# INCREASING GENERAL CHEMISTRY STUDENT PASS RATE:

# IMPROVING COURSE REQUIREMENTS AND

# IMPLEMENTING METACOGNITIVE

# TRAINING

by

Brock Lewis Casselman

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry

The University of Utah

May 2018

## The University of Utah Graduate School

# STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Brock Lewis Casselman**

has been approved by the following supervisory committee members:

| | | |
|---|---|---|
| **Charles Atwood** | , Chair | **3/27/2018**<br>Date Approved |
| **Richard Ernst** | , Member | **3/27/2018**<br>Date Approved |
| **Jordan Gerton** | , Member | **4/4/2018**<br>Date Approved |
| **Thomas Richmond** | , Member | **3/27/2018**<br>Date Approved |
| **Ryan Steele** | , Member | **4/2/2018**<br>Date Approved |

and by **Cynthia Burrows** , Chair/Dean of

the Department/College/School of **Chemistry**

and by David B. Kieda, Dean of The Graduate School.

# ABSTRACT

This research project improved general chemistry performance through four major efforts: 1) identifying factors predictive of students failing general chemistry, 2) adjusting course structure, 3) changing course curriculum to align with chemical education research, and 4) implementing metacognitive training.

Low math ability and poor attendance at discussion sections most significantly contributed to students failing the course. Course prerequisites were then implemented and attendance in discussion became required, resulting in a 3% increase in final exam median score and a semester decrease in student enrollment. Yearly enrollment did not change in response to the addition of course prerequisites.

Greater structure was added to the established flipped classroom through online homework questions, which linked students to videos and offered tutorials on the topic. Practice tests were added to this system, resulting in a 4% increase in median final exam score and a 6% improvement in pass rate.

Metacognitive training within practice tests was added to the homework, with students predicting their ability, receiving feedback of their chemistry ability, and developing study plans. This training overcame low-achieving students’ tendency to overestimate their chemistry ability but did not improve assessment scores.

Regular metacognitive training was implemented within weekly quizzes, in addition to practice tests, including the system of prediction, feedback, and study plans.

Student tendency to overestimate assessment scores was overcome and scores on each course exam improved. After factoring out other effects, overall final exam average improved by 4%, with the bottom quartile of students improving by 10%. This training was implemented over the year of general chemistry, improving the bottom quartile’s final exam score by 2.7% for each semester of training received. Analysis of this training indicated that regular completion of study plans and accurate prediction of end-of-semester assessment scores correlated to positive semester trajectory.

By the end of this study, student enrollment rates, the number of students passing the course, and the percent pass rates for General Chemistry I and II reached the highest levels seen since the outset. Final exam scores generally increased over the study, with a minor decline witnessed in the most recent semester of General Chemistry I.

Is the measure of a course how much a bright student learns
or how much someone who is “lost”
can be made to comprehend?

from *They’re Not Dumb, They’re Different, Stalking the Second Tier* by Sheila Tobias[1]

# TABLE OF CONTENTS

ABSTRACT ........ iii

LIST OF TABLES ........ x

LIST OF FIGURES ........ xi

ACKNOWLEDGMENTS ........ xiii

Chapters

1. INTRODUCTION ........ 1

    1.1 Historic Performance in General Chemistry ........ 1
    1.2 Research Goals ........ 2
    1.3 Research Summary ........ 3

2. EARLY CHANGES TO CHEMISTRY COURSE STRUCTURE ........ 6

    2.1 Development of Chemistry Prerequisites ........ 6
        2.1.1 Math Test Score Data Analysis ........ 6
        2.1.2 Math Course Grade Analysis for Students with Low Math Ability ........ 8
        2.1.3 Chemistry Preparatory Course Grade Analysis for Students with Low Math Ability ........ 9
        2.1.4 Implementation of Chemistry Prerequisites ........ 10
    2.2 Requirement of Discussion Attendance ........ 11
        2.2.1 Discussion Attendance and General Chemistry II Pass Rates ........ 12
    2.3 Results and Discussion ........ 15

3. PROVIDING GREATER STRUCTURE FOR THE FLIPPED CLASSROOM ........ 19

    3.1 Previous Research ........ 19
    3.2 Implementation of the Flipped Classroom at the University of Utah ........ 20
        3.2.1 Results and Discussion ........ 23

4. EARLY ITERATIONS OF IMPLEMENTING METACOGNITIVE TRAINING ........ 26

    4.1 Previous Research ........ 26
    4.2 Methods and Framework ........ 33
    4.3 Results ........ 37
        4.3.1 Effect of Metacognitive Training on Student Score Prediction Ability 37

4.3.2 Results: Effect of Metacognitive Training on Exam Scores .................. 40
4.4 Early Iteration of Metacognitive Training in General Chemistry II .................. 44
4.5 Summary of Initial Iterations of Metacognitive Training .................................. 49

5. IMPLEMENTATION OF WEEKLY METACOGNITIVE TRAINING EXERCISES .................................................................................................. 51

5.1 Previous Research .................................................................................. 51
5.2 Study Overview ...................................................................................... 52
5.3 Course Details ........................................................................................ 53
5.3.1 Control Semester ................................................................................. 53
5.3.2 Experimental Semester ........................................................................ 54
5.3.3 Metacognitive Training ........................................................................ 55
5.4 Experimental Details ............................................................................... 57
5.5 Results .................................................................................................... 61
5.5.1 Results: The Effect of Metacognitive Training on Score Prediction Ability ....................................................................................................... 61
5.5.2 Results: Effect of Metacognitive Training on Assessment Scores ....... 63
5.5.3 Results: Accounting for Other Factors: Incoming Ability and Attrition Rate .......................................................................................................... 66
5.5.4 Results: Accounting for Other Factors: Interaction Regression ............ 67
5.6 Summary of the Effects of Regular Metacognitive Training ............................. 71

6. THE CUMULATIVE EFFECT OF ONLINE METACOGNITIVE TRAINING ACROSS A YEAR OF GENERAL CHEMISTRY ........................................................ 74

6.1 Methods and Framework ........................................................................ 74
6.1.1 Details of the Control and Experimental Semesters .............................. 75
6.1.2 Details of the Metacognitive Training .................................................. 77
6.2 Experimental Details ............................................................................... 78
6.2.1 Total Number of Semesters of Metacognitive Training ........................ 79
6.2.2 Regression Details ............................................................................... 79
6.3 Results and Discussion ........................................................................... 82

7. THE EFFECT OF ONLINE METACOGNITIVE TRAINING ON GENERAL CHEMISTRY STUDENT COURSE TRAJECTORY ................................................... 88

7.1 Methods and Framework ........................................................................ 89
7.2 Experimental Details ............................................................................... 92
7.2.1 Course Trajectory ................................................................................ 92
7.2.2 Study Plans .......................................................................................... 94
7.2.3 Score Prediction Ability ....................................................................... 95
7.2.4 Accounting for Other Factors ............................................................... 96
7.3 Results and Discussion ........................................................................... 96
7.3.1 Course Trajectory ................................................................................ 96
7.3.2 Semester Trajectory Correlations ......................................................... 98
7.3.3 Regression Results ............................................................................... 99

7.3.4 Statistical Mediation ........................................................................ 103
7.3.5 Course Assessment Completion and Student Trajectory..................... 104
7.3.6 Discussion ............................................................................................ 106

8. THE LONG-TERM EFFECTS OF THIS RESEARCH EFFORT ............................ 109

8.1 Changes to General Chemistry I .................................................................. 111
8.1.1 Changes to General Chemistry I Enrollment ...................................... 111
8.1.2 Number of Students Passing General Chemistry I ............................. 111
8.1.3 General Chemistry I Pass Rate .......................................................... 113
8.1.4 General Chemistry I Final Exam Scores ............................................. 114
8.2 Changes to General Chemistry II ................................................................. 116
8.2.1 Changes to General Chemistry II Enrollment ..................................... 116
8.2.2 Number of Students Passing General Chemistry II ............................ 117
8.2.3 General Chemistry II Pass Rate ......................................................... 118
8.2.4 General Chemistry II Final Exam Scores ........................................... 120
8.3 Year-Long Enrollment and Pass Rates .......................................................... 120
8.3.1 Enrollment and Pass Rates in General Chemistry I ............................ 122
8.3.2 Enrollment and Pass Rates in General Chemistry II ........................... 123
8.4 Enrollment and Pass Rates in Organic Chemistry .......................................... 125
8.5 Conclusion .................................................................................................. 127

9. CONCLUSION ........................................................................................................ 130

9.1 Identifying and Resolving Early Challenges to Passing General Chemistry.. 130
9.2 Providing Greater Structure for the Flipped Classroom ................................. 132
9.3 Early Efforts toward Implementing Metacognitive Training .......................... 133
9.4 The Effect of Regular Metacognitive Training in General Chemistry I ......... 134
9.5 The Cumulative Effect of a Year of Metacognitive Training in General Chemistry ........................................................................................................ 136
9.6 The Effect of Metacognitive Training on General Chemistry Course Trajectory ....................................................................................................... 137
9.7 General Changes to Student Enrollment and Performance in General Chemistry ........................................................................................................ 139
9.8 Concluding Remarks ..................................................................................... 143

Appendices

A. SAMPLE ONLINE FLIPPED CLASSROOM HOMEWORK QUESTION ........... 145

B. SCREENSHOTS OF THE METACOGNITIVE TRAINING CYCLE ................... 147

C. CALCULATING INTERACTION SLOPE VALUES: A DIFFERENCE OF DIFFERENCES ........................................................................................................ 155

D. QUESTIONS FROM THE INTRO QUIZ FOR FALL 2016 .................................. 156

E. VENN DIAGRAM TO VISUALIZE THE EFFECTS ON TRAJECTORY ............ 159

F. SIX-YEAR CHANGES TO COURSE PASS RATE................................................ 160

REFERENCES .......................................................................................................... 161

# LIST OF TABLES

Tables

1.1 Timeline of research changes........................................................................... 4

2.1 Summary of General Chemistry I prerequisites......................................................... 11

3.1 Course structure for the control and flipped classroom semesters ............................ 23

4.1 Course structure comparison for the metacognitive and control sections .................. 37

4.2 Course structure comparison for the General Chemistry I metacognitive and control sections by year.................................................................................................. 43

4.3 Course structure comparison for the General Chemistry II metacognitive and control sections by year.................................................................................................. 46

5.1 Course structure comparison for the General Chemistry I metacognitive and control sections by year.................................................................................................. 59

5.2 Mann-Whitney *U* and effect size assessment comparison by course section............ 65

5.3 ACS final exam percent interaction regression results............................................. 70

6.1 Number of total possible semesters of metacognitive training.................................. 80

6.2 Analysis groups and number of semesters of metacognitive training ........................ 82

6.3 Regression results of the cumulative effect of metacognitive training....................... 83

7.1 Student percentile ranks over time and corresponding course trajectory slopes ........ 94

8.1 Timeline of research changes...................................................................... 110

# LIST OF FIGURES

Figures

1.1. Six year graduation rate vs general chemistry grade .................................................. 2

2.1. ACT math score vs. General Chemistry I pass rate .................................................... 7

2.2. General Chemistry I pass rate for low-achieving math students by college algebra grade. ......................................................................................................................... 9

2.3. General Chemistry I pass rate for low-achieving math students by their “Prep for General Chemistry” grade ............................................................................................ 10

2.4. General Chemistry II pass rate by General Chemistry I grade, divided into three general grade categories ................................................................................................ 12

2.5 Discussion percentage earned in General Chemistry II compared to student General Chemistry I grade category ............................................................................................. 13

2.6 General Chemistry II pass rate vs discussion percentage and General Chemistry I grade category ................................................................................................................ 15

2.7 General Chemistry I discussion percentage comparison between the fall 2013 and fall 2014 semester ................................................................................................................ 16

2.8 Comparison of student percentages on the ACS general chemistry final exam during the fall 2013 and fall 2014 semesters ................................................................................ 18

3.1 Locations of course structure for traditional and flipped classroom ........................... 20

3.2 General Chemistry II histogram comparison of the control and flipped classroom semesters ........................................................................................................................ 24

4.1 The metacognitive learning cycle, adapted from Dunlosky, 2009 ............................. 27

4.2 The Dunning-Kruger effect in general chemistry ...................................................... 28

4.3 Change in exam score averages by language ability and practice test type ................ 30

4.4 Perceived learning and actual ability across cycles of studying and testing ............... 31

4.5 The metacognitive cycle completed during the practice test week ............................. 35

4.6 Practice test prediction accuracy for the first set of practice tests .............................. 38

4.7 Practice test prediction accuracy for the final set of practice tests. ............................ 39

4.8 Dummy coding of course sections within the interaction.......................................... 42

5.1 Score prediction comparison between the intro and final quizzes............................. 62

5.2 *Z*-score comparison of the metacognitive and control sections on each major assessment.................................................................................................................. 64

5.3 Histogram of final exam percentages between the control and metacognitive sections ..................................................................................................................... 66

5.4 Final exam interaction plots for the overall course and by quartile ........................... 72

7.1 Histogram of trajectory slope values for General Chemistry I.................................. 97

7.2 Three perspectives of a 3D scatterplot.................................................................... 102

8.1 Fall General Chemistry I enrollment over time.. ..................................................... 112

8.2 Fall General Chemistry I number of students enrolling over time.. ......................... 113

8.3 Fall General Chemistry I pass rate over time.......................................................... 114

8.4 Overall and bottom quartile General Chemistry I final exam averages over time ... 115

8.5 General Chemistry II enrollment over time.. .......................................................... 117

8.6 Number of students passing General Chemistry II over time.................................. 118

8.7 General Chemistry II pass rates over time............................................................... 119

8.8 Overall and bottom quartile General Chemistry II final exam averages. ................. 121

8.9 Yearly total of students enrolled in and passing General Chemistry I. .................... 124

8.10 Yearly total of students enrolled in and passing General Chemistry II. ................. 125

8.11 Number of students enrolled in and passing organic chemistry I over time........... 126

8.12 Pass rate in organic chemistry I over time ............................................................ 127

# ACKNOWLEDGMENTS

First, I wanted to thank the individuals at Madra Learning who provided the homework platform that made this research possible. I am very grateful to my research group and my committee, all of whom offered advice and criticism that was essential in the progress of this work. Most of all, I need to gratefully acknowledge the support of Dr. Atwood, my committee chair, research advisor, and mentor. His continual support on this project has been invaluable, and he has taught me much about life through his wonderful example. I have grown immensely as a researcher and a person through him.

I am forever grateful for my parents who have been a continual support throughout my life and who provided support during the most challenging time of this process. Finally, I want to acknowledge the continual support of my loving wife, who has been my rock during the final years of my graduate work. Without her, this project would never have come to fruition.

# CHAPTER 1

# INTRODUCTION

General Chemistry Is required for a substantial number of students on the University of Utah campus. From 2000 to 2012 over 16,000 students enrolled in General Chemistry I. Of these students, survey results indicate that approximately 80% must pass a minimum of two semesters of general chemistry to meet the prerequisites for their degree. In 2012, Dr. Atwood was employed to improve student performance in general chemistry, and his direction has guided the efforts of this research. His efforts are timely, as the University of Utah created the Comprehensive Task Force on Enhancing Graduation Rates in 2013 to improve retention and graduation rates among all students. The purpose of this dissertation is to summarize the efforts and results of the research efforts to improve general chemistry student performance and pass rates.

## 1.1 Historic Performance in General Chemistry

Historically, General Chemistry I has been one of the most failed courses on campus. As evidence of this, during the 2009-2010 academic year, the Office of Budget and Institutional Analysis (OBIA) reported that General Chemistry I had a staggering 37% fail rate (D, E, or W). These students were tracked over the next six years in order to determine which of these students had graduated within that time period. Results indicated that 67.7% of students passing general chemistry ultimately graduated from the U of U within 6 years after taking the course. However, only 33.9% of students who failed general chemistry graduated from the University of Utah within six years of taking General Chemistry I (see Figure 1.1).

## 1.2 Research Goals

The main goal of this research project was to improve the student pass rate in first semester general chemistry without lowering expectations of student performance. Several research efforts were implemented in both semesters of general chemistry with hopes of improving the retention and performance of students through General Chemistry II. To accomplish this goal, we set out to accomplish three tasks:

1) Identify factors that are highly predictive of student failure.
2) Change the general chemistry course structure to improve student performance.
3) Create a system to help students practice effective learning skills in the classroom.

