygen gas.
    c. You are running low on water vapor.

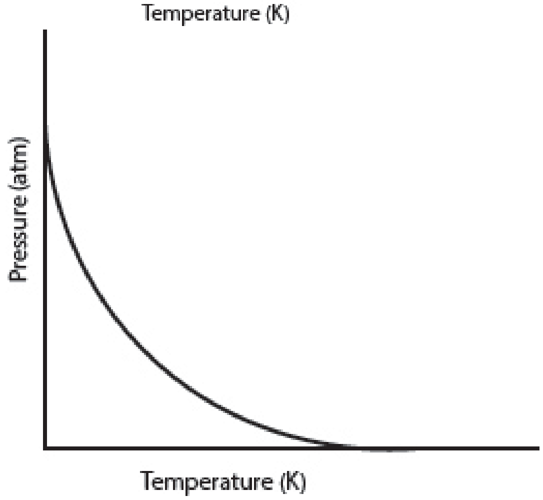
11. The fuel in the FCV is stored as a gas. How is the temperature of the system related to the pressure of the gas in the tank if we assume the gas is ideal? Select the correct graph that shows this relationship.


Pressure (atm)
Temperature (K)

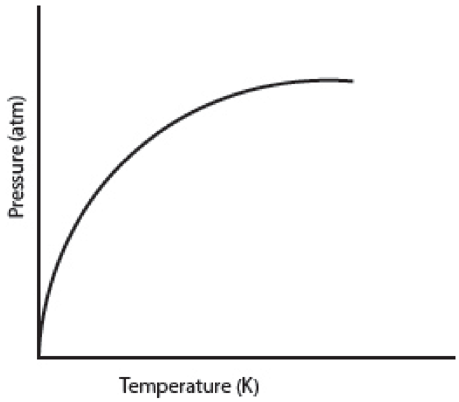
a.


Pressure (atm)
Temperature (K)

b.


Pressure (atm)
Temperature (K)

c.

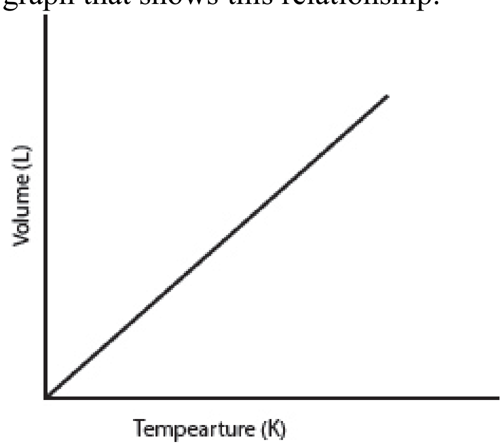



d.

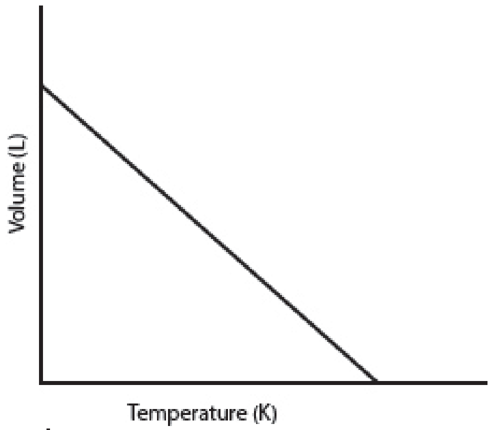
12. Today is a particularly cold day outside. What would happen to the pressure inside the tank if the tank was rigid and the temperature decreased?
    a. Increase
    b. Decrease
    c. Pressure is not influenced by changes in temperature

13. Explain your answer.

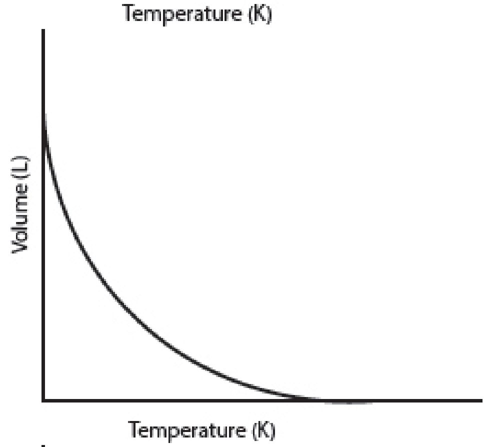
14. Fuel tanks on a FCV are flexible and adjust the volume to keep the pressure constant. What is the relationship between temperature and volume for an ideal gas? Select the correct graph that shows this relationship.



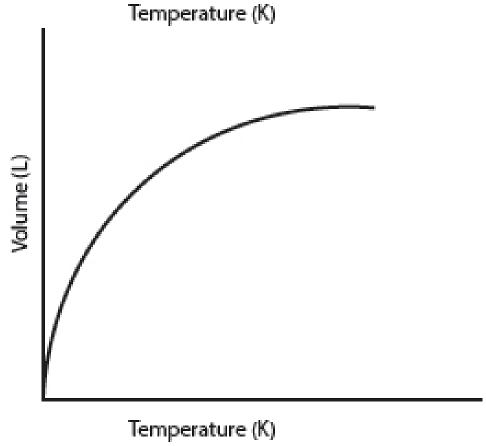

a.




b.




c.




d.