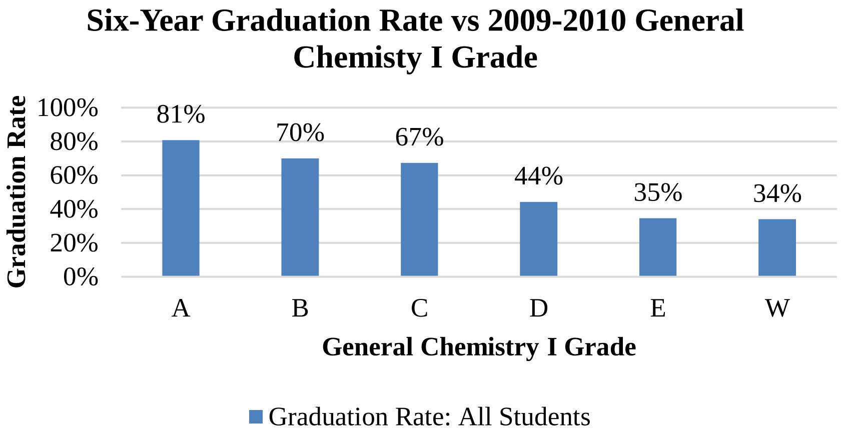


Figure 1.1. Six-year graduation rate vs general chemistry grade. 67.7% of those passing General Chemistry I graduated vs. 33.9% of students failing the course (*N* = 2336)

We hoped that implementing these tasks would help students be better prepared to enter general chemistry and be more successful in the course, resulting in an increase in course pass rate and retention in general chemistry.

## 1.3 Research Summary

To accomplish these goals, four major changes were made to the general chemistry structure and curriculum: institution of prerequisites based on student data, the requirement of attendance in TA discussion help sessions, the addition of more online homework structure to the flipped classroom system, and the implementation of a system of metacognitive training within the homework program. These changes were implemented in one or more sections of General Chemistry I and II over the past five years at the University of Utah. The timing and details of these changes are summarized in Table 1.1. In addition to these research efforts, it should also be noted that two other changes not part of this research project influenced student enrollment over the semesters. The first was the implementation of an online section of both semesters of general chemistry, which occurred during the 2016-2017 and 2017-2018 school years. Additionally, the university significantly increased freshman and sophomore enrollment during the fall 2017 semester.

In this document, Chapter 2 will summarize the theory, data, and results of the addition of prerequisites and the requirement of discussion attendance. Chapter 3 will cover the theory, incorporation, and results of the greater flipped classroom structure. Chapter 4 will provide the background for metacognitive theory and will describe the efforts and results of the early implementation of metacognitive training at the university.

Table 1.1 Timeline of Research Changes

| Semester | Summary | Details |
|---|---|---|
| Fall 2013 | Standardization of General Chemistry | All sections of general chemistry standardized with a common, flipped classroom curriculum |
| Fall 2014, Spring 2015 | Prerequisites and discussion attendance | General Chemistry I math and chemistry prerequisites established, discussion attendance required |
| Fall 2015 | General Chemistry I practice tests and metacognitive training | Implementation of practice test system, including ability feedback. Metacognitive Section: addition of score and ability prediction and study plan creation. |
| Spring 2016 | General Chemistry II practice tests and metacognitive training, increased flipped classroom structure | Implementation of practice tests, with feedback of ability by topic.<br>Metacognitive Section: addition of score and ability prediction and study plan creation.<br>Flipped Classroom: greater structure within the homework system during General Chemistry II |
| Fall 2016, Spring 2017 | Increased metacognitive training including weekly quizzes and practice tests, flipped classroom structure, study cycle homework | Implementation of practice tests and weekly quizzes, with feedback of ability by topic. Metacognitive Section: addition of score and ability prediction and study plan creation. Flipped Classroom: greater structure according to the study cycle within the homework system during General Chemistry I. |
| Fall 2017 | Study Assignments | Addition of targeted study assignments to the flipped classroom and regular metacognitive training. All sections receiving metacognitive training |

The following three chapters will summarize the results of the regular metacognitive training implemented during the fall 2016 and spring 2016 semester: Chapter 5 will describe the methods and results in General Chemistry I during the fall 2016 semester, with Chapter 6 summarizing the results of the spring 2017 semester and indicating a long-term and cumulative effect of the metacognitive training across a full year of general chemistry. Chapter 7 will analyze the effect of the metacognitive training on student trajectory during the semester. Chapter 8 summarizes the overall effects of the research efforts on changes in student enrollment, pass rates, number of students passing, and final exam scores over time. Chapter 9 will summarize the findings of this work. Finally, it should be noted that the data from Chapters 2-5 have been published previously.[2,3]

# CHAPTER 2

# EARLY CHANGES TO CHEMISTRY COURSE STRUCTURE TO IMPROVE GENERAL CHEMISTRY PASS RATE

Predicting student performance in general chemistry has long been of interest for chemical education researchers. Early research in chemistry education demonstrated that student math ability—especially measured by Scholastic Aptitude Test (SAT) or ACT math scores—was identified as a key factor in predicting student pass rates in general chemistry.[4,5] With these early findings in mind, we set out to determine student factors that were most predictive of pass rates in general chemistry.

## 2.1 Development of Chemistry Prerequisites

### 2.1.1 Math Test Score Data Analysis

Within the context of early findings in chemistry education research, we worked in conjunction with the OBIA to correlate student math background to general chemistry pass rates at the University. To determine the most predictive factors of student success in general chemistry, data were obtained from the OBIA from the 2000 to 2012 school years. For each set of data, students' final grade in General Chemistry I were compared to high school math test scores and previous courses taken at the University of Utah in math and other preparatory chemistry courses. First, Math ACT and SAT scores and grades for all math courses from Intermediate Algebra through Calculus II were obtained for all students.

The results indicated that ACT and SAT math test scores were highly correlated to General Chemistry I pass rate, with ACT math score having the greatest correlation ($R^2 = 0.936$, $p < 0.001$, $N = 12858$, see Figure 2.1) to pass rate in first semester general chemistry. Results of the SAT Math scores demonstrate a similar correlation between test score and pass rate ($R^2 = 0.759$, $p < 0.001$, $N = 2126$). It should be noted that the lower correlation for SAT math scores is likely due to decreased number of students in each score bin due to more score categories and that fewer students enrolled at the University of Utah take the SAT for enrollment. Additionally, test differences and differences in student demographics likely played a role in the decreased correlation. These results indicate that at the University of Utah, incoming high school math scores are highly predictive of chemistry pass rates.

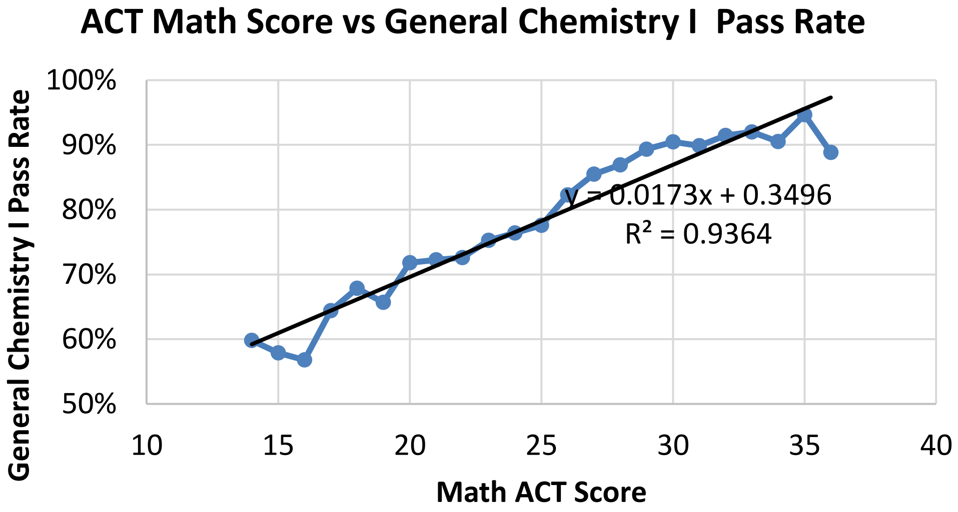


Figure 2.1. ACT math score vs. General Chemistry I pass rate ($N = 12858$).

### 2.1.2 Math Course Grade Analysis for Students with Low Math Ability

After math test score results were determined, students’ math and other course grades at the University of Utah were reviewed and compared to student scores on the ACT. To simplify the analysis, students scoring below 25 on the Math ACT were considered to have low math ability, while students scoring 25 and above were determined to have adequate math ability to enter General Chemistry I. From 2000 to 2012, the “low math ability” students were found to have a historic 74.2% pass rate in first semester general chemistry, while 89.6% of students who scored 25 or above passed General Chemistry I.

The goal of this process was to determine courses at the university that improved “low math ability” students’ probability of passing General Chemistry I. As part of the analysis, grades from several math and chemistry prep courses were obtained, and students’ grades in these courses were compared to their first semester chemistry grade received after taking the preparatory course. It should be noted that only students who received below a 25 on the Math ACT were included in the below analysis.

For students with low math ability, the course that was found to most significantly improve general chemistry pass rate was college algebra. For students scoring below 25 on the math ACT, passing college algebra appreciably improved General Chemistry I pass rates (see Figure 2.2, $N = 3472$). Results from OBIA indicated that students passing college algebra, having lower math ACT scores, had an 81.8% pass rate in General Chemistry I, in comparison to an overall historic 74.2% pass rate for all students scoring below 25 on the math ACT.

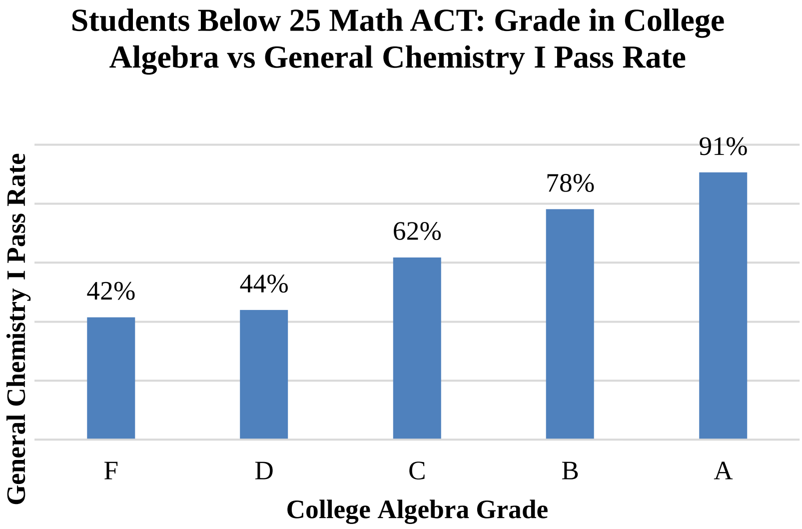


Figure 2.2. General Chemistry I pass rate for low-achieving math students by college algebra grade (*N* = 3472).

### 2.1.3 Chemistry Preparatory Course Grade Analysis for Students with Low Math Ability

In the past, the chemistry department at the University of Utah has implemented a number of preparatory courses to improve student pass rates in general chemistry. As with college algebra, students' grades in the various chemistry preparatory courses were obtained for students who scored below 25 on the Math ACT. Their grade in these prep courses was then compared to the grade received in General Chemistry I. Results indicate that the best preparatory course was Prep for General Chemistry. This course is a full-semester course that teaches students basic chemistry math skills necessary for general chemistry. Students who passed this course before taking General Chemistry I ultimately had an 83.3% pass rate in the first semester general chemistry (*N* = 624, see Figure 2.3).

These results indicate that students who receive below 25 on their math ACT are at-risk of failing general chemistry, having a historical 74.2% pass rate during the years

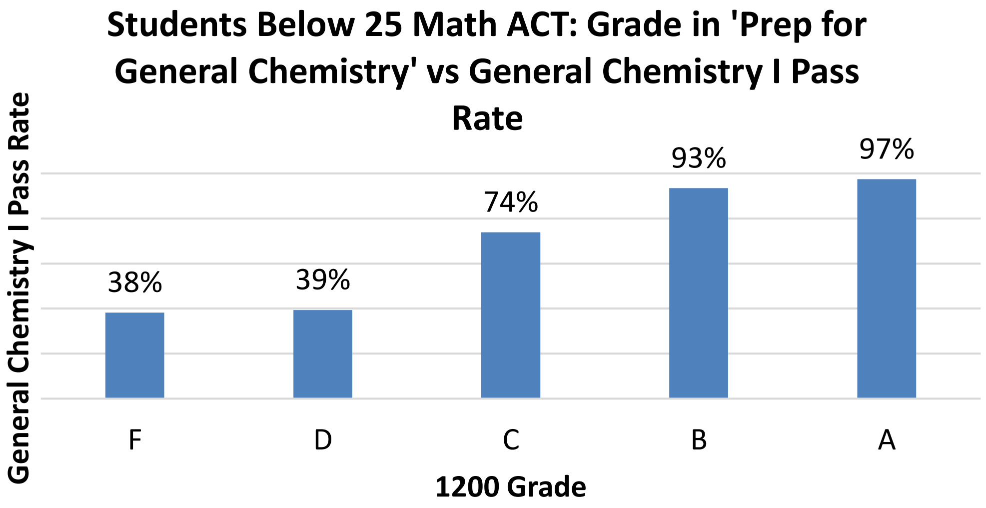


Figure 2.3. General Chemistry I pass rate for low-achieving math students by their “Prep for General Chemistry” grade ($N = 624$).

from 2000-2012. However, two key courses can assist “low math ability” students in developing adequate math and/or chemistry ability in preparation for General Chemistry I: college algebra and prep for general chemistry. “Low math ability” students who pass either of these courses before taking General Chemistry I ultimately improve their pass rate in first semester general chemistry to above 80%.

### 2.1.4 Implementation of Chemistry Prerequisites

In response to these findings, prerequisites were established for the General Chemistry I class, as summarized in Table 2.1, enforcing that students have adequate math and/or chemistry ability before entering the course. In order to enter the general chemistry class, students must either meet minimum math test scores or pass one of a number of math courses at the University of Utah. The first possible prerequisite was a minimum Math ACT, or corresponding Math SAT or Math Accuplacer, test score. The

Table 2.1 Summary of General Chemistry I Prerequisites

| Test or Course | Details |
|---|---|
| **Test Scores** | |
| -Math ACT | Score of 25 or greater |
| -Math SAT | Score of 600 or greater |
| -Math Accuplacer | Score of 75 or greater |
| **Course Grade** | |
| -College Algebra or Higher | Passing Grade |
| -Prep for General Chemistry | Passing Grade |

second prerequisite was a passing grade in any math class equal to or following College Algebra. Finally, students could also pass the Prep for General Chemistry course. In order to enroll in General Chemistry, I students must fulfill one of these requirements

## 2.2 Requirement of Discussion Attendance

In addition to the efforts resulting in prerequisite implementation, we were also interested in other course aspects that may impact student performance in general chemistry after enrolling in the course. To explore these factors, survey results of the General Chemistry II course during the spring 2014 semester were obtained regarding student demographic, test score, and university grade data. We then compared this data to pass rates in second semester general chemistry.

Comparing these results to students' end of semester final grades indicated that the best predictor of General Chemistry II pass rate was students' General Chemistry I grade ($R^2 = 0.456$, $p < 0.001$). At the end of the semester, students were retroactively grouped into three grade categories based on their incoming first semester general chemistry grade (see Figure 2.4). In particular, students who received below a C+ in General Chemistry I were considered at-risk in General Chemistry II, students with

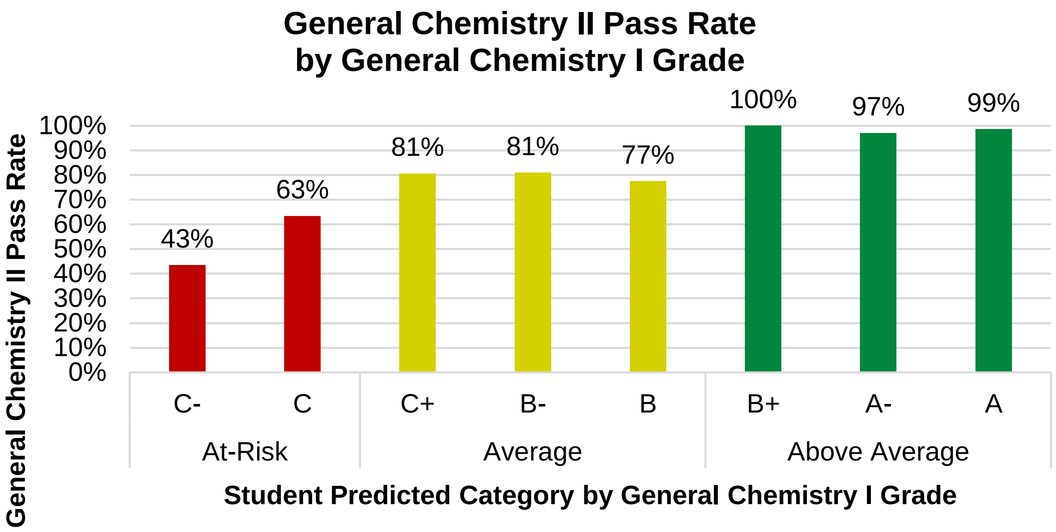


Figure 2.4. General Chemistry II pass rate by General Chemistry I grade, divided into three general grade categories (*N* = 483)

grades between a C+ and a B in General Chemistry I were considered average, and students who received a B+ or higher were designated as above average in the second semester chemistry. At-risk students had by far the lowest pass rates with only 59% passing General Chemistry II, average students had intermediate pass rates in the course, and nearly all above average students passed second semester chemistry

### 2.2.1 Discussion Attendance and General Chemistry II Pass Rates

As part of the General Chemistry II course structure, teaching assistant discussion help sessions were offered to students. Students enrolled for a particular discussion time and location before the semester began. During the course of the spring 2014 semester, attendance in discussion was not required, but attendance points in discussion were given as extra credit in an effort to boost attendance: students could receive up to 5% additional points added to their final grade at the end of the semester based on their attendance in discussion. Student response clickers were used in discussion to receive students' answers to questions, and students generally answered end-of-chapter type questions.

As part of the analysis to determine factors that are predictive of General Chemistry II pass rate, students' performance on discussion clicker scores were compared to the grade that students received in General Chemistry I. Students were then grouped into four categories based on the total percent of discussion points that they earned in the semester: 0% to 25%, 25% to 50%, 50% to 75%, and 75% to 100%. The results of comparing discussion clicker score group, students' incoming General Chemistry I grade, and General Chemistry II pass rates revealed a concerning trend (see Figure 2.5). Not surprisingly, students placed in the above average category tended to earn between 75% the possible discussion points. In fact, the majority of at-risk students received 0% to 25% of the possible discussion points, with relatively fewer above average students in each other discussion point category. In comparison, there was an over-representation of

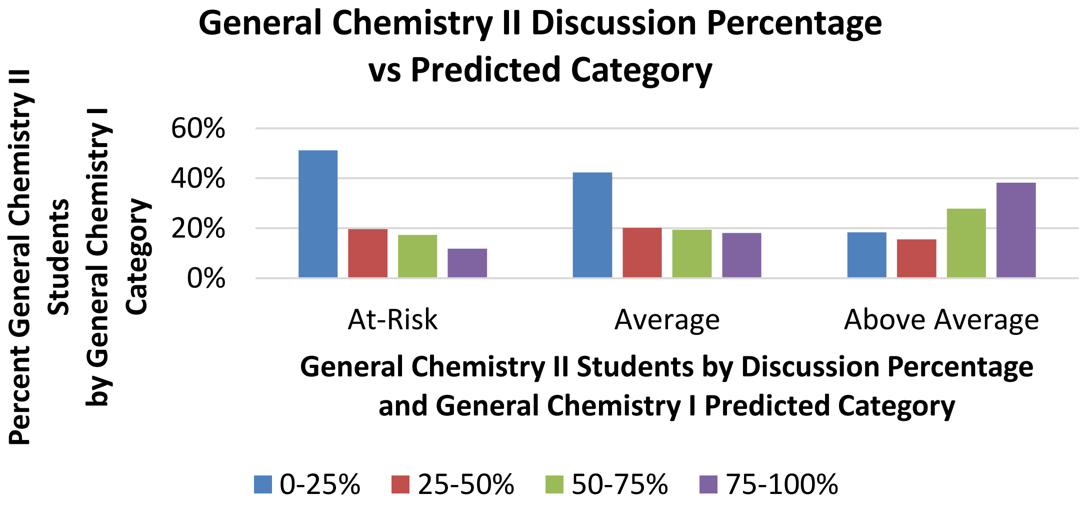


Figure 2.5 Discussion percentage earned in General Chemistry II compared to student General Chemistry I grade category (*N* = 483)

at-risk and average students who received between 0% and 25% of the possible discussion points.

After this analysis, students' end-of-semester General Chemistry II grades were compared to their incoming General Chemistry I grade category and their performance in the General Chemistry II discussion sessions. Note that discussion session extra credit was excluded from the final grade calculation to make this calculation independent of points earned in discussion attendance.

For students that were considered above-average by their General Chemistry I grade, pass rates did not significantly change based on their discussion attendance. However, for students designated as average or at-risk, pass rate improved dramatically as they increased in their percentage of discussion clicker points earned. For example, at-risk students who received less than 25% of the total discussion clicker points had a 39.6% pass rate in General Chemistry II, while students in this category who received more than 75% of total discussion points had a 92.9% pass rate in the second semester general chemistry course (see Figure 2.6). These results indicate that at-risk students are the least likely students to attend discussion when not required. However, when these at-risk students regularly attend discussion help sessions, their pass rates in General Chemistry II significantly improve.

In response to these findings, discussion attendance was made a required portion of the course. In subsequent semesters of General Chemistry I and II, students' course grades were calculated to include their points received in discussion. In particular, final course grades were calculated with discussion attendance representing 10% of their total grade.

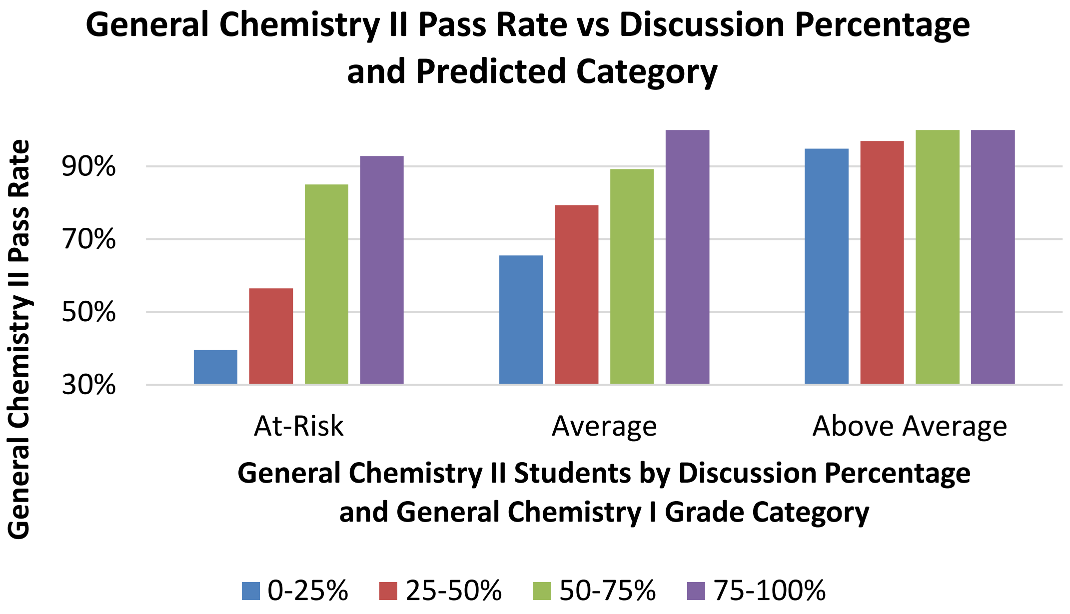


Figure 2.6 General Chemistry II pass rate vs discussion percentage and General Chemistry I grade category (*N* = 483)

## 2.3 Results and Discussion

Prerequisites for General Chemistry I and the requirement of discussion attendance were both implemented during fall 2014 semester. To track the effect of these changes to the general chemistry course setup, General Chemistry I course performance scores during the fall 2014 semester were compared to corresponding performance during the fall 2013 semester. In particular, course enrollment, overall percentage in discussion, and percent on the ACS exam were compared between the two semesters.

As prerequisites were implemented for enrollment in General Chemistry I during the fall 2014 semester, some decrease in enrollment during this semester was expected. Enrollment figures obtained from OBIA indicate that while 994 students enrolled in first semester general chemistry during the fall 2013 semester, only 921 students enrolled in the subsequent fall semester: a 7.3% decrease in enrollment. This decrease was likely influenced by fewer students able to enroll in the fall semester as a result of the implementation of course prerequisites for General Chemistry I.

During the fall 2014 semester, attendance in discussion was made a required portion of the course, representing 10% of students' total grade. To test the effect of this change, student percentages of total discussion points earned in General Chemistry I during the fall 2014 semester were compared to the discussion attendance during fall 2013 semester in first semester chemistry. Results indicate that the average discussion percentage rose substantially (see Figure 2.7).

At the end of each semester, the 2009 form of the ACS General Chemistry I final exam was used to evaluate students' overall knowledge regarding the first semester chemistry content. Because the final exam results are negatively skewed, traditional *t*-Test comparisons could not be made. Instead, the Mann-Whitney *U* statistic was generated to compare final exam scores between the two semesters. This statistic

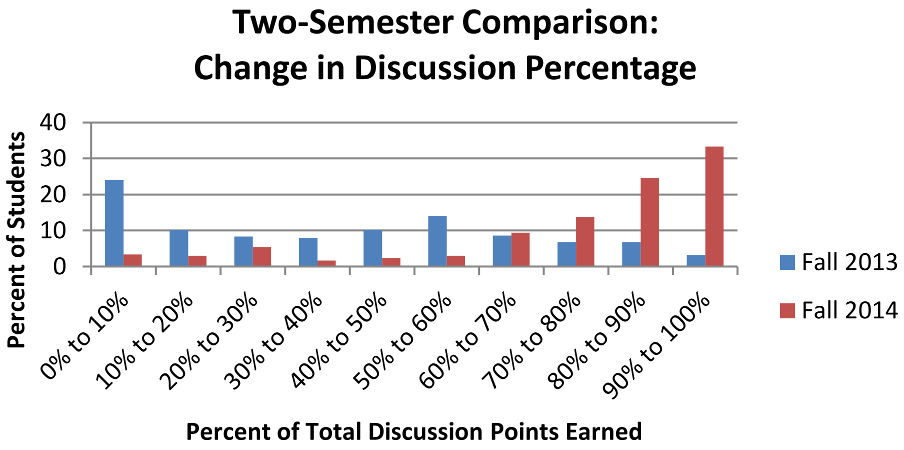


Figure 2.7 General Chemistry I discussion percentage comparison between the fall 2013 and fall 2014 semester (*N* = 609). Note the substantial increase in attendance during the fall 2014 semester as a result of the requirement of discussion in fall 2014.

compares the medians of each sample set to determine if the distributions are statistically significantly different from each other. Results indicate that the median score rose from 64.3% during the fall 2013 semester to 67.1% in the fall 2014 semester. Results of the Mann-Whitney *U* comparison indicate that the distributions are statistically significantly different from each other ($U = 291879.5$, $p < 0.001$).

Additionally, the *r* effect sizes were calculated according to (1), where *Z* is the Mann–Whitney *U* or Wilcoxon signed rank test *Z*-score and *N* is the total sample size:

$$r = \frac{Z}{\sqrt{N}} \tag{1}$$

Effect size *r* values are interpreted as follows: 0.1 is considered a small effect, 0.3 is considered a medium effect, and 0.5 is considered a large effect.[6] Results of this comparison yielded an effect size of 0.12, indicating that these changes resulted in a small effect on final exam score. These results indicate that student performance on the ACS final exam significantly improved as a result of the changes made during the fall 2014 semester (see Figure 2.8).

These results demonstrate that performance in discussion percentage significantly increased as a resulted of making discussion a required portion of the course.

Additionally, results indicate that the combined effect of implementing prerequisites and requiring discussion significantly improved ACS final exam average percent. It should be noted, however, that there was a 7% decrease in enrollment as a result of implementing the prerequisites, which may be influencing these results. Further analysis of changes in student enrollment over time will be discussed in Chapter 8.

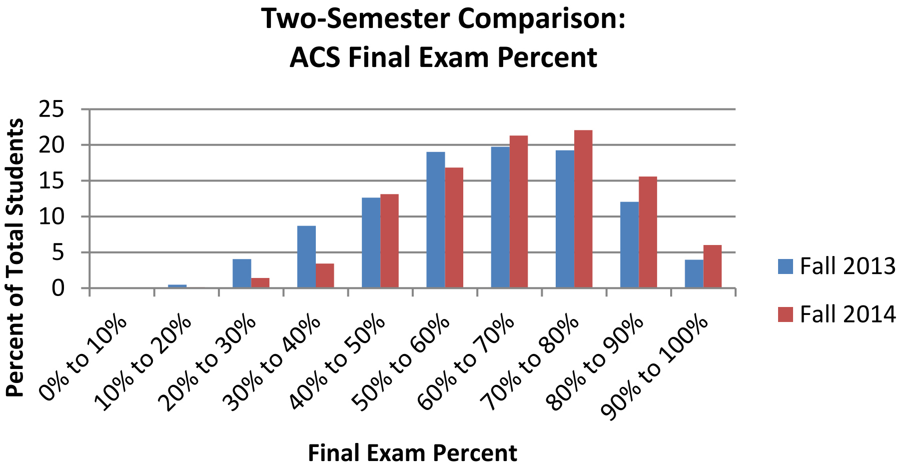


Figure 2.8 Comparison of student percentages on the ACS general chemistry final exam during the fall 2013 and fall 2014 semesters (*N* = 1644). During fall 2014, prerequisites were established and discussion was required, resulting in a 2.8% increase in median score.

# CHAPTER 3

# PROVIDING GREATER STRUCTURE FOR THE FLIPPED CLASSROOM

## 3.1 Previous Research

The flipped classroom is a relatively new instructional methodology, which was first described in chemical education research in 2014 by Michael Christiansen, though he used the term "inverted teaching." In general, the methodology involves a reversal of the locations of where course materials—especially lecture and problem solving—are typically presented. In his study, the flipped classroom was implemented in a sophomore organic chemistry course. Students watched recorded lectures, took notes, recorded their questions, and studied at home. As a result, class time was then opened up for students to work together on problem solving and ask questions from the instructor. In the study, he found that performance on course exams increased while not increasing time spent by students on course material (see Figure 3.1).[7]

Since this study, the flipped classroom has been widely researched in chemistry education. This system was found to improve exam performance, especially among the bottom-performing students, and reduce percentage of failing students within general chemistry.[8] Additionally, it was found that students' initial perceptions to the flipped classroom were mixed, but over the semester students responded increasingly positively

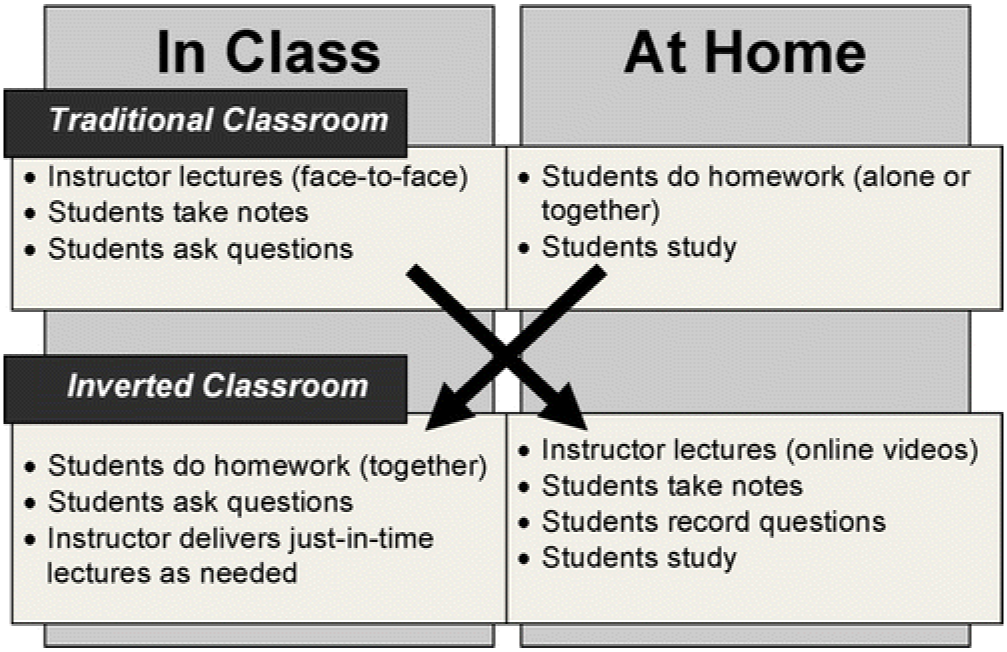


Figure 3.1 Locations of course structure for traditional and flipped classroom[5]

to the flipped classroom implementation.[9-11] The improvement in course performance and course satisfaction as a result of the flipped classroom system has been ascribed to the decrease of cognitive load on students, resulting in greater intellectual accessibility in the course layout.[12]

## 3.2 Implementation of the Flipped Classroom at the University of Utah

In the years previous to Dr. Atwood's appointment as the Ragsdale Professor of Chemical Education, general chemistry courses at the University of Utah were generally taught using the traditional lecture format. Additionally, student response devices (clickers) were only used by a few professors. However, during the 2013 semester, Dr Atwood implemented an early iteration of the flipped classroom in the general chemistry courses at the University of Utah.

As part of this flipped classroom format, lecture videos of content across general chemistry were recorded by Jeff Statler. Additionally, the curriculum was built around the arrangement of topics within the first edition of Zumdahl's *Chemistry: An Atom's First Approach*.[14] Students were directed to watch the lecture videos outside of class, while time spent in class was primarily focused on problem solving. To facilitate this, a common set of lecture slides were developed for the general chemistry courses, and students responded to in-class questions using clicker devices. As a result of this early implementation, the pass rate in General Chemistry I increased from 67.0% in the fall 2012 semester to 72.4% in the fall 2013 semester. Upon review of this implementation, survey results indicated that few of the students utilized the videos and textbooks in their studying for the course, perhaps due to the lack of accountability regarding utilizing this content. Additionally, though anecdotal, a discussion with students during this semester providing an interesting insight: though these students wanted to watch the videos, the lack of organization and structure made it difficult to find the appropriate videos related to course content, decreasing their motivation to utilize the flipped classroom structure.