15. Today is a particularly cold day. What happens to the volume of the flexible tank if the temperature decreased in order to keep the pressure inside the tank constant?
    a. Contracts
    b. Expands
    c. Volume is not influenced by changes in temperature

16. Explain your answer.

Scoring (1 point for each answered correctly):

≥50% (8-15 questions) – Scenario 1 Questions

<50% (0-8 questions) – repeat with a note to check the hints provided

---

**Scenario 1 Questions**:

**Table C.7 Activity 2 (fuel cells) scenario 1 questions**

| Section | Number of Questions in Pool | Number of Questions Pulled |
|---|---|---|
| Galvanic Cells | 10 | 1 |
| Cell Potential | 8 | 1 |
| System/Surroundings | 8 | 1 |
| Macroscopic-Gases | 15 | 2 |
| Total | 41 questions | 5 questions |

Scoring (1 point for each answered correctly):

100% (5 questions) – Scenario 3

≥50-99% (3-4 questions) – Scenario 2

<50% (0-2 questions) – repeat

---

**Scenario 2**:

**Introduction**: **Your focus in this scenario will be on symbolic representations which will involve some calculations.**

**Please use the hints provided as they are designed to help you with answering the questions. Good luck!**

**You have been chosen to test drive a hydrogen fuel cell car (referred to as fuel cell vehicle or FCV). You may have heard that these cars are more efficient and better for the environment than a car that runs on gasoline as the fuel (referred to as a standard vehicle or SV). The hydrogen used in your car is stored in a flexible tank that keeps the pressure at 10,000 psi. The reactant gases undergo catalytic reactions that produce energy that powers your car. The energy output is less than 80% efficient and results in a fuel economy of roughly 70 mpk (miles per kilogram of hydrogen). The car is rated for 300 miles per tank of gas with a maximum temperature rating of 125ºC (257ºF). Based on your experience in your chemistry class you are going to figure out how big the tank is and whether hydrogen can be considered an ideal gas.**

**The hydrogen gas used to fuel the FCV is expensive. Thinking about the efficiency of the fuel cell, you contemplate the reaction between hydrogen and oxygen that allows your car to run.**

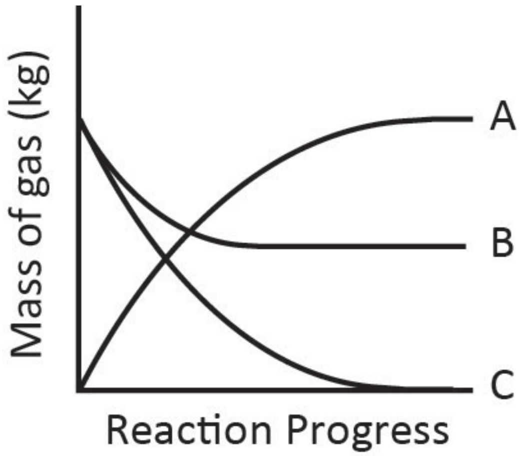
1. Ignoring the catalyst, what is the symbolic representation (balanced equation) for the reaction that occurs between hydrogen and oxygen in the fuel cell?
   a. $H(g) + O(g) \rightarrow H_2O(g)$
   b. $2H_2(g) + O_2(g) \rightarrow 2H_2O\ (g)$
   c. $H_2(g) + O_2(g) \rightarrow H_2O\ (g)$
   d. $2H(g) + O(g) \rightarrow 2H_2O\ (g)$
   e. $H(g) + O(g) \rightarrow H_2O\ (l)$
   f. $2H_2(g) + O_2(g) \rightarrow 2H_2O\ (l)$
   g. $H_2(g) + O_2(g) \rightarrow H_2O\ (l)$
   h. $2H(g) + O(g) \rightarrow 2H_2O\ (l)$

2. How should the reaction in number 1 be classified?
   a. Double displacement reaction
   b. Combustion reaction
   c. Decomposition reaction
   d. Oxidation-reduction reaction

3. The type of reaction you identified in question 2 has other formats that the reaction could be written. Fill in the various elements and coefficients of the reduction reaction taking place. You must enter a numerical value for a coefficient (including if the coefficient is 1, but remember this can also be 0).

4. Fill in the various elements and coefficients of the oxidation reaction taking place. You must enter a numerical value for a coefficient (including if the coefficient is 1, but remember this can also be 0).

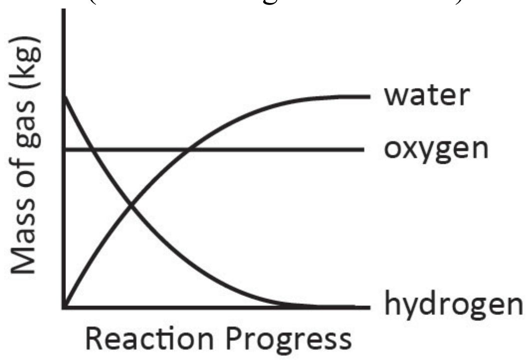
5. The reaction in the fuel cell could also be represented in a plot showing the depletion of reactants and production of products. Using the reaction that you identified in number 1, between hydrogen and oxygen to form water vapor, and the given plot, which letter corresponds to each substance of the reaction?




A. Water vapor
B. Oxygen
C. Hydrogen

6. Looking at the image also used in the previous question, which letter corresponds to the limiting reactant?
   a. A
   b. B
   c. C

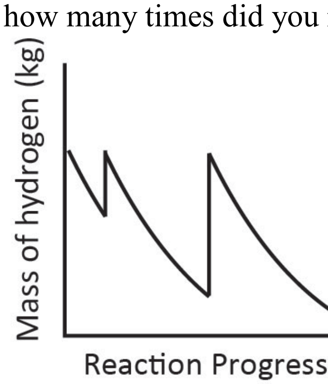
7. Let's look back at a different plot that also shows the depletion of the reactants and production of the products. Which statement best describes the difference between this plot and the one given in number 5 (also shown again in the hint)?




   a. Oxygen is the limiting reactant
   b. Oxygen is being constantly supplied
   c. Oxygen is not present
   d. Oxygen is now the product

**Hint**: Plot given in question number 5. A, B, and C are referring to substances from the reaction of hydrogen and oxygen to form water vapor. [Image: 2_5.jpg]


8. The reaction in the fuel cell could also be represented in a plot showing the depletion of hydrogen. When you refuel a car the amount of hydrogen is replenished in the storage tank. Assuming oxygen is in excess, how many times did you refuel the FCV?



**In chemistry class, you've been learning about Galvanic cells and remember that the definition of a galvanic cell is "an electrochemical cell that generates electricity by means of a spontaneous redox reaction" (p. 669).**

9. Is the fuel cell in your car a galvanic cell?
   a. Yes
   b. No

10. Explain your answer.

11. Galvanic cells use spontaneous oxidation-reduction reactions to produce electrical energy. The amount of energy produced by the cell that "is available to do work" is called Gibbs free energy (p. 644). Given a constant energy output and using the sign convention for Gibbs free energy that you are familiar with what's the relationship between Gibbs free energy and efficiency?
    a. The more positive the Gibbs free energy the more efficient the reaction
    b. The more negative the Gibbs free energy the more efficient the reaction
    c. Gibbs free energy is not related to the reaction efficiency

12. Each car comes with an efficiency rating. An efficiency of 100% means that 100% of the change in Gibbs free energy is available to use. You know that the FCV runs at about 80% efficiency meaning that 80% of the hydrogen fuel can successfully be converted to usable energy. What is one reason the efficiency is not 100%?

13. Calculate the change in Gibbs free energy for one mole of the system of hydrogen and oxygen combining to form water vapor at room temperature (25°C) where the change in enthalpy is -241.8 kJ and the change in entropy is -147.3 J/K.
    a. 43680 kJ
    b. 3441 kJ
    c. -197.9 kJ
    d. -238.1 kJ

**Hint**: The equation for the change in Gibbs free energy (ΔG) is:

$$\boldsymbol{\Delta G = \Delta H - T\Delta S}$$

where H is enthalpy, T is temperature, and S is entropy.

For example: $N_2(g) + 3H_2(g) \rightarrow 2NH_3(g)$

has a ΔH of -92.22 kJ and a ΔS of -198.75 J/K

at 298.15 K what is ΔG?

$$\boldsymbol{\Delta G = -92.22\ kJ - (298.15\ K)\left(-198.75\ \frac{J}{K}\right)}$$

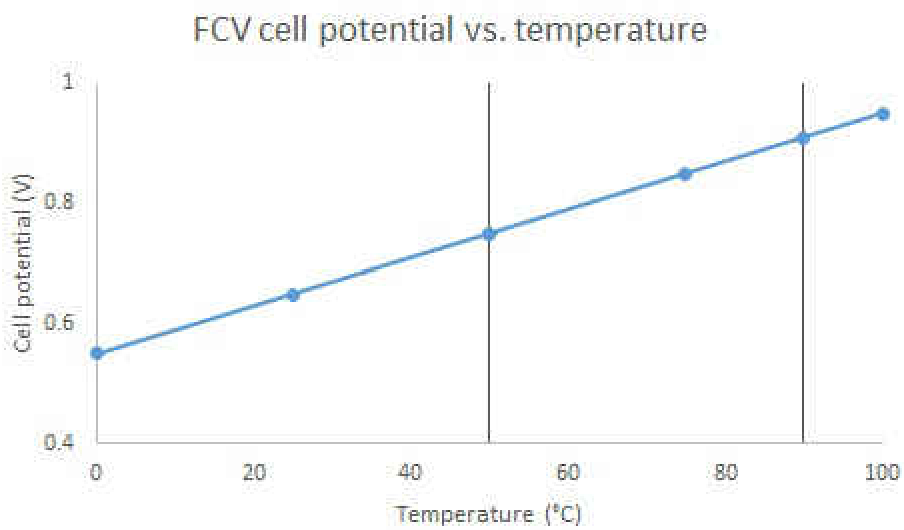
$$\Delta G = -92.22\ kJ - (-59257.315\ J) * \left(\frac{1\ kJ}{1X10^3 J}\right)$$

$$\Delta G = -92.22\ kJ - (-59.26\ kJ) = -32.96\ kJ$$

A negative ΔG means the reaction is spontaneous as written. A positive ΔG means the reaction is spontaneous in the reverse direction.