With these challenges in mind, a homework system was designed to provide students with greater accountability and structure regarding the flipped classroom resources already in place. In particular, during the spring 2016 semester in the General Chemistry II course, several homework questions were developed within the Madra Learning online homework system to provide this necessary structure and accountability.[15]

As part of these questions, students were first directed to watch one or more of the previously generated videos regarding a particular topic. Additionally, students were directed to read a section of the textbook to preview a particular chemistry topic. The structure and phrasing of the questions indicated that students should complete this reading and watch the videos before moving forward in answering the flipped classroom questions.

After watching and reading this material, students answered a number of tutorial questions regarding the topic. These questions were designed to scaffold students in a possible thought process to guide them while solving the problems. An example of this type of question can be found in Appendix A. These questions were embedded into homework assignments due before the class and were meant to prepare students for the flipped classroom questions to be covered in the upcoming class material. As such, completing this homework assignment helped students preview the relevant material to prepare them for class time spent on problem solving regarding these topics. After class as part of the end-of-week homework, students also completed a limited number of review questions on the topics from previous week’s material. These questions were more traditional online homework questions, either conceptual or algorithmic, covering as many of the topics from the previous week’s material as possible. Additionally, the homework included practice tests, with one course section receiving metacognitive training and the other acting as a control (see Chapter 4).

This setup was first implemented during the spring 2016 semester in General Chemistry II, with the homework structure including flipped classroom preview questions. The differences between the sections of these two years are summarized in Table 3.1. To measure the effect of these changes, the 2009 form of the General

Table 3.1 Course structure for the control and flipped classroom semesters

| | | Flipped Classroom | |
|---|---|---|---|
| Course Structural Components: Fall 2015 | Control Year | Control | Metacognitive |
| Computer-based midterm exams | X | X | X |
| ACS conceptual final | X | X | X |
| ACS General Chemistry I final | X | X | X |
| Flipped classroom: Lecture videos and reading outside of class | X | X | X |
| Flipped classroom: Class time spent mostly solving problems | X | X | X |
| Flipped classroom: Homework video “preview” tutorial questions | — | X | X |
| Homework: Once weekly | X | X | X |
| Practice tests | — | X | X |
| Practice test: Topic feedback | — | X | X |
| Practice test: Score prediction | — | — | X |
| Practice test: Study plans | — | — | X |

Chemistry II ACS final exam was implemented during the two semesters. Doing so enabled researchers to compare students’ understanding of the chemistry material from one semester to the next. It should be noted that the same professors taught General Chemistry II during the spring 2015 and spring 2016 semester.

### 3.2.1 Results and Discussion

Results of the second semester ACS general chemistry exam demonstrated a significant improvement in exam performance from spring 2015 to spring 2016 (see Figure 3.2). Because the spring 2016 data was significantly negatively skewed, the nonparametric Mann-Whitney *U* test was used once more to compare the change in median exam percent from one semester to the next. From spring 2015 to spring 2016,

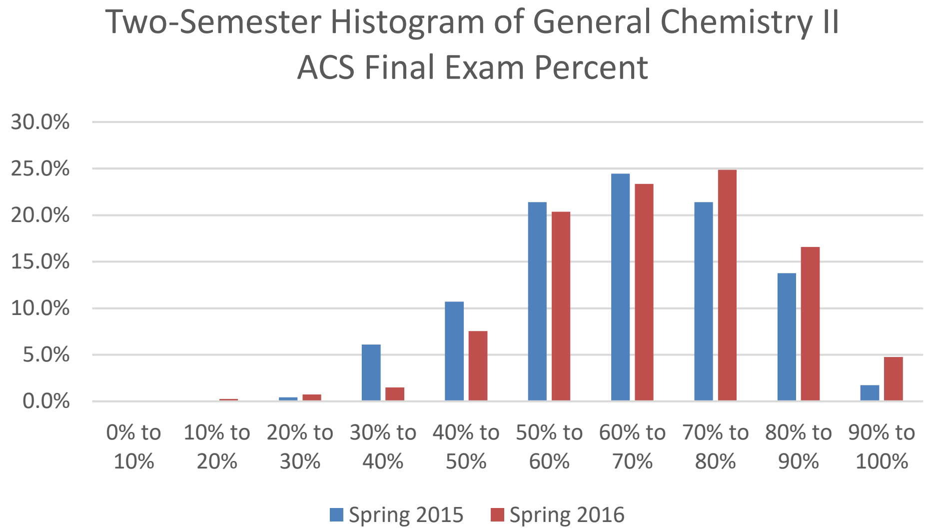


Figure 3.2 General Chemistry II histogram comparison of the control and flipped classroom semesters (*N* = 848). During the spring 2016 semester, greater structure was incorporated into the class, in addition to practice tests, resulting in a 4.3% increase in median score.

the ACS General Chemistry II final exam median percent increased from 64.3 to 68.6, a statistically significant 4.3% increase ($U = 77152.5$, $p < 0.001$). Effect sizes were once more calculated according to (1), with the results demonstrating a small effect size ($r = 0.13$). These results indicate that the addition of the flipped classroom structure and practice tests significantly improved final exam performance in General Chemistry II, with a statistically small effect.

One challenge with the results is the difficulty in separating out the individual effects of the addition the practice tests, metacognitive training, and the flipped classroom structure on student performance on the final exam. This being said, these results indicate that the changes made to the course structure, including enforcing the flipped classroom and the addition of metacognitive training and practice tests, significantly improved final exam performance. Though this result has demonstrated in previous chemical education research, it demonstrates a significant effort in aligning general chemistry education at the University of Utah with more statistically validated instructional methods. These results indicate that final exam scores continued to rise as the University of Utah brought the general chemistry course closer in alignment with the practices recommended in the latest chemical education research. Note that the individual effect of the metacognitive training as part of practice tests will be further discussed in Chapter 4.

# CHAPTER 4

# EARLY ITERATIONS OF IMPLEMENTING METACOGNITIVE TRAINING

## 4.1 Previous Research

Metacognition is defined as "thoughts about one's own thoughts and cognitions."[16] Metacognition involves two general concepts: metacognitive awareness, which is "assessing or evaluating the ongoing progress or current state of a particular cognitive activity"[16] and metacognitive control, which is "regulating an ongoing cognitive activity, such as stopping the activity, deciding to continue it, or changing it midstream."[16] Students utilize both aspects of metacognition during the learning process as follows: first, students compare their current level of understanding to their desired level through metacognitive awareness, and second they may adjust their study patterns, through metacognitive control, to accomplish their learning goals (see Figure 4.1).[16]

Students with high levels of metacognitive ability are more likely to "recognize when their ideas are not productive or cannot be reconciled with data or ideas presented by others."[17] Metacognition allows these students to "realize in what ways their understanding of concepts is incomplete."[17] The result is that highly metacognitive students are "more able and more likely to refine naive ideas in the face of contradictory experimental results. … Thus, students' own monitoring of their developing

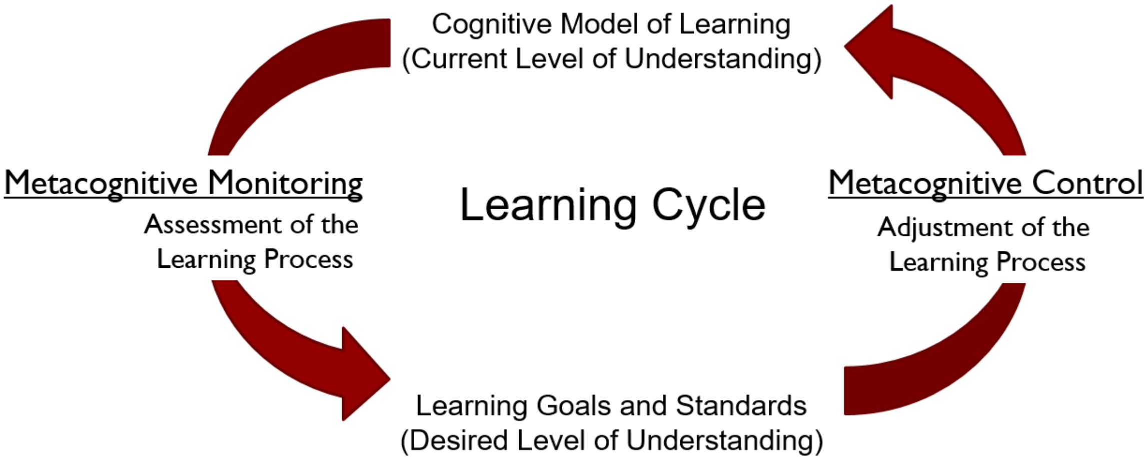


Figure 4.1 The metacognitive learning cycle, adapted from Dunlosky, 2009[16]

understanding of new concepts is essential for effective learning."[17]

Highlighting the importance of metacognitive awareness, Dunning and Kruger developed a study where participants predicted their ability on various psychological assessments. After completion of the assessment, the researchers compared participants' perceived ability to their actual ability. Not surprisingly, individuals who scored well on the assessments also accurately predicted their ability. However, "participants scoring in the bottom quartile…grossly overestimated their test performance and ability."[18] In other words, poor performing individuals also had poor metacognitive awareness of their ability.[18] Since this study, the Dunning-Kruger effect has been demonstrated repeatedly within chemical education research. One example had students predict their ability on each course exam, comparing their percentile ability to other students in the class. The high-achieving students were found to somewhat underestimate their percentile, while the bottom 25% of students dramatically overestimated their ability relative to other students in the course (see Figure 4.2).[19] More recently, another study had students "post-dict"

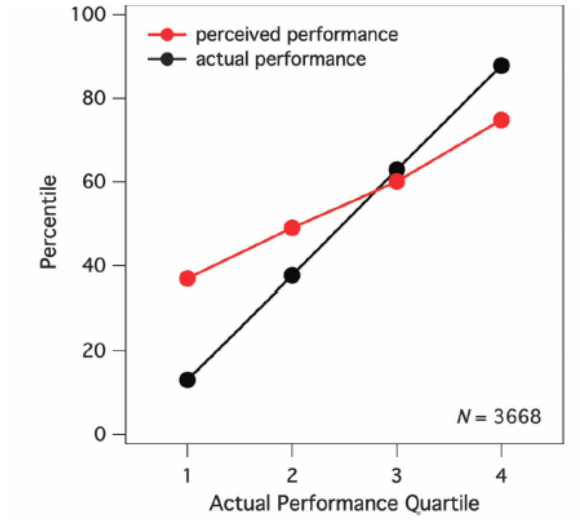


Figure 4.2 The Dunning-Kruger effect in general chemistry[19]

their exam scores after completing each general chemistry exam across the semester. The study not only demonstrated that the bottom students overestimated their score; it also found that their score overestimation remained stable following the first exam.[20] These results demonstrate that poor performing students will overestimate their chemistry ability, and this overestimation remains consistent, following the first exam, regardless of the number of exams taken throughout the semester.

The findings of Dunning, Kruger, and others demonstrate that there is a correlation with students' metacognitive awareness and their performance in the course. In other words, one reason that students may perform poorly is because they do not know what they do not know. In response to this need in chemical education, the following has been noted:

> If the goal of enhanced student understanding of chemistry is to be achieved, chemistry instructors will need to include instruction on the use of relevant thinking strategies in their courses. … Infusion of teaching of metacognitive skills in subject matter areas, in addition to general teaching of metacognitive skills in schools, is a promising approach to helping students learn to use their content knowledge more appropriately and flexibly. (ref *17*, p. 917)

Metacognition research within chemical education has generally focused on generating assessments to measure metacognitive ability and improving metacognitive ability during the problem-solving process.[21-24] However, there have been limited studies within chemical education on improving students' metacognitive ability in regard to preparation for course assessments. In one such study, students took a practice test before taking the American Chemical Society (ACS) general chemistry final exam, having students indicate the relative "mental effort" required for each question. The "test" group was considered those students who took the practice test, with the "control" group made up of students who did not take the practice test. Surprisingly, general results indicated that those who took the practice test generally did worse than those who did not take the exam. Upon further analysis of students' previous test scores, the researchers found that students who performed poorly on previous exams, who also took the practice exam, improved in final exam score as a result of taking the practice test when compared to the control students. The researchers attributed this improvement on the final exam to an improvement in metacognitive awareness among poor performing students that took the practice test. However, these researchers also found that students who took the practice test and who did well on previous exams actually did worse on final exam score when compared to the control students. These results indicate a mixed effect of the practice test on improving student exam scores on the ACS final exam.[25]

A more recent study used the SAT critical reading test scores to group students into two language comprehension categories—low and high language comprehension. The students then received one of two types of practice tests before each midterm: "elaborative interrogation" pretests, which required students to explain the pretest concepts in depth, and traditional multiple-choice pretests. Students were tested during the semester using multiple choice exams, and their exam results were compared to student exam averages in a control class where students had the same course layout but did not complete pretests. “Elaborative interrogation” pretests had no significant effect on test scores, but students who completed the multiple-choice pretests were found to score significantly higher on course exams than the students in the control course. They also found that this “testing effect,” or improvement in exam score after taking multiple-choice practice tests, was greatest among students with low language comprehension[26] (see Figure 4.3).

Outside of chemical education, multiple educational psychology studies have demonstrated additional methods of improving participants’ metacognitive ability. In the learning: they predicted how likely they would recall the pair on a test. They then took a

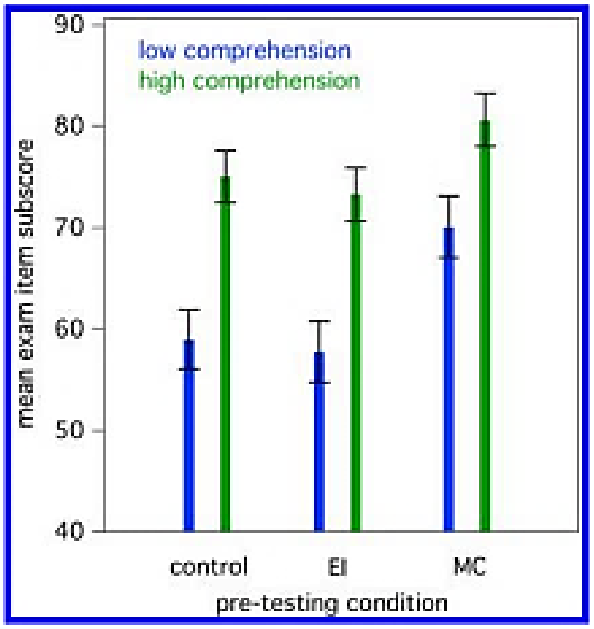


Figure 4.3 Change in exam score averages by language ability and practice test type[26]

test on the material, receiving feedback on the accuracy of their predictions. For the first test, individuals, in general, were overconfident in predicting their learning. After, participants completed a few more study, prediction, and testing cycles. Though participants overestimated their ability on the first exam, on the second and subsequent test, individuals generally under predicted their ability. Additionally, across the four tests, individuals consistently increased in ability, in addition to becoming better on the last two tests at predicting their actual ability (see Figure 4.4).[27] Results of this study demonstrate that with repeated cycles of studying, predicting, and testing, overconfidence can rapidly be overcome, with individuals becoming increasingly aware of their actual ability over time. Additionally, repeated cycles of study and testing tend to improve individuals' exam ability over time.

In another study, researchers had individuals take a test on general knowledge. After answering each question individuals rated how confident they felt about the accuracy of their confidence ratings. After receiving this feedback, they took two

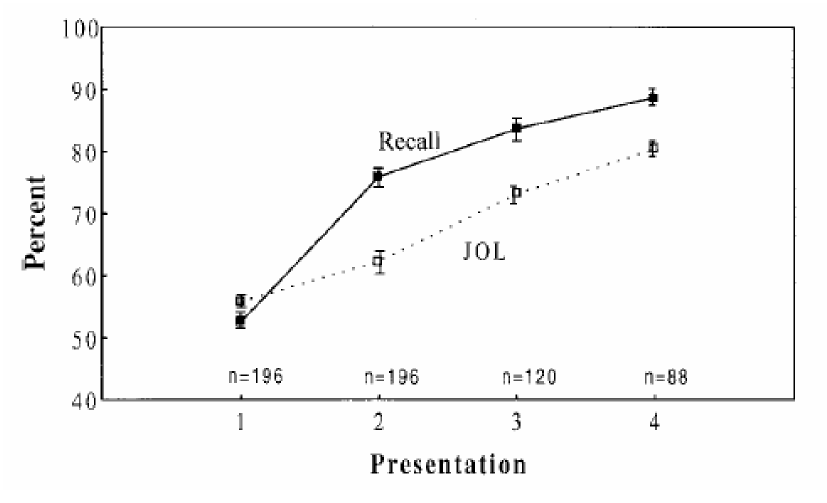


Figure 4.4 Perceived learning and actual ability across cycles of studying and testing[27]

subsequent tests, each time rating their confidence on test items, after which receiving feedback on the accuracy of their judgments. Notably, after taking the second test the accuracy of their confidence judgments markedly improved: while most of the individuals were overconfident in the first test session, individuals on average were not overconfident on the second test.[28] This study demonstrates that testing, while predicting ability, followed by extensive feedback on prediction accuracy, improves individuals' capability of predicting ability.