14. Is this reaction spontaneous based on the number you calculated in the previous problem?
    a. Yes
    b. No

15. As you know from class, Gibbs free energy is related to cell potential. Use the plot to describe what's happening in a normal FCV.




    a. At the normal FCV operating temperature, the cell potential is lower and the efficiency is lower.
    b. At the normal FCV operating temperature, the cell potential is lower and the efficiency is higher.
    c. At the normal FCV operating temperature, the cell potential is higher and the efficiency is lower.
    d. At the normal FCV operating temperature, the cell potential is higher and the efficiency is higher.

**Hydrogen gas is the fuel used in the FCV and oxygen gas is constantly supplied from the atmosphere. Gases behave ideally at sufficiently low pressure and high temperature.**

16. Using the information above, assuming the tank in your car has a maximum temperature rating of 125°C, how big is the tank in your car if you have 3.0 kg of hydrogen at 10,000 psi (680 atm) in the tank?
    a. 0.14 L
    b. 9.8 L
    c. 45 L
    d. 72 L

**Hint**: Calculate the volume of oxygen gas if the amount of oxygen weighs 2.0 kg under 8,000 psi of pressure, at 100°C. In order to calculate the volume of an ideal gas the combined gas law equation can be used.

$$\boldsymbol{PV = nRT}$$

This can be rearranged to read:

$$\boldsymbol{V = \frac{nRT}{P}}$$

The gas law constant, R, is 0.082057 with units L•atm/mol•K. Convert the information in the question so that the units match the gas law constant.

$$\boldsymbol{2.0\ kg\ O_2 * \frac{1000\ g}{1\ kg} * \frac{1\ mol\ O_2}{32.00\ g\ O_2} = 62.\ mol\ O_2}$$

$$\boldsymbol{100^\circ\mathrm{C} + 273.15 = 373.15\ K}$$

$$\boldsymbol{8{,}000\ psi * \frac{0.0680\ atm}{1\ psi} = 544\ atm}$$

$$\boldsymbol{V = \frac{(62.\ mol\ O_2)(0.082057\ \frac{L * atm}{mol * K})(373.15\ K)}{544\ atm} = 3.4\ L\ O_2}$$

17. Assuming just a volume of 1.00 L, how many hydrogen molecules are in this tank at STP?
    a. $2.7 \text{ X } 10^{22}$ hydrogen molecules
    b. $5.9 \text{ X } 10^{22}$ hydrogen molecules
    c. $6.2 \text{ X } 10^{24}$ hydrogen molecules
    d. $2.0 \text{ X } 10^{25}$ hydrogen molecules

18. Assuming just a volume of 1.0 L, how many hydrogen molecules are in this tank at 10,000 psi (680 atm) at 125°C?
    a. $1.2 \text{ X } 10^{25}$ hydrogen molecules
    b. $4.0 \text{ X } 10^{25}$ hydrogen molecules
    c. $1.8 \text{ X } 10^{26}$ hydrogen molecules
    d. $5.9 \text{ X } 10^{26}$ hydrogen molecules

19. At STP, hydrogen molecules are approximately 3800 pm apart and at 680 atm they compress to approximately 440 pm apart. How many times closer together are the molecules at high pressure than at low pressure? Round to the nearest whole number.

20. If hydrogen is stored at 10,000 psi in your vehicle, is it realistic to consider hydrogen as an ideal gas at this pressure?
    a. Yes
    b. No
21. Explain your previous answer.

Scoring (1 point for each answered correctly):

≥50% (11-21 questions) – Scenario 2 Questions

<50% (0-10 questions) – repeat with a note to check the hints provided

---

**<u>Scenario 2 Questions</u>**:

**Table C.8 Activity 2 (fuel cells) scenario 2 questions**

| Section | Number of Questions in Pool | Number of Questions Pulled |
|---|---|---|
| Symbolic-Gases | 5 | 1 |
| Symbolic-Reaction | 15 | 2 |
| Nernst Equation | 9 | 1 |
| Spontaneity and Temperature | 15 | 1 |
| Total | 44 questions | 5 questions |

Scoring (1 point for each answered correctly):

≥50-100% (3-5 questions) – Scenario 3

<50% (0-2 questions) – repeat

---

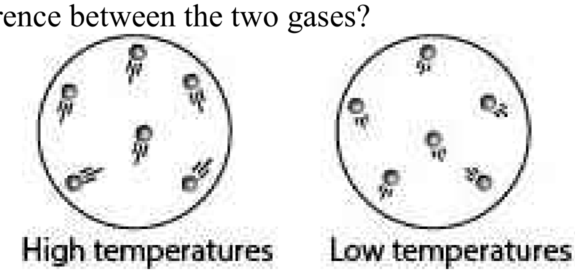
**<u>Scenario 3</u>**:
**<u>Introduction</u>: Your focus in this scenario will be on the particulate level.**

**Please use the hints provided as they are designed to help you with answering the questions. Good luck!**

**You have been chosen to test drive a hydrogen fuel cell car (referred to as fuel cell vehicle or FCV). You may have heard that these cars are more efficient and better for the environment than a car that runs on gasoline as the fuel (referred to as a standard vehicle or SV). Today you are car-pooling to chemistry class with a friend and discussing your upcoming chemistry exam on energy. Your friend says that because your car is using energy to drive, the reactions occurring inside the fuel cell must all be exothermic. He says lots of chemical energy gets released when bonds are broken due to the energy stored in the bonds the car then converts the chemical energy into electrical energy. You tell your friend that you remember hearing your chemistry professor say that even though a reaction may be exothermic overall, energy is still required to break the bonds of the reactants before the atoms can rearrange and form new bonds. You aren't sure who is right, but start to discuss the enthalpy and entropy involved in the reactions occurring in your FCV.**

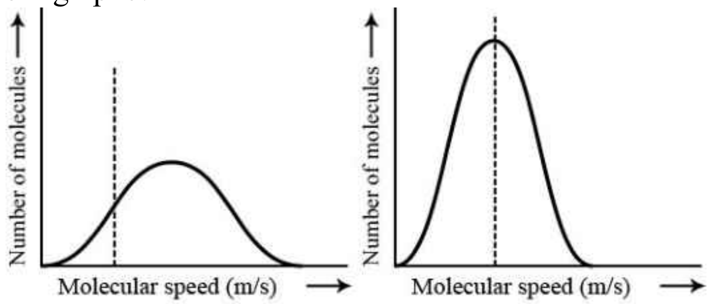
**A reaction is the result of molecular collisions. A reaction cannot occur without sufficient kinetic energy and proper orientation of the molecules. As the temperature increases, the gas particles gain more energy which causes a greater number of collisions. If we compare the reaction inside a fuel cell to a much simpler process, the combustion reaction of hydrogen, $H_2 + \frac{1}{2} O_2 \rightarrow H_2O$ then $\Delta G° = -228.6$ kJ/mol.**

1. What is the difference between the two gases?




   a. At higher temperatures, the particles, on average, are moving fast.
   b. At lower temperatures, the particles, on average, are moving fast.
   c. At higher temperatures, every particle is moving fast.
   d. At lower temperatures, every particle is moving fast.

2. Another representation we can use to discuss the speed of molecules is a Maxwell speed distribution curve. A speed distribution graph shows the number of molecules that are moving at a particular speed. What is the difference between the two graphs if hydrogen is the gas in both graphs?

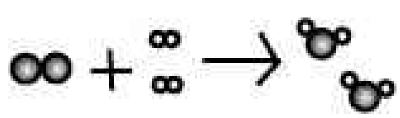



Graph A Graph B

   a. Graph A has a greater fraction of gas particles moving at or above the marked speed.
   b. Graph B has a greater fraction of gas particles moving at or above the marked speed.
   c. Graph A and Graph B have the same fraction of gas particles moving at or above the marked speed.

3. Which particles must collide for the combustion reaction of hydrogen to start?
   a. 1 molecule of $H_2$ and 1 molecule of $O_2$
   b. 1 molecule of $H_2$ and 1 O atom
   c. 1 $H^+$ ion and 1 $O^{2-}$ ion
   d. 2 $H^+$ ions and 1 $O^{2-}$ ion

**The combustion reaction of hydrogen, $H_2 + \frac{1}{2} O_2 \rightarrow H_2O$ $\Delta G° = -228.6$ kJ/mol, one molecule of hydrogen and one molecule of oxygen must collide with sufficient kinetic energy and proper orientation. The tank of the FCV keeps hydrogen at constant pressure meaning heat can be equated with enthalpy. In order for the FCV to run, the overall reaction occurring in the fuel cell should be exothermic.**



**Your friend, who is not driving, calculates the enthalpy of the reaction**.

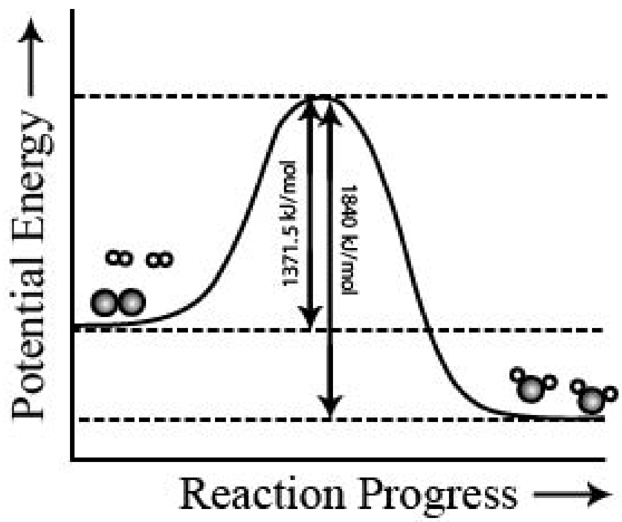
| Type of bonds broken | Number of bonds broken | Bond enthalpy (BE) (kJ/mol) | Energy change (kJ/mol) |
|---|---|---|---|
| H-H | 2 | 436.4 | 872.8 |
| O=O | 1 | 498.7 | 498.7 |
| **Type of bonds formed** | **Number of bonds formed** | **Bond enthalpy (BE) (kJ/mol)** | **Energy change (kJ/mol)** |
| O-H | 2 | 460 | 1840 |

**$\Delta H° = \Sigma BE$ (reactants) – $\Sigma BE$ (products)**

**$\Delta H°$= (872.8 kJ/mol + 498.7 kJ/mol) – 1840 kJ/mol = -469 kJ/mol**

4. Based on the calculated enthalpy, is this reaction endothermic or exothermic?
    a. Exothermic because the enthalpy is negative.
    b. Exothermic because the enthalpy is positive.
    c. Endothermic because the enthalpy is negative.
    d. Endothermic because the enthalpy is positive.