A final study involved a section of undergraduate educational psychology students who were regularly trained in metacognitive monitoring and control. Each week, the students were given a worksheet to rate their understanding of the day's content. They assessed which concepts were difficult and determined what they would do to improve understanding. Finally, they took a number of quiz questions and practiced rating their ability on these questions. Later, these students were given feedback on the accuracy of their monitoring. In the exams for both classes, students rated how confidently they felt about their answers for each question. A control section of the course received the same course content and exams but did not receive the weekly metacognitive training. Results of the first test demonstrated no significant difference between the two sections' score or monitoring ability. However, as the semester continued, students in the treatment section demonstrated significant improvement in accuracy predictions and test scores compared to the control.[29]

From these studies, researchers determined several aspects necessary to improve student metacognitive ability in test taking: repeated cycles of study and testing, prediction of ability and detailed feedback on prediction accuracy, regular assessment of ability by topic, making plans to improve that ability, and receiving feedback on this assessment.

## 4.2 Methods and Framework

During the fall 2015 semester, metacognitive training was implemented within the General Chemistry I homework structure. In preparation for the semester experiment, details of the experiment were submitted to the University's Institutional Review Board (IRB) for approval. A status of IRB exempt was received for the experiment. Course sizes in both sections were large, each having approximately 300 students enrolled in the course. Additionally, the courses were largely composed of college freshman and sophomores who were primarily seeking degrees in engineering or natural sciences or who were taking the course as a prerequisite for a professional health degree.

During the semester, two professors taught the two sections and will be referenced throughout the next two chapters as the "control" and "metacognitive" professors. The course layout between the sections, besides the below experimental conditions, was held as constant as possible. Both the control and metacognitive sections shared the same syllabus, were presented the same slides in class, completed the same weekly homework assignments, and were assessed by the same practice tests and exams. Additionally, during this semester, the course structure in both sections involved an early university attempt at implementing the flipped classroom, where students were directed to watch relevant lecture videos and read textbook sections in preparation for class.[7-12] During class, the bulk of the time was spent having students work through chemistry questions using audience response clicker devices. At the end of each week, students completed an online homework assignment answering several questions regarding the previous week's material.

In addition to this course structure, practice tests were implemented with the goal of training students in improving their metacognitive ability. These practice tests were designed after a review of an early draft of the exam, with the topics on each practice test mirroring those covered on the midterm and final exams. However, special effort was put into selection of different types of questions on the practice tests when compared to the midterms. As such, these practice tests were meant to encourage students to study the topics instead of memorizing problem types. Additionally, though students could take up to four versions of the practice tests, they were informed that their practice test score for each unit would be based on their best score received on the set of practice tests. Each practice test, as well as the homework assignments, were developed within the Madra Learning homework program.[15]

In the metacognitive section, three interventions were made to train students in metacognition: assessment score prediction, topic ability feedback, and topic study plans (see Figure 4.5). The first two interventions were incorporated into each practice test, topic feedback was emailed to students after the assessment closed, and students completed a study plan as part of a subsequent homework assignment.

These interventions were carried out as follows, with sample screenshots of each step included in Appendix 2.

1. Students opened the online practice test. Before being allowed to view any assessment questions, they were asked to predict their score on the assessment.

Figure 4.5 The metacognitive cycle completed during the practice test week

They were also informed that they would receive a small increase in their homework score, out of four possible points, based on the accuracy of their closest score prediction. This score was awarded as part of a separate assignment, independent of the assessment score, to ensure that the method of scoring practice tests between the two sections was identical. Additionally, for each practice test, students predicted their general Likert-scale ability on the concepts, calculations, problem-solving ability, and general ability of the upcoming assessment as follows: well below average, below average, average, above average, and well above average.

2. Students then took the assessment, "post-dicting" their score and Likert-scale ability after completing the assessment.
3. After the test closed, item response theory (IRT) was used to analyze each student's overall ability and their ability on questions regarding problem solving, calculation, and conceptual questions. Additionally, IRT analysis was used to determine students' ability on each chemistry topic in the assessment. This information was individually emailed to each student using their university email account.
4. Using the emailed feedback, students completed an online study plan, selecting

categories they would focus on in exam preparation. Students were given a series of possible methods to improve their ability by category, which they selected from to focus their studying. The feedback email and study plans were designed to improve metacognitive awareness by topic as well improving metacognitive control of how they would improve their knowledge in these topics.

The effect of metacognitive training was isolated by analyzing two sections:

1. Metacognitive Section: this section completed all practice tests. For each set of practice tests, the students in this section predicted their assessment score and Likert-scale ability on the assessment. After completing the assessment, these students post-dicted their score and ability once more. Once the assessment closed, students were emailed with feedback on their assessment score, their score prediction accuracy, and their ability on each assessment topic. These students were then asked to create a study plan on the subsequent homework assignment. Here, they selected topics to focus their study upon as well as activities to improve their ability in these topics.
2. Control Section: students in this section received the same practice tests. Additionally, this section received emails of feedback on their practice test score and their ability by assessment topic. However, students in the control section did not predict their assessment scores, and they did not create study plans.

When comparing the two course sections, there were two key differences between the sections: score and ability prediction and the creation of study plans (see Table 4.1). Consequently, the results of this study isolate the effect of these two interventions on student performance and predictive ability. Researchers hypothesized that the support for

Table 4.1 Course structure comparison for the metacognitive and control sections

| Course Structural Components: Fall 2015 | Control | Metacognitive |
|---|---|---|
| Computer-based midterm exams | X | X |
| ACS conceptual final | X | X |
| ACS General Chemistry I Final | X | X |
| Flipped classroom: Lecture videos and reading outside of class | X | X |
| Flipped classroom: Class time spent mostly solving problems | X | X |
| Homework: Once weekly | X | X |
| Practice test: Topic feedback | X | X |
| Practice test: Score prediction | — | X |
| Practice test: Study plans | — | X |

the metacognitive training would improve students' prediction abilities over time. In addition, test scores were hypothesized to be higher for the students receiving the metacognitive training than those who were not.

## 4.3 Results

### 4.3.1 Effect of Metacognitive Training on Student Score Prediction Ability

During the fall 2015 semester, the homework program was created that implemented practice tests designed to improve students' metacognitive abilities. Two sections of the course were compared during the 2015 semester: the metacognitive section, which received practice tests, predicted scores, received ability feedback, and created study plans, and the control section that only received the practice tests and topic feedback. Each section was grouped into quartiles based on exam performance to demonstrate the students who were most influenced by this study.

To test the effect of the score prediction during each practice test on students' metacognitive awareness during midterm 1 preparation, students in the metacognitive section were grouped into quartiles based on their performance on the first midterm. After, students' predicted score for each practice test was compared to their actual practice test score. Prediction scores were calculated for each student as "predicted score minus actual score," with a positive score representing an overestimation of score. Results indicated that all quartiles overestimated their ability on the first practice test. However, by the fourth practice test, all students were predicting their score on average to within 5% of their actual score (see Figure 4.6).

The same process was carried out for the final exam, with students once more grouped into quartiles by their score on the exam. These students' prediction scores were once more calculated by comparing their predicted score to their actual score on the practice tests. It should be noted that because students were grouped by specific exam quartile, the grouping of students into quartiles for the final exam does not necessarily

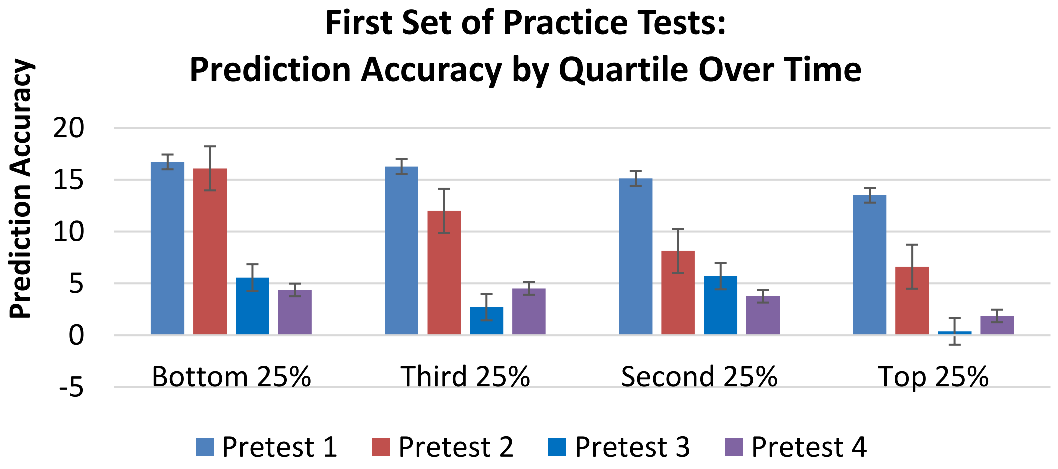


Figure 4.6 Practice test prediction accuracy for the first set of practice tests (*N* = 260). Note the significant score overestimation during the first practice test and the accurate prediction by the final practice test.

reflect the grouping of students by quartile for the first exam.

The results of the first practice test prediction scores for the final exam indicate the important result that overestimation of ability essentially disappeared for all students. Additionally, by the last practice test for the final exam, students in all quartiles tended to underestimate their scores, including students performing at the bottom 25% of the class on the final exam (see Figure 4.7). The Dunning-Kruger effect states that poor performing individuals will tend to overestimate their ability, and research in chemical education has demonstrated that this over-prediction remains stable regardless of the number of exams taken by students. However, this result indicates the first time we are aware of within chemical education research that the Dunning-Kruger effect was overcome: poor performing students were able to accurately predict or even under-estimate their scores on practice tests within a system of repeated practice tests and regular score prediction.

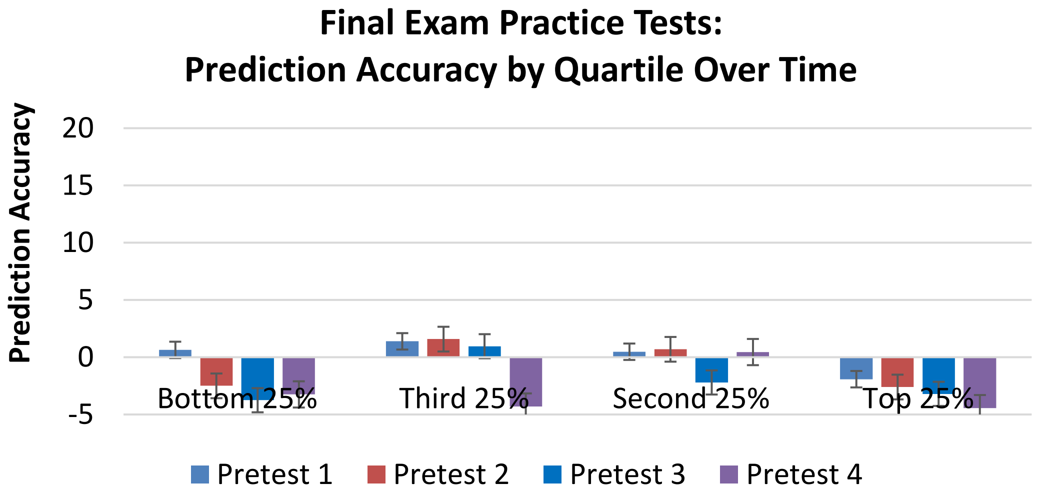


Figure 4.7 Practice test prediction accuracy for the final set of practice tests (*N* = 260). Note, no performance quartile overestimated scores for any final exam practice test.

### 4.3.2 Results: Effect of Metacognitive Training on Exam Scores

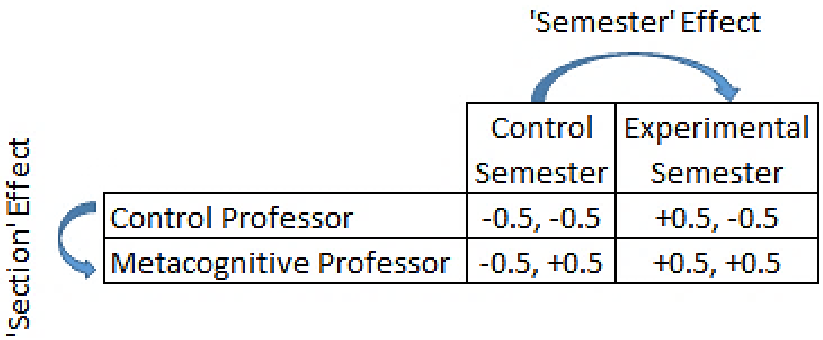
During the fall 2015 semester, the control and metacognitive professors taught the two sections at different times of day. These professors had different teaching styles, and the sections tended to have general differences in student demographics. For example, survey results have indicated a greater percentage of freshman students in the control section and a greater percentage of married students in the metacognitive section. Additionally, the professor of the control section generally has a greater tendency to lecture on occasion when compared to the metacognitive professor. Though results of a previous semester indicated that the professors of these sections perform statistically identically on the 2009 form of the ACS General Chemistry I final exam ($p = 0.42$), students taking the course from the professor who taught the metacognitive section performed slightly better than the control teacher's students on the final exam. As such, researchers were interested in factoring out any effect due to different teaching styles and demographics between the sections. To accomplish this, an interaction regression was performed, comparing the final exam performance between the two sections in this semester to a previous control semester.

In interaction regressions, "two variables are said to interact in their accounting for variance in [the independent variable] when *over and above* any additive combination of their separate effects, they have a *joint* effect" (italics in original text).[30] In the case of binary (two-level) variables, the "joint effect" represents "that over and above whatever average effects the two research factors have…a third source of [independent variable] variation, namely their joint or interaction effect is operating in the latter two data sets."[30] From a practical standpoint, the interaction effect is the *change* in the *average difference* of one effect when measured across a second effect.[30] To perform the regression, the final exam averages in the current experimental semester for both sections was compared to the same professors' final exam performance in a previous control semester. The regression evaluated two main effects: a "semester" effect and a "section" effect. The "semester" effect represents the general difference in student performance between the control and experimental semesters, and the "section" effect represents the general difference between the sections due to difference in teaching style and section demographics between the control and metacognitive sections. In addition, as part of the interaction regression, the "interaction" between the semester and section effects represents the additional effect of the metacognitive training after factoring out the other two effects. The regression model is given in eq 2, with $B_1$, $B_2$, and $B_3$ representing the unstandardized $B$ coefficients of regression effects, with $X_1$ representing the semester effect, $X_2$ representing the section effect, and $X_1 \times X_2$ representing the interaction between the semester and section effects:

$$Y = B_1 \mathrm{X}_1 + B_2 X_2 + B_3 (X_1 \times X_2) + Intercept \tag{2}$$

The variables for each were "dummy coded" accordingly: "–0.5" for the control semester and control section and "+0.5" for the experimental semester and metacognitive section. Finally, an "interaction" effect was created by multiplying the corresponding semester and the section dummy-coded values for a given set of data (see Figure 4.8).

During the control semester, the control and metacognitive professors taught, with



| | Control Semester | Experimental Semester |
|---|---|---|
| Control Professor | -0.5, -0.5 | +0.5, -0.5 |
| Metacognitive Professor | -0.5, +0.5 | +0.5, +0.5 |

Figure 4.8 Dummy coding of course sections within the interaction

class held at the same time as during the fall 2015 semester. Most of the course components remained generally the same between the control and fall 2015 semesters. The notable exception to this was the addition of practice tests, and corresponding metacognitive training, during the fall 2015 semester (see Table 4.2)

Regarding this study, if a significant interaction is observed between the semester and section effects, this interaction represents the effect of whatever was changed *between* the two professors during the experimental semester that was not present during the control semester. Though some teaching style and demographic differences were present during the control semester, these changes were perpetuated in the experimental semester. As such, the primary change between the sections during the experimental semester was the addition of the metacognitive training. Consequently, if a significant interaction is obtained, this value is interpreted as the effect of the addition of the metacognitive training during the experimental semester.

Results of this interaction regarding overall performances yielded no significant interaction results for any of the midterms or final exam. Notably, the interaction value for midterm 1 was 1.5%, indicating that metacognitive training may have accounted for a

Table 4.2 Course structure comparison for the General Chemistry I metacognitive and control sections by year

| | Control Semester | | Fall 2015 | |
|---|---|---|---|---|
| Course Structural Components: Fall 2015 | Control Professor | Metacognitive Professor | Control Professor | Metacognitive Professor |
| Computer-based midterm exams | X | X | X | X |
| ACS conceptual final | X | X | X | X |
| ACS General Chemistry I final | X | X | X | X |
| Flipped classroom: Lecture videos and reading outside of class | X | X | X | X |
| Flipped classroom: Class time spent mostly solving problems | X | X | X | X |
| Homework: Once weekly | X | X | X | X |
| Practice test: Topic feedback | — | — | X | X |
| Practice test: Score prediction | — | — | — | X |
| Practice test: Study plans | — | — | — | X |

slight improvement in overall student performance on the first midterm. However, this result was not statistically significant ($p = 0.28$), indicating a weak effect that could be due to chance. Additionally, the value of the interaction regression for the subsequent semesters decreased over time, with a slightly negative nonsignificant interaction value for the final exam ($B = -0.28$, $p = 0.88$). These results indicate that metacognition did not significantly improve exam performance for the class as a whole.