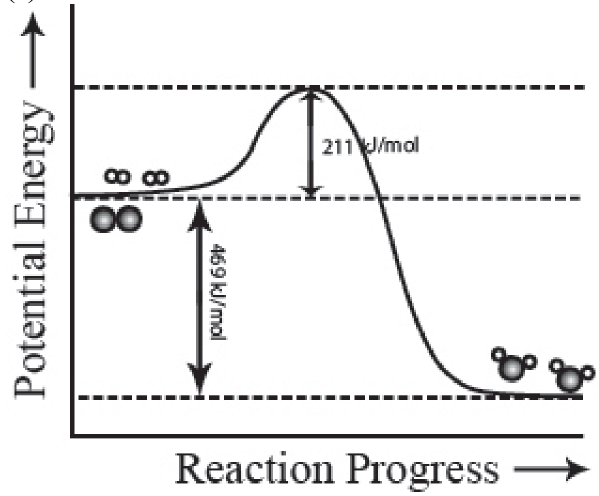
5. Another way to display the information contained in the table above is with an energy diagram. Looking at this energy diagram, how much energy is needed to reach the transition state?




    a. -469 kJ/mol
    b. 1371.5 kJ/mol
    c. 3211.5 kJ/mol
    d. Not enough information

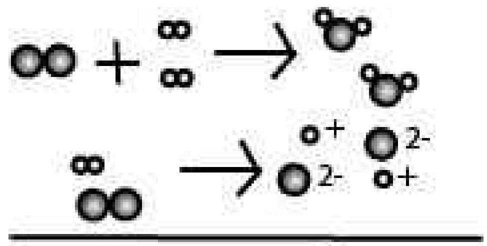
6. Where does the value for activation energy come from?

7. The Ea for this reaction is 211 kJ/mol what does that mean in terms of the energy diagram? Include in your answer an explanation of why the Ea and the ΣBE are not the same in terms of the intermediate(s) formed.

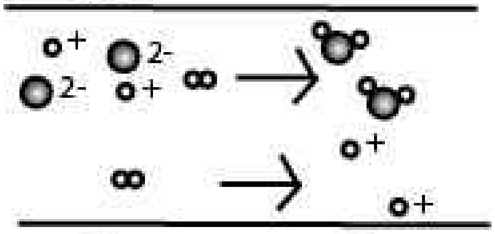



8. If the rate law for this reaction is rate = k[$H_2$], select the most plausible mechanism for this reaction.

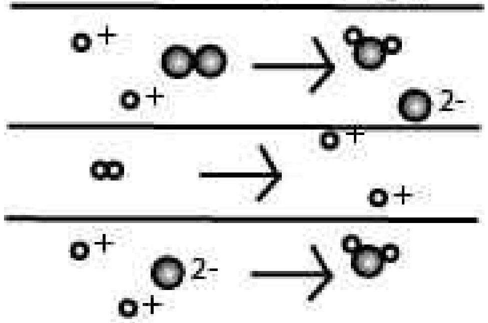
a.



b.

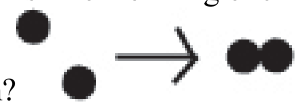


c.



9. Thinking back at your answer for number 8, explain your reasoning of which mechanism you selected.

10. Simplifying this process and just looking at the mechanism for forming one H-H bond,



what is the sign of the entropy change for this reaction?
   a. Positive because the number of microstates is reduced.
   b. Positive because the number of microstates is increased.
   c. Negative because the number of microstates is reduced.
   d. Negative because the number of microstates is increased.

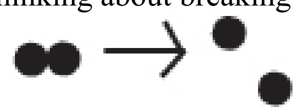
11. Entropy can help us predict the spontaneity of a reaction. Is this reaction of bond formation spontaneous?
   a. Yes because the reaction decreases the entropy of the universe.
   b. Yes because the reaction increases the entropy of the universe.
   c. No because the reaction decreases the entropy of the universe.
   d. No because the reaction increases the entropy of the universe.
   e. Cannot be determined from the information given.

12. Now that we know about the spontaneity of the reaction, lets focus on the energy changes of the system. What is the sign of the enthalpy change for this bond formation reaction?
   a. Positive because heat is released from the system.
   b. Positive because heat is absorbed from the surroundings.
   c. Negative because heat is released from the system.
   d. Negative because heat is absorbed from the surroundings.