Researchers were interested in whether the metacognitive training might affect students differently based on their performance levels in the course. To measure this, students were grouped into performance quartiles based on their scores for each individual course exam. Interaction results of the first midterm demonstrated an encouraging 3.3% interaction value for the bottom 25% of the class, indicating that metacognitive training seemed to improve these students' first exam score by 3%.

Though this value was just outside of statistical significance (*p* = 0.053), the results seem to indicate that metacognitive training as part of practice tests might benefit the poorest performing subset of students. However, interaction effects for midterms 2 and 3 decreased dramatically, with metacognitive training no longer having any significant effect ($p_{midterm\ 2} = 0.814$, $p_{midterm\ 3} = 0.708$). On the final exam, the bottom quartile of students once more experienced a notable, though not statistically significant, improvement in final exam score as a result of metacognitive training. The interaction slope value for this population of students was 2.7% for the final exam (*p* = 0.16). Though outside of the range that is considered statistically significant, this result demonstrates that metacognitive training may have slightly improved final exam performance among the poorest performing students.

## 4.4 Early Iteration of Metacognitive Training in General Chemistry II

A similar study was performed during the spring 2016 semester in General Chemistry II, with a few additions to the fall 2015 study. In particular, greater structure was added to the study plan portion of metacognitive training and the flipped classroom was enforced to a greater degree as part of the online homework system. Details of the homework structure of the flipped classroom are presented in Chapter 3 of this work.

During this semester, greater structure was incorporated into the study plan portion of the metacognitive training with a goal of improving student metacognitive control. Once more, students completed practice tests before the actual exam, predicted their ability, and received feedback regarding their ability by chemistry topic. In this iteration, students were asked to identify chemistry topics they would like to improve their ability in, and then they selected specific tasks—such as reading the book or completing practice problems—they would do to improve their ability in those chemistry topics. Finally, students responded to a free response question, replying to what, where, when, and how they would study in preparation for the upcoming exam. As with the fall semester, one section of students received this metacognitive training, with the other section acting as a control. The control section completed the practice tests and received feedback on their ability by topic, but students in this section did not predict their scores or create the more detailed study plans. As with the fall 2016 semester, the course structures between the two sections, outside of the experimental conditions, were held as constant as possible, including a common syllabus, identical assessments, and identical classroom slides. Homework assignments and practice tests were once more were created within the Madra Learning homework program.[15]

Researchers were interested in quantifying the individual effects of the addition of greater structure to the flipped classroom, the addition of practice tests, and the implementation metacognitive training with more structured study plans. To do so, the final exam scores of the spring 2016 semester were compared to the final exam scores of a control spring 2015 semester, similar to the method performed for the fall 2015 semester. The teachers who taught during the spring 2016 semester also taught during the control semester. Online exams and the ACS final exams were utilized during both semesters, and similar course slides, focusing on problem solving, were presented to students during each year. The primary differences, then, between the control and spring 2016 semesters was the addition of the flipped classroom structure and the addition of practice tests with ability feedback by topic. Additionally, during the spring 2016 semester, the primary differences between the control and metacognitive sections were the addition of score prediction and study plans during the week of practice tests (see Table 4.3).

As with the fall 2015 semester, we created an interaction regression to separate out three effects—the semester, section, and interaction effects. These three effects were once more calculated by comparing the final exam data from the spring 2015 and spring 2016 General Chemistry II courses. As before, the "interaction" effect represents the effect of the metacognitive training. Also, the "section" effect represents the general difference in final exam score between professors, independent of the semester taught or the incorporation of metacognitive training. Finally, the "semester" effect represents the general difference in final exam scores from the control to the spring 2016 semester

Table 4.3 Course structure comparison for the General Chemistry II metacognitive and control sections by year

| | | Flipped Classroom, Study Cycle | |
|---|---|---|---|
| Course Structural Components: Fall 2015 | Control Year | Control | Metacognitive |
| Computer-based midterm exams | X | X | X |
| ACS conceptual final | X | X | X |
| ACS General Chemistry I final | X | X | X |
| Flipped classroom: Lecture videos and reading outside of class | X | X | X |
| Flipped classroom: Class time spent mostly solving problems | X | X | X |
| Flipped classroom: Homework video "preview" tutorial questions | — | X | X |
| Homework: Once weekly | X | X | X |
| Practice tests | — | X | X |
| Practice test: Topic feedback | — | X | X |
| Practice test: Score prediction | — | — | X |
| Practice test: Study plans | — | — | X |

independent of professor or the addition of metacognitive training. This final effect, then, represents the change in student final exam average as a result of the increased structure within the flipped classroom and the addition of practice tests in the course.

Once again, the interaction of the semester and section effects, which would have represented the effect of the addition of metacognitive training, demonstrated no significant overall change in final exam average scores (*p* = 0.90). In other words, the greater structure added to the metacognitive training during this semester had no significant effect on improving students' final exam average scores, similar to the results during the fall 2015 semester.

In hopes of finding the metacognitive training targeting a particular ability group, students were grouped into final exam score quartiles. The results of the fall 2015 semester demonstrated a weak but important improvement in midterm 1 and final exam ability by students scoring in the bottom 25% on the final exam. However, during the spring 2016 semester, the bottom quartile had no significant change in response to the metacognitive training (*p* = 0.92). Additionally, students in the 3rd quartile, who generally get C's in the course, experienced a significantly negative interaction value of 2.0% (*p* = 0.029). In other words, the C students in the course seemed to perform 2% worse on the final exam as a result of the addition of metacognitive training. The only potentially encouraging results of the metacognitive training occurred among the top 25% of students. The interaction regression slope for these students was 2.6%, a value just outside of statistical significance, indicating a potential improvement in these students' final exam scores as a result of the metacognitive training (*p* = 0.097). In other words, the only potential positive result of the metacognitive training during the spring 2016 semester was a slight improvement in performance among the already successful students in the course.

However, the other effects of the interaction regression demonstrated that the general changes made during the spring 2016 semester did help to improve student performance. Recall that the semester effect measures the effect of the general changes in the chemistry course from one year to the next, independent of professor or the addition of the metacognitive training. In particular, this value would measure the improvement in student performance on the ACS final exam average as a result of the addition of practice tests with topic feedback and the increased structure of the flipped classroom, both of which were not present during the control semester. Results of the regression of the overall final exam scores demonstrated a significant 3.8% semester effect for the overall class ($p < 0.001$). In other words, the addition of the flipped classroom structure, in conjunction with the addition of practice tests, improved student performance on the final exam average performance by 3.8%. In reviewing the results determined in Chapter 3, we found that there was a 4.3% improvement on final ACS exam median score from the control semester to spring 2016. The particular statistical test in Chapter 3 represents a change to students' final exam median from one year to the next, while this latter measure represents the change in the mean of final exam scores between years. However, these results seem to indicate that the majority of the improvement in General Chemistry II final exam performance during the spring 2016 semester occurred as a result of the incorporation of practice tests, the increased structure for the flipped classroom, and the homework support regarding the study cycle.

Encouragingly, all four quartiles experienced a significant semester effect, with the greatest improvement occurring among the bottom 25% of students. This category of students significantly improved by 4.5% on the ACS General Chemistry II final exam, with the 3rd quartile improving by 3.6%, the 2nd quartile by 3.5%, and the top 25% by 3.7% ($p < 0.001$ for all values). These results indicate that the addition of practice tests and the increased homework structure for the flipped classroom combined to improve all quartiles of students' final exam average score, with the bottom students improving to the greatest degree.

## 4.5 Summary of Initial Iterations of Metacognitive Training

Previous studies in chemical education have indicated that poor students tend to overestimate their ability on assessments according to the Dunning-Kruger effect.[19] Additionally, this overestimation of ability remains stable following the first exam regardless of the number of assessments taken.[20] In comparison, the results of this first iteration of the study indicated that the Dunning-Kruger effect can be overcome. Results of the fall 2015 iteration of metacognitive training initially indicated a general overestimation of ability by all students. However, by the final practice test before taking the first midterm, all students were generally predicting their practice test scores accurately. Even more importantly, by the time students took the practice tests for the final exam, they were able to accurately predict, or in some cases underpredict, their scores for each practice test. This was the first time, to our knowledge, that the Dunning-Kruger effect was found to be overcome within chemical education research.

In addition to this, there is some evidence that a system of metacognitive training, within practice tests, improved student performance, with the bottom 25% of the class improving by about 3% on the first midterm and final exam in the interaction regression. However, neither of these results were quite statistically significant. Additionally, there was no similar improvement in exam score on the second and third midterm. This indicates that the metacognitive training provided during the fall 2015 semester was not sufficient to consistently or significantly improve student performance across the first semester of general chemistry.

During the spring 2015 semester in General Chemistry II, no improvement in student final exam score was observed in the bottom quartile of students as a result of metacognitive training. In fact, the 3rd quartile of students seemed to decrease by 2% in final exam average performance as a result of metacognitive training. Only the top 25% of students seemed to improve as a result of the metacognitive training, potentially scoring 2.5% higher as a result of the training, though this improvement was not statistically significant. However, results of the semester effect of the interaction indicate that the addition of the flipped classroom and practice tests improved student performance across all quartiles by 3% to 4% on the final exam.

# CHAPTER 5

# IMPLEMENTATION OF WEEKLY METACOGNITIVE TRAINING EXERCISES

## 5.1 Previous Research

Before providing the details and results of this next study, it should be noted that, during this semester, the homework structure in the course was altered to enforce that students work through homework according to the study cycle. This cycle shows promising potential to improve student performance in the general chemistry classroom and involves several steps that students cycle through as part of test preparation: previewing course material before coming to class, attending class, reviewing the material after class, and studying for upcoming assessments. Previous research has demonstrated that teaching this cycle to students improves test performance. In one such study, researchers prepared a presentation on the study cycle immediately following students' first exam. Students in attendance that day were taught the steps of the study cycle and committed, in writing, to implementing some or all of these steps. As part of this assignment, they indicated which aspects of the study cycle they would utilize the remainder of the semester. The students not in attendance during this day of class did not receive the training and were considered the control group. They found that there was a significant improvement in later test scores among students that received training on the study cycle when compared to the control group.[31]

## 5.2 Study Overview

Results of the metacognitive training during the fall 2015 and spring 2016 semesters demonstrated that metacognitive training through practice tests resulted in an encouraging improvement in student metacognitive awareness, as measured by practice test score prediction ability. However, this training was found to have no significant effect on student performance on the ACS final exam, though results of the first midterm and the final exam demonstrate encouraging, though not statistically significant, improvements in final exam score by the lowest quartile of students in general chemistry 1 and the highest quartile of students in General Chemistry II.

Researchers hypothesized that these preliminary improvements in student chemistry performance could be increased through greater opportunity for student metacognitive reflection. As a result, we developed a third iteration of the experiment that was meant to expand students' opportunities for metacognitive training. This iteration included weekly quizzes—including ability prediction, ability feedback, and the creation of study plans—in addition to the already established metacognitive training as part of practice tests.

As with previous semesters, all homework, quizzes, practice tests, and midterm exams were required and completed within the Madra Learning online homework system.[15] Additionally, students took the 2009 form of the first term ACS general chemistry exam once more at the end of the semester.

As with the fall 2015 and spring 2016 semesters, the study compared the performance of the students in two professors' course sections across two semesters. During both semesters, the same two professors—the "control and "metacognitive" professors—taught the same two sections. Additionally, as with previous experiments, the results of the second "experimental" semester were compared against the results in an earlier "control" semester.

## 5.3 Course Details

### 5.3.1 Control Semester

During the control semester, both the metacognitive and control professors taught course sections. The course structure in both sections involved the early university attempt at implementing the flipped classroom, with students directed to watch relevant lecture videos and read textbook sections in preparation for class.[7-12] During class, the bulk of the time was spent having students work through chemistry questions using audience response clicker devices. At the end of each week, students completed an online homework assignment answering several questions regarding the previous week's material.

Efforts were made, once more, to ensure that the course setup was as similar as possible between the two course sections. Both professors utilized the same course syllabus, used the same flipped classroom slides, had identical homework assignments, and assessed students with the same exams. As with the fall 2015 semester, the control professor had a greater tendency to briefly lecture on the topics than the other. As the courses were taught at the same time of day from one year to the next, researchers assumed that the differences in demographics between the course sections, as summarized in Chapter 4, were perpetuated into the fall 2016 semester. Recall that results from the control semester indicated that there was no statistically significant difference between the average student performance on midterms or the final exam between the professors' sections. Finally, end-of-semester survey results indicated that students generally were not adequately utilizing the flipped classroom resources of the course during the control semester.

### 5.3.2 Experimental Semester

During the experimental semester, both the control and metacognitive professors taught once more. In addition to the experimental details of the semester, summarized below, a few general course changes were made in the experimental semester when compared to the control semester. As with the fall 2015 semester, students in both the control and metacognitive sections completed practice tests before each midterm and before the final exam to assist them in exam preparation. In addition to these practice tests, students in both sections completed weekly quizzes to assess their ability on the previous week's material. These quizzes were completed during each week of the course except for weeks where the practice tests were being taken. Each weekly quiz and homework assignment was required and included in their overall homework grade. Additionally, though students could take up to three versions of the practice tests, they were informed that their practice test score for each unit would be based on their best score received on the set of practice tests.

As utilization of flipped classroom resources was poor during the control semester, greater structure was implemented in the flipped classroom model during the experimental semester, mirroring the structure implemented during the spring 2016 semester: students completed “preview” questions that directed them to relevant textbook sections, linked them to online lecture videos, and had students work through a brief set of tutorial questions.

Within the general framework of the study cycle, students prepare for class, review the material after class, study, and are regularly assessed on the material. In the implementation of this structure, students completed preview video questions before class to prepare for the upcoming topics. In the subsequent homework assignment, which closed before the following class day, students also completed a limited number of review questions on the topics from previous week’s material. These questions were more traditional online homework questions, either conceptual or algorithmic, covering as much of the previous lecture material possible. Additionally, the homework included practice tests, with one course section receiving metacognitive training and the other acting as a control, as summarized in Chapter 4.

As with the control semester, the course layout between the sections, besides the below experimental conditions, was held as constant as possible between the two sections. Both the control and metacognitive sections shared the same syllabus, were presented the same flipped classroom slides in class, completed the same homework assignments, and were assessed by the same weekly quizzes, practice tests, and exams.

### 5.3.3 Metacognitive Training

As with the fall 2015 semester, three specific metacognitive training interventions were implemented: assessment score prediction, topic ability feedback, and topic study plans. The first two interventions were incorporated into each weekly quiz and during each practice test, while the study plans were incorporated into subsequent homework assignments after the completion of the quiz or practice test. These interventions were carried out with a few changes from the fall 2015 semester as follows, with sample screenshots of each step included in the Appendix B:

1. Students opened the online weekly quiz or practice test. Before being allowed to view any assessment questions, they were asked to predict their score on the assessment. For the practice tests, though not for weekly quizzes, students were informed that they would receive a small increase in their homework score, out of four possible points, based on the accuracy of their closest score prediction. This score was awarded as part of a separate assignment, independent of the assessment score, to ensure that the method of scoring practice tests between the two sections was identical. Additionally, for each weekly quiz and practice test, students predicted their general Likert-scale ability on the concepts, calculations, problem-solving ability, and general ability of the upcoming assessment as follows: well below average, below average, average, above average, and well above average.
2. Students then took the assessment, “post-dicting” their score and Likert-scale ability after completing the assessment.
3. Following the assessment, students received detailed feedback within the Madra Learning system on their ability on the assessment. This feedback included their score on the assessment and the accuracy of their score predictions. They were also given feedback of their ability on each assessment topic, as calculated by the

percent of assessment questions within a specific chemistry topic that they answered correctly. Finally, the homework system provided students with a list of potential study topics that they could focus future studying upon.

4. Once the assessment closed, students were allowed to review their results on specific quiz questions. On the following homework assignment, which was due two days after the assessment closed, students answered a required homework question to assist them in building a study plan. For this plan, students indicated how well they did (good, average, or poor) on the major chemistry concepts covered on the assessment. Additionally, for study plans generated after each practice test, students were also provided with a list of the specific topics, divided by the major chemistry area, covered on the assessment. From this list, they selected specific topics to focus their future studying upon. They were given a small completion-based score on this homework assignment for answering these study-plan homework questions.

## 5.4 Experimental Details

To isolate the effects of the metacognitive training, two course sections were analyzed:

1. Metacognitive Section: this section completed all weekly quizzes and practice tests. For each weekly quiz and each set of practice tests, the students in this section predicted their assessment score and Likert-scale ability on the assessment. After completing the assessment, these students post-dicted their score and ability once more. Students were then provided with feedback on their

assessment score, their score prediction accuracy, their accuracy on each assessment topic, and key areas they should focus their studying upon. Once the assessment results were released, these students created a study plan on the subsequent homework assignment. Here, they indicated their ability on the general chemistry concepts covered on the assessment. Additionally, for the practice tests, they selected specific topics within each chemistry area to focus their future studying upon.

2. Control Section: students in this section received the same weekly quizzes and practice tests. Additionally, this section received identical feedback on their quiz score, their ability by assessment topic, and key areas they should focus their studying upon. However, students in the control section did not predict their assessment scores, and they did not create study plans.

When comparing the two course sections, there were two key differences between the sections: score and ability prediction and the creation of study plans (see Table 5.1). Consequently, the results of this study isolate the effect of these two interventions on student performance and predictive ability.

To test the effect of the metacognitive training, we were first interested in whether metacognitive monitoring improved in the metacognitive section as a result of the regular score prediction for each assessment. In particular, we were interested in whether the Dunning–Kruger effect—in particular the general over-prediction of assessment scores by poor students—would be overcome once more with regular score prediction.

Two quiz predictions were compared in this analysis: the intro quiz, which was completed the first week of the semester, and the final quiz completed by students in the

Table 5.1 Course structure comparison for the General Chemistry I metacognitive and control sections by year

| Course Structural Components | Control Semester Sections | | Experimental Semester Sections | |
|---|---|---|---|---|
| | Control | Metacognitive | Control | Metacognitive |
| Computer-based midterm exams | X | X | X | X |
| ACS conceptual final | X | X | X | X |
| ACS General Chemistry I final | X | X | X | X |
| Flipped classroom: Lecture videos and reading outside of class | X | X | X | X |
| Flipped classroom: Class time spent mostly solving problems | X | X | X | X |
| Flipped classroom: Homework video “preview” tutorial questions | — | — | X | X |
| Homework: Once weekly | X | X | — | — |
| Homework: Three per week | — | — | X | X |
| Weekly quiz | — | — | X | X |
| Weekly quiz: Topic feedback | — | — | X | X |
| Weekly quiz: Score prediction | — | — | — | X |
| Weekly quiz: Study plan | — | — | — | X |
| Practice tests | — | — | X | X |
| Practice test: Topic feedback | — | — | X | X |
| Practice test: Score prediction | — | — | — | X |
| Practice test: Study plans | — | — | — | X |

semester. As with the fall 2015 semester, students’ actual score was compared to their predicted score, calculating “prediction scores” by subtracting their actual score from their predicted score. The average prediction scores were then compared between the intro and the final quiz to see if there was a general improvement in prediction ability across the semester. Next, students were retroactively grouped into “ability” quartiles, this time based on their performance on the final exam. Finally, once divided into quartiles, students’ prediction scores were averaged and compared between the intro and the final quiz in the course. This comparison allowed for a determination of the effect, over time, of the score prediction on prediction accuracy for the same subset of students.

In particular, this allowed for the determination of potential changes in the Dunning–Kruger effect among low-achieving students across the semester. Note, only students who took the first and last quiz of the semester and the final exam were included in the comparison.

Additionally, we were interested in how assessment scores compared between the control and metacognitive sections across the semester. First, the subset of students in both sections who had taken each of five major assessments were determined: the intro quiz, all three midterms, and the ACS general chemistry final exam. As each distribution in these comparisons was significantly negatively skewed, the nonparametric Mann–Whitney *U* test was applied once more to compare the median values of the control and metacognitive sections for each assessment. This result was used to determine if there was a difference in assessment performance between the control and metacognitive sections. It was hypothesized that if the metacognitive section performed better than the control, this difference could be attributed to improved metacognitive control as a result of regular metacognitive training.

Finally, as with the fall 2015 semester, the metacognitive and control professors taught course sections at different times of day, with different teaching styles. As such, it was necessary to perform a further statistical analysis to separate out any differences due to teaching styles and demographics between the two sections. As before, an interaction regression was calculated by comparing the final exam averages for each experimental section to final exam averages of the same professors during the control semester. The resulting interaction separated the effect of the metacognitive training on final exam averages from the general effect of teaching styles and demographics between the sections. The interaction regression was also calculated for each final exam quartile to isolate the effect of the metacognitive training for each student performance level.

## 5.5 Results

### 5.5.1 Results: The Effect of Metacognitive Training on Score Prediction Ability

In the metacognitive section, students predicted their scores before taking each weekly quiz and practice test. The prediction scores were calculated with the same method as those in the fall 2015 semester and represents students' score prediction accuracy before taking the assessment: zero is a perfect prediction and a positive value represents an over-prediction of ability. These prediction scores were calculated for the intro quiz, which was taken during the first week of the course, and for the last quiz that students took in the semester. For each quiz, the prediction scores were averaged for the class overall, and the average prediction scores were compared for the two quizzes to determine the students' change in average prediction ability over time. It should be noted that each of these comparisons only involved students who took the intro quiz, the final quiz, and the final exam. Overall results of the intro quiz, which was given at the end of the first week of the course, demonstrate that students in the course tended to overestimate their score by 11%, on average. In comparison, on the final quiz students tended to underestimate their score by 4% on average.

Next, students were retroactively grouped into performance quartiles, this time based on their score on the ACS general chemistry exam, and each quartile's prediction scores were averaged. The result of having students placed into performance quartiles retroactively by final exam score allowed for a comparison in prediction ability of the same subset of students across the semester. When divided into final exam performance quartile, the results of prediction scores for the intro quiz almost perfectly mimicked the Dunning Kruger effect: students performing in the bottom 25% of the course on the final exam dramatically overestimated their intro quiz score by 22%, while students who scored in the top 25% of the course on the final exam almost perfectly predicted their intro quiz score.[4] In comparison, during the final quiz of the semester, students in all ability quartiles, on average, underestimated their score (see Figure 5.1). Noting that a prediction score of zero represents a perfect prediction, the results also demonstrate that the students scoring in the bottom 25% and top 25% on the final exam, on average, had nearly perfect score predictions, with the middle two quartiles slightly under-predicting their final quiz score.

The purpose of having students predict their score was to improve student metacognitive awareness and overcome the Dunning–Kruger effect. In particular, we

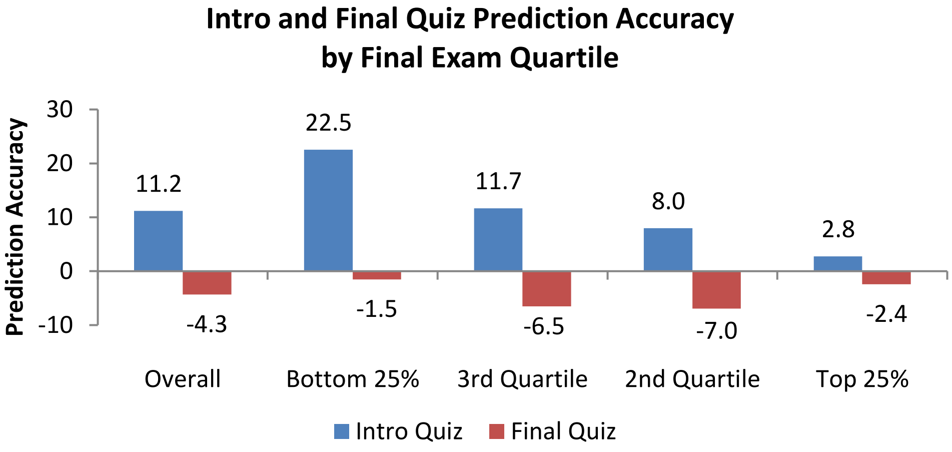


Figure 5.1 Score prediction comparison between the intro and final quizzes (*N* = 217). Note the strong initial overestimation of ability by the bottom quartile and the accurate prediction ability at the end of the semester.

hoped to decrease the tendency toward over-prediction among poor-performing students. Results of the first quiz of the semester indicated, on average, a general overestimation of score, with the most dramatic over-prediction occurring among students who performed the worst on the final exam. This result reflects the general trend in poor prediction ability noted by Dunning and Kruger. Additionally, this result indicates the interesting finding that students who ultimately do poorly on the final exam begin the course by over-predicting assessment scores within the first week of the semester. In comparison, results of the final quiz demonstrate that the poor students' initial tendency to over-predict their score was completely overcome. In other words, the Dunning–Kruger effect of over-prediction among poor students was once more eliminated through the process of score prediction at regular intervals throughout the course. This result reflects in novel result during the fall 2015 semester, where score prediction across multiple practice tests also eliminated the Dunning-Kruger effect.

### 5.5.2 Results: Effect of Metacognitive Training on Assessment Scores

To test the cumulative effect of all metacognitive training, including score prediction and the creation of study plans, section scores on key assessments in the semester—including the intro quiz, each midterm, and the ACS general chemistry exam—were compared between the control and metacognitive sections. All assessments were identical for both sections. Additionally, this analysis only compared the subset of students who took every assessment during the semester.

We were interested in whether training in metacognition, through score prediction and regular study plans, would improve the metacognitive section's assessment scores over time as compared to the control section. For clarity of representation over time, *Z* scores were calculated to compare the difference in assessment scores over time between the course sections. This was done by first subtracting the overall average of both sections of students from the individual section assessment average. After, this difference was divided by the overall standard deviation of both sections on the assessment. This result, then, represents the number of standard deviations each section's average score was away from the overall mean.

Results of these five assessments first indicate the important result that the metacognitive and control sections scored nearly identically on the intro quiz. However, results on all of the subsequent midterm exams and final exam demonstrated that the metacognitive section consistently outperformed the control section on course exams the remainder of the semester (see Figure 5.2). Results of the Mann-Whitney *U* test indicate that while the sections scored identically on the intro quiz, the metacognitive section

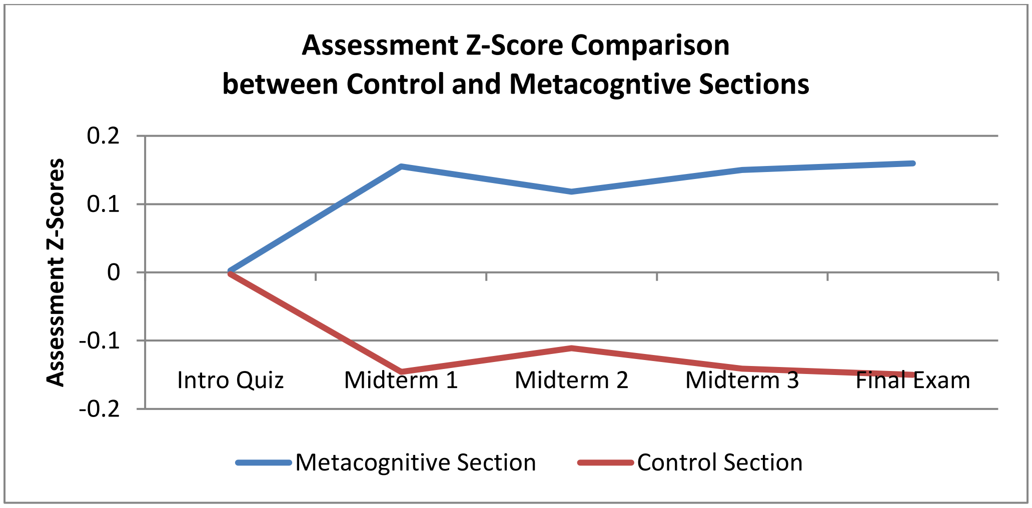


Figure 5.2 *Z*-score comparison of the metacognitive and control sections on each major assessment (*N* = 217). Note that the sections scored similarly on the first quiz but that the metacognitive section outperformed the control on each subsequent assessment.

performed significantly better on each subsequent exam (see Table 5.2). Results of effect sizes, according to the equation given in Chapter 4, indicate that each exam difference was a small statistical effect.

When comparing histogram results of the 2009 form of the ACS General Chemistry I final exam, results indicate a large decrease among students in the metacognitive section, compared to the control section, who scored below 50% on the final. Additionally, there was a corresponding increase among students in the metacognitive section who scored above 70% on the ACS general chemistry exam when compared to the control section (see Figure 5.3).

This first demonstrates the important result that the control and metacognitive sections performed identically on the intro quiz, the first assessment of the semester. However, midterm and final exam score results demonstrate that the metacognitive section consistently and significantly outperformed the control course on each subsequent midterm and on the ACS general chemistry final. As the primary difference between the control and metacognitive sections was the incorporation of metacognitive training, including score prediction and the creation of study plans, we concluded that the

Table 5.2 Mann-Whitney *U* and effect size assessment comparison by course section

| Assessment | *U* Statistic | *p* Value[a] | *r* Effect Size |
|---|---|---|---|
| Intro Quiz | 36735 | 0.962 | 0.00 |
| Midterm 1 | 30937 | 0.001 | 0.14 |
| Midterm 2 | 32648 | 0.022 | 0.10 |
| Midterm 3 | 31206 | 0.002 | 0.13 |
| ACS Final | 30973 | 0.001 | 0.14 |

[a] *N* = 543.

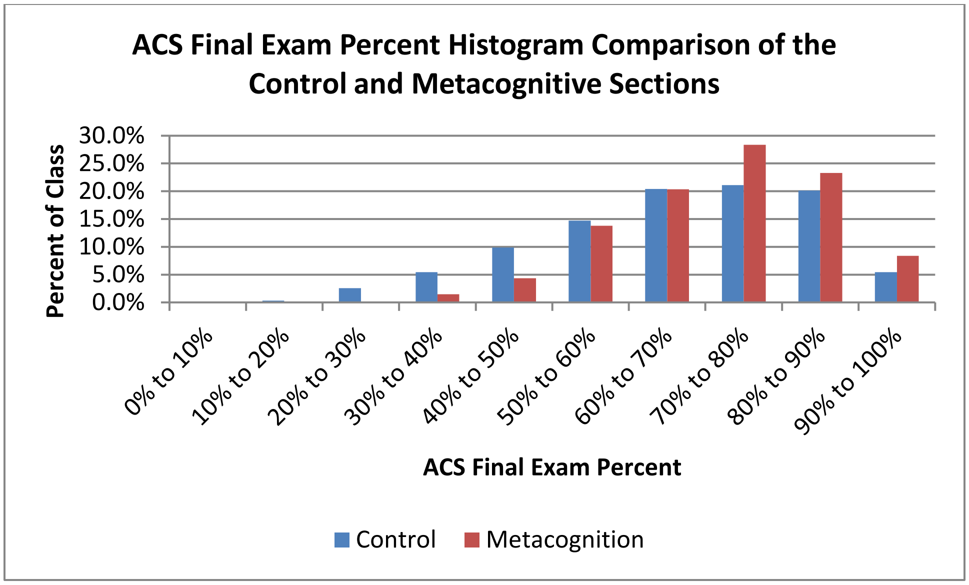


Figure 5.3 Histogram of final exam percentages between the control and metacognitive sections (*N* = 588)

improvement in the metacognitive section's performance, relative to the control section, was likely a direct result of the metacognitive training.

### 5.5.3 Results: Accounting for Other Factors: Incoming Ability and Attrition Rate

To isolate the effect due to metacognitive training, a number of other comparisons were made between the two classes. Already noted, student average performance on the first quiz was not statistically different ($p = 0.962$) between the two sections. Student performance on this quiz reasonably correlated to their performance on the final exam ($r^2 = 0.27$, $p < 0.001$), indicating that the first quiz was a decent measure of student incoming ability. These combined results indicated that student incoming ability was not significantly different between the two sections and did not influence subsequent midterm results.

Additionally, we were interested in whether the regular training in the metacognitive section caused poor students to become keenly aware of their abilities, potentially causing a greater attrition rate in course enrollment among students in the metacognitive section than in the control section. To measure the enrollment attrition rate in each section, the number of students originally enrolled in the online homework was compared to the number of students who took the final exam. In the control course, of 390 students originally enrolled in the homework, 313 took the final exam, a 19.3% attrition rate. In comparison, 22.5% of the metacognitive section dropped between initial enrollment ($n$ = 355) and final exam participation ($n$ = 275). These students dropped the course, withdrew, or remained enrolled in the class while no longer participating in course material. To compare these values, initial and final course enrollment values were converted to percentages of initial enrollment. Analysis using a chi-squared test indicated that the percentage change in the metacognitive section enrollment was not significantly different than the change in the control section enrollment ($p$ = 0.548).

### 5.5.4 Results: Accounting for Other Factors: Interaction Regression

As with the fall 2015 semester, different teachers taught the two sections at different times of day, with different teaching styles, and different in student demographics. Additionally, the professor of the control section generally has a greater tendency to lecture on occasion when compared to the metacognitive professor. Though results of a control semester indicated that the students in each section performed statistically identically on the 2009 form of the ACS General Chemistry I final exam ($p$ = 0.42), students taking the course from the professor who taught the metacognitive section performed slightly better than the control teacher's students on the final exam. As such, researchers were interested in factoring out any effect due to different teaching styles and demographics between the sections. To accomplish this, an interaction regression was performed once more, comparing the final exam performance between the two sections in the present semester to a previous semester where both teachers previously taught. All details of the interaction regression are identical to those found in Chapter 4. Recall that from a practical standpoint, the interaction effect is the *change* in the *average difference* of one effect when measured across a second effect.[22] As such, if the two professors had a 1.0% difference on the final exam average score during the control semester and a 6.0% average difference during the experimental semester, the interaction effect would be the change in average difference over time: 5.0%.