13. Based on your answer to number 12 above, is this reaction exothermic or endothermic?
   a. Exothermic
   b. Endothermic

14. Explain you answer.

15. Reversing this process and thinking about breaking one H-H bond. Is this reaction



exothermic or endothermic?
   a. Exothermic
   b. Endothermic

16. Explain your answer.

17. Based on your previous answers, who was right? You saying not every process is exothermic even though the overall reaction is exothermic, or your friend saying every process that contributes to the overall reaction is exothermic in order for the overall reaction to be exothermic?
   a. You
   b. Your friend

18. Explain your answer.

Scoring (1 point for each answered correctly):

≥60% (11-18 questions) – Scenario 3 Questions

<60% (0-10 questions) – repeat with a note to check the hints provided

---

**Scenario 3 Questions**:

**Table C.9 Activity 2 (fuel cells) scenario 3 questions**

| Section | Number of Questions in Pool | Number of Questions Pulled |
|---|---|---|
| Particulate-Gases | 5 | 1 |
| Energy Diagrams | 12 | 1 |
| Reaction Mechanisms | 4 | 2 |
| Energy/Bonding | 15 | 1 |
| Total | 36 questions | 5 |

Scoring (1 point for each answered correctly):

≥60-100% (4-5 questions) – Final questions

<60% (0-3 questions) – repeat

---

**Final Questions**:

**Table C.10 Activity 2 (fuel cells) final questions**

| Sections | Number of Questions in Pool | Number of Questions Pulled |
|---|---|---|
| Galvanic Cells | 10 | 1 |
| Cell Potential | 8 | 1 |
| System/Surroundings | 8 | 1 |
| Macroscopic-Gases | 15 | 1 |
| Symbolic-Gases | 5 | 1 |
| Particulate-Gases | 5 | 1 |
| Symbolic-Reaction | 15 | 1 |
| Nernst Equation | 9 | 1 |
| Spontaneity | 15 | 1 |
| Energy Diagrams | 12 | 1 |
| Reaction Mechanisms | 4 | 1 |
| Total | 106 questions | 11 questions |

---

# Appendix D: Anatomy and Physiology I correlation matrix

## Table D.1 Pearson correlation matrix of Anatomy and Physiology I course measures

*n* = 111

| | | ACT COMP | ACT READ | ACT ENGL | ACT MATH | ACT SCIRE | SLST Pre | SCI Pre | SLS Pre | SLST Pre_adj | SLS Pre adj | SLST pre math | SLST pre math adj | SLST pre theo | SLST pre theo adj | Final Exam Score | Course Score |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ACT COMP | Pearson Correlation | 1 | .808** | .823** | .774** | .822** | .500** | .320** | .527** | .481** | .511** | .328** | .296** | .528** | .510** | .554** | .314** |
| ACT READ | Pearson Correlation | .808** | 1 | .628** | .402** | .563** | .409** | .249** | .428** | .391** | .413** | .297** | .269** | .416** | .397** | .460** | .244** |
| ACT ENGL | Pearson Correlation | .823** | .628** | 1 | .490** | .523** | .383** | .251** | .405** | .367** | .392** | .194* | 0.164 | .438** | .425** | .461** | .291** |
| ACT MATH | Pearson Correlation | .774** | .402** | .490** | 1 | .642** | .392** | .290** | .424** | .387** | .420** | .336** | .322** | .368** | .359** | .402** | .249** |
| ACT SCIRE | Pearson Correlation | .822** | .563** | .523** | .642** | 1 | .435** | .292** | .462** | .412** | .443** | .240* | .202* | .485** | .467** | .452** | .228* |
| SLST Pre | Pearson Correlation | .500** | .409** | .383** | .392** | .435** | 1 | .319** | .966** | .988** | .957** | .840** | .811** | .949** | .926** | .321** | 0.047 |
| SCI Pre | Pearson Correlation | .320** | .249** | .251** | .290** | .292** | .319** | 1 | .554** | .310** | .549** | .297** | .275** | .285** | .278** | 0.138 | 0.012 |
| SLS Pre | Pearson Correlation | .527** | .428** | .405** | .424** | .462** | .966** | .554** | 1 | .953** | .991** | .819** | .787** | .912** | .890** | .320** | 0.044 |
| SST Pre_adj | Pearson Correlation | .481** | .391** | .367** | .387** | .412** | .988** | .310** | .953** | 1 | .965** | .832** | .817** | .937** | .939** | .311** | 0.039 |
| SLS Pre adj | Pearson Correlation | .511** | .413** | .392** | .420** | .443** | .957** | .549** | .991** | .965** | 1 | .813** | .794** | .903** | .902** | .312** | 0.037 |
| SLST pre math | Pearson Correlation | .328** | .297** | .194* | .336** | .240* | .840** | .297** | .819** | .832** | .813** | 1 | .984** | .625** | .599** | .205* | 0.063 |
| SLST pre math adj | Pearson Correlation | .296** | .269** | 0.164 | .322** | .202* | .811** | .275** | .787** | .817** | .794** | .984** | 1 | .592** | .569** | .192* | 0.040 |
| SLST pre theo | Pearson Correlation | .528** | .416** | .438** | .368** | .485** | .949** | .285** | .912** | .937** | .903** | .625** | .592** | 1 | .983** | .343** | 0.030 |
| SLST pre theo adj | Pearson Correlation | .510** | .397** | .425** | .359** | .467** | .926** | .278** | .890** | .939** | .902** | .599** | .569** | .983** | 1 | .330** | 0.031 |
| Final Exam Score | Pearson Correlation | .554** | .460** | .461** | .402** | .452** | .321** | 0.138 | .320** | .311** | .312** | .205* | .192* | .343** | .330** | 1 | .506** |
| Course Score | Pearson Correlation | .314** | .244** | .291** | .249** | .228* | 0.047 | 0.012 | 0.044 | 0.039 | 0.037 | 0.063 | 0.040 | 0.030 | 0.031 | .506** | 1 |

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

## Table D.2 Pearson correlation matrix of Anatomy and Physiology I course measures and math placement and sub-scores

*n* = 50

| | ACT COMP | ACT READ | ACT ENGL | ACT MATH | ACT SCIRE | ALG | TRG | MBSC | SLST Pre | SCI_Pre | Scale Literacy Pre | SLST pre_adj | SLS Pre adj | SLST pre math | SLST pre math adj | SLST pre theo | SLST pre theo adj | Final Exam Score | Course Score |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ACT COMP | 1 | .862** | .845** | .797** | .858** | .703** | .636** | .623** | .602** | .364** | .627** | .579** | .608** | .446** | .410** | .608** | .578** | .714** | .351* |
| ACT READ | .862** | 1 | .740** | .525** | .620** | .434** | .400** | .400** | .519** | 0.246 | .522** | .507** | .513** | .399** | .378** | .516** | .495** | .591** | .328* |
| ACT ENGL | .845** | .740** | 1 | .470** | .581** | .560** | .422** | .454** | .489** | .291* | .508** | .469** | .492** | .322* | .285* | .519** | .496** | .678** | .354* |
| ACT MATH | .797** | .525** | .470** | 1 | .754** | .782** | .815** | .792** | .494** | .418** | .546** | .483** | .538** | .452** | .426** | .450** | .433** | .503** | 0.248 |
| ACT SCIRE | .858** | .620** | .581** | .754** | 1 | .646** | .554** | .509** | .536** | .359* | .568** | .507** | .543** | .349* | .306* | .571** | .537** | .599** | 0.241 |
| ALG | .703** | .434** | .560** | .782** | .646** | 1 | .778** | .765** | .380** | .557** | .482** | .374** | .479** | .350* | .336* | .344* | .332* | .523** | .381** |
| TRG | .636** | .400** | .422** | .815** | .554** | .778** | 1 | .733** | .366** | .442** | .440** | .374** | .448** | .333* | .311* | .334* | .347* | .323* | .282* |
| MBSC | .623** | .400** | .454** | .792** | .509** | .765** | .733** | 1 | .543** | .418** | .589** | .564** | .609** | .523** | .503** | .478** | .502** | .515** | 0.264 |
| SLST pre | .602** | .519** | .489** | .494** | .536** | .380** | .366** | .543** | 1 | .329* | .968** | .990** | .961** | .846** | .813** | .950** | .923** | .478** | 0.023 |
| SCI_Pre | .364** | 0.246 | .291* | .418** | .359* | .557** | .442** | .418** | .329* | 1 | .555** | .319* | .549** | .336* | .303* | 0.278 | 0.273 | 0.085 | 0.007 |
| SLS Pre | .627** | .522** | .508** | .546** | .568** | .482** | .440** | .589** | .968** | .555** | 1 | .957** | .993** | .834** | .797** | .910** | .886** | .443** | 0.022 |
| SLST pre_adj | .579** | .507** | .469** | .483** | .507** | .374** | .374** | .564** | .990** | .319* | .957** | 1 | .967** | .834** | .810** | .943** | .939** | .476** | 0.005 |
| SLS Pre adj | .608** | .513** | .492** | .538** | .543** | .479** | .448** | .609** | .961** | .549** | .993** | .967** | 1 | .826** | .795** | .906** | .901** | .442** | 0.006 |
| SLST pre math | .446** | .399** | .322* | .452** | .349* | .350* | .333* | .523** | .846** | .336* | .834** | .834** | .826** | 1 | .987** | .636** | .600** | 0.275 | 0.037 |
| SLST pre math adj | .410** | .378** | .285* | .426** | .306* | .336* | .311* | .503** | .813** | .303* | .797** | .810** | .795** | .987** | 1 | .597** | .558** | 0.267 | -0.005 |
| SLST pre theo | .608** | .516** | .519** | .450** | .571** | .344* | .334* | .478** | .950** | 0.278 | .910** | .943** | .906** | .636** | .597** | 1 | .983** | .529** | 0.012 |
| SLST pre theo adj | .578** | .495** | .496** | .433** | .537** | .332* | .347* | .502** | .923** | 0.273 | .886** | .939** | .901** | .600** | .558** | .983** | 1 | .516** | 0.010 |
| Final Exam Score | .714** | .591** | .678** | .503** | .599** | .523** | .323* | .515** | .478** | 0.085 | .443** | .476** | .442** | 0.275 | 0.267 | .529** | .516** | 1 | .479** |
| Course Score | .351* | .328* | .354* | 0.248 | 0.241 | .381** | .282* | 0.264 | 0.023 | 0.007 | 0.022 | 0.005 | 0.006 | 0.037 | -0.005 | 0.012 | 0.010 | .479** | 1 |

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).