As with the previous semesters, the regression evaluated two main effects: a semester effect and a section effect. The semester effect represents the general difference in student performance between the control and experimental semesters, and the section effect represents the general difference between the sections due to difference in teaching style and section demographics between the control and metacognitive sections. In addition, as part of the interaction regression, the interaction between the semester and section effects represents the additional effect of the metacognitive training after factoring out the other two effects. The regression model is the same as the model given in (2) found in Chapter 4. Additionally, the variables for each were "dummy coded" identically to the regression performed in the fall 2015 semester: "–0.5" for the control semester and control section and "+0.5" for the experimental semester and metacognitive section. An interaction effect was created by multiplying the corresponding semester and the section dummy-coded values for a given set of data. Finally, recall that if a significant interaction is observed between the semester and section effects, this interaction represents the effect of whatever was changed *between* the two professors during the experimental semester that was not present during the control semester. Though some teaching style and demographic differences were present during the control semester, these changes were perpetuated in the experimental semester. As such, the primary change between the sections during the experimental semester was the addition of the metacognitive training. Consequently, if a significant interaction is obtained, this value is interpreted as the effect of the addition of the metacognitive training during the experimental semester.

Results of the interaction regression indicated a significant interaction effect between the section and semester effects for the overall exam averages: approximately 4% of improvement in the overall exam averages can be accounted for by the interaction between the section and semester effects. Additionally, quartile results demonstrate an increasing interaction effect for the bottom three quartiles: the unstandardized slope value of the interaction between the section and semester effects increase substantially as course performance decreases. Most importantly, the interaction term of the bottom quartile demonstrates that approximately 10% of the improvement in overall exam average can be accounted for by the interaction effect. These results are interpreted to indicate that, after factoring out the effect of the semester or the section, metacognitive training accounts for about a 4% improvement in ACS final exam average in the class overall. Regarding course quartiles, the bottom three quartiles of the course significantly improved due to metacognitive training, with the bottom 25% of the course increasing, on average, by 10% on the ACS final exam average (see Table 5.3).

As noted previously, the interaction result is interpreted as the *change in the difference* between professors over time. For example, during the control semester, the control professor and metacognitive professor averaged 66.8 and 68.3%, respectively, on the ACS final exam: a difference of 1.5%. In comparison, during the experimental semester, the control professor averaged 67.1%, with the metacognitive professor averaging 72.9% on the final exam: a difference of 5.8%. Comparing the differences in average performance between the control and experimental semester demonstrates that the difference, over time, increased by approximately 4.3%. Within rounding error, this value is identical to the unstandardized *B* value of the interaction portion of the "overall" regression given in Table 5.3. This same process is applied for each quartile to obtain the other unstandardized *B* values found in Appendix C.

To visualize the effect, "interaction plots" were created: average final exam percent was plotted against a change in semester, with separate lines representing the change in final exam average for each section. In addition to the overall averages, the interaction plot included the final exam averages for each quartile. As a qualitative

Table 5.3 ACS final exam percent interaction regression results

| Parameters | Unstandardized *B* Values | | | | |
|---|---|---|---|---|---|
| | Overall[a] | Bottom Quartile | 3rd Quartile | 2nd Quartile | Top Quartile |
| Semester | 2.5[c] | 3.0[c] | 3.5[d] | 2.7[d] | — |
| Section | 3.6[d] | 5.7[d] | 3.6[d] | 3.6[d] | 2.4[c] |
| Interaction (Semester × Section) | 4.2[b] | 10.3[d] | 5.2[d] | 1.4[b] | — |
| Regression $r^2$ | 0.023 | 0.201 | 0.353 | 0.370 | 0.069 |

*N* = 1169, [a]Nonsignificant values excluded; *N* = 588. [b]$p < 0.05$, [c]$p < 0.01$. [d]$p < 0.001$.

measure, an interaction can be observed when the slopes of the two lines differ, with a greater difference in the slopes representing a more significant interaction (see Figure 5.4). Note in the plots the differing slopes of the section and semester effects for the overall class and for the bottom three quartiles, representing significant interactions for these populations. Reviewing the above table, one result that initially seems peculiar is the relatively small $r^2$ value of the overall interaction regression when compared to the $r^2$ values of the individual quartile interaction regressions. A review of the interaction plot, however, demonstrates the cause for this difference in effect size. When analyzing a single quartile, the range of scores for the assessment is relatively limited, resulting in a larger explanatory power of the interaction regression slope values. However, considering the large range of scores for the entire class, the "overall" regression slope values explain a relatively small portion of the total variance observed across all of the data points.

## 5.6 Summary of the Effects of Regular Metacognitive Training

Previous studies in chemistry have demonstrated the Dunning–Kruger effect in the chemistry classroom: the bottom quartile of students tends to dramatically over-predict their assessment ability.[18,19] Additionally, research in chemical education has indicated that the over-prediction among poor students remains stable across the semester.[20] In comparison, results of the fall 2015 and fall 2016 semesters indicate that, with regular metacognitive training, metacognitive awareness can improve, resulting in improvement in assessment score predictions from the beginning to the end of the semester. During the fall 2016 semester, this effect was most pronounced for students performing in the bottom quartile of the ACS final exam in the metacognitive section.

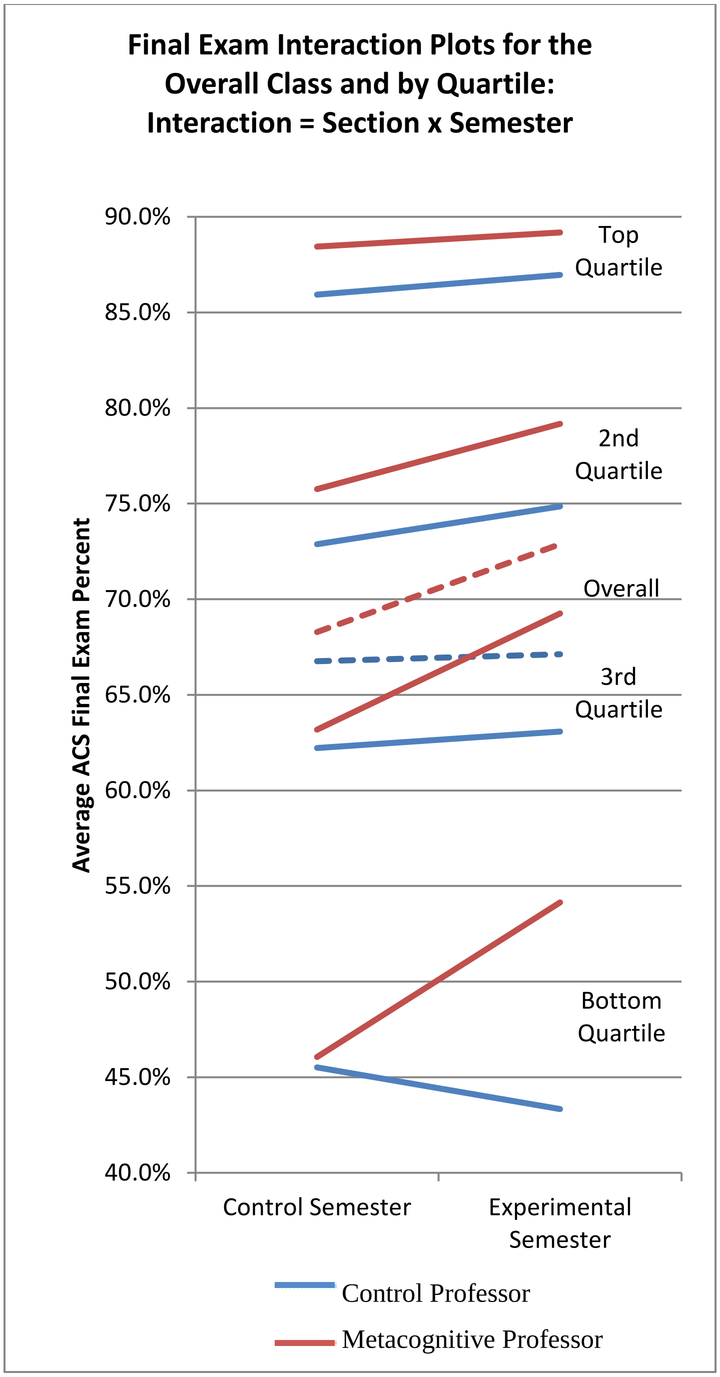


Figure 5.4 Final exam interaction plots for the overall course and by quartile (*N* = 1169)

Note that the overall course is represented by dotted lines, and the quartile results are shown with solid lines. Also note that difference in slopes in the overall data and for the bottom three quartiles, indicating the presence of an interaction between the semester and section effects for these populations of students. Finally note that the largest interaction effect is observed in the bottom quartile of students.

These students, during this study, went from dramatically overestimating their score, on average, on the first semester quiz to almost perfectly predicting their score on the last quiz of the semester. The results of the fall 2015 and fall 2016 semesters indicate the novel finding that, with regular metacognitive training, the Dunning–Kruger effect can be overcome on average for the low-performing students in the course.

Previous chemical education studies have utilized practice tests with the purpose of improving student performance on exams.[25,26] These studies have demonstrated mixed results concerning the effect of these practice tests on exam performance, with the most successful study demonstrating improvement on instructor-generated exams through the use of practice tests.[26] With these previous studies in mind, we know of no previous study within chemical education research that has utilized regular metacognitive training throughout the semester. Additionally, this is the first study known by researchers to demonstrate a course-wide improvement, on average, on the ACS general chemistry final exam as a result of metacognitive training. Most importantly, results of this study indicate an improvement in performance on the ACS general chemistry exam by targeting the bottom three quarters of the class. Previous studies have indicated that the bottom-achieving students are most prone to exam overconfidence, as predicted by the Dunning–Kruger effect.[18-20] This over-prediction of ability may be perpetuating poor student exam performance over time. In comparison, the results of the current study demonstrate that regular metacognition training can overcome this overconfidence in poor-performing students, with a corresponding significant improvement in these students' performance on ACS general chemistry exams.

# CHAPTER 6

# THE CUMULATIVE EFFECT OF ONLINE METACOGNITIVE TRAINING ACROSS A YEAR OF GENERAL CHEMISTRY

## 6.1 Methods and Framework

The study in Chapter 5 was part of a year-long effort to incorporate metacognitive training in the general chemistry classroom. The second semester general chemistry course was taught during the spring 2017 semester, with approximately 525 students enrolled between two major course sections: one section having approximately 350 students and the other with roughly 175 students in the section. As with the first semester study, all homework, quizzes, practice tests, and midterm exams were required and completed within the Madra Learning online homework system.[15] At the end of the spring semester, students took the 2010 form of the second term ACS general chemistry exam. With a full-year of metacognitive training implemented across general chemistry, we were interested in the long-term effect of the incorporation of metacognitive training across two semesters of freshman general chemistry.

The second-semester study compared the performance of the students in two professors' General Chemistry II course sections across the spring 2015 and spring 2017 semesters. As with the study in Chapter 5, student performance in the latter semester was compared to performance to an earlier semester to account for any changes in student performance from one year to the next. During both semesters, the same two professors taught the two sections and will be referenced throughout this section as the “control II” and “metacognitive II” professors.

In addition to a comparison of these two professors, further analysis was performed to separate out any effect on student performance that may have occurred as a result of the professor from whom they took General Chemistry I. As such, student performance was also tracked based on whether they took General Chemistry I from the metacognitive General Chemistry I professor referenced in Chapter 5 or from another General Chemistry I professor. Unlike the analysis in Chapter 5, this analysis will consider students taking chemistry from any other professor besides the metacognitive professor as coming from a “control” General Chemistry I professor. This change is to increase the statistical power of the analysis, as a limited number of students who took General Chemistry I from the control professor from Chapter 5 also took General Chemistry II from the metacognitive second semester professor. As a result, any students who took General Chemistry I from the metacognitive professor of Chapter 5 will be termed as coming from the “metacognitive I” professor, while any other students will be considered as coming from the “control I” professors’ sections. For this analysis, the spring 2017 semester will be named the “experimental semester” and the spring 2015 semester will be termed the “control” semester.

### 6.1.1 Details of the Control and Experimental Semesters

The metacognitive II and control II professors both taught General Chemistry II during the control and the experimental semesters. During the control semester, the course structure in both sections involved the early university attempt at implementing the flipped classroom, with students directed to watch relevant lecture videos and read textbook sections in preparation for class. Students then spent class time working through chemistry questions.[7-12] At the end of each week, students completed an online homework assignment answering several questions regarding the previous week's material.

During the experimental semester, paralleling the changes made in the fall 2016 semester, students completed three homework assignments throughout the week instead of one major assignment at the end of the week. Students in both the control and metacognitive sections completed practice tests before each exam to assist them in their studying. Students in both sections also completed weekly quizzes to assess their ability on the previous week's material. Once more, greater structure was implemented in the flipped classroom model during the experimental semester, with students completing preview video questions to prepare them for the upcoming day's topics and working through study cycle review questions in the homework assignment following a particular lecture.

During the control and experimental semesters, the course layout between the sections, besides the below experimental conditions, was held as constant as possible between the two sections. These structural similarities including a shared syllabus, use of the same flipped classroom slides in class, assignment of the same homework assignments, and assessment using the same weekly quizzes, practice tests, and exams. Additionally, the time of day that the professors taught during the control semester was maintained during the experimental semester. It should be noted that the General

Chemistry II control II professor has had more experience teaching general chemistry than the General Chemistry II metacognitive II professor. However, the metacognitive II professor frequently observed the control II professor's lectures and patterned the lecture, as closely as possible, with the control II professor's process.

An additional note: the professor acting as the metacognitive I professor was different than the professor acting as the metacognitive II professor. Additionally, as already mentioned, though one professor acted as the control II professor during General Chemistry II, students coming from any section other than the metacognitive I section were considered as coming from the control I sections.

### 6.1.2 Details of the Metacognitive Training

Paralleling the efforts of the fall 2016 semester, metacognitive training was implemented in the metacognitive section of General Chemistry II during the experimental semester. This training included the same three specific interventions as the previous semester: assessment score prediction; topic ability feedback; and topic study plans. These interventions were carried out as follows:

1. Students opened the online weekly quiz or practice test, predicting their assessment score and their Likert-scale chemistry ability before viewing any questions.
2. Students took the assessment, afterward "post-dicting" their score and Likert-scale ability on the assessment.
3. Following the assessment, students received detailed feedback on their ability on the assessment score, their prediction accuracy, and their ability by assessment

topic. Once more, the homework system provided students with a list of potential study topics that they could focus future studying upon.

4. On the following homework assignment, students answered a required homework question to assist them in building a study plan, indicating how well they did on the major chemistry concepts covered on the assessment and selecting specific chemistry topics to focus their future studying upon.

## 6.2 Experimental Details

As with the first semester study, to isolate the effects of the metacognitive training, two course sections were analyzed:

1. Metacognitive II Section: this section completed all weekly quizzes and practice tests. For each weekly quiz and each set of practice tests, these students were provided with all of the metacognitive training indicated above including ability and score prediction; prediction accuracy, score, and topic ability feedback; and the creation of study plans.
2. Control II Section: students in this section received the same weekly quizzes and practice tests as the metacognitive II section. Additionally, this section received identical feedback on their quiz score, their ability by assessment topic, and key areas they should focus their studying upon. However, students in the control section did not predict their assessment scores, and they did not create study plans.

When comparing the two course sections, there were two key differences between the sections: score and ability prediction and the creation of study plans. Consequently, the results of this study isolate the effect of these two interventions on student performance and predictive ability.

### 6.2.1 Total Number of Semesters of Metacognitive Training

During the control semester students had no access to metacognitive training for either General Chemistry I or General Chemistry II. However, during the experimental year, students in General Chemistry II had the option of taking General Chemistry I from either the metacognitive I professor or the control I professors. Likewise, in General Chemistry II, from either the control II or the metacognitive II professors. As a result, during the experimental year, students had the possibility of having zero, one, or two full semesters of metacognitive training by the end of second semester freshman chemistry (see Table 6.1).

### 6.2.2 Regression Details

As with the experiment in Chapter 5, we were interested in separating out the effect of metacognitive training experienced in the metacognitive sections from the effect of the professor who was teaching the courses and the semester during which the courses were taught. Additionally, we were interested in determining the cumulative effect on student performance that resulted from a full year of metacognitive training in general chemistry. With these details in mind, a regression was created with four different independent variables, regressed against the unstandardized percent that students received on the ACS General Chemistry II final exam. In this regression, four effects were compared: the “semester” students took general chemistry, the “General Chemistry I

Table 6.1 Number of total possible semesters of metacognitive training

| General Chemistry I | General Chemistry II | | | |
|---|---|---|---|---|
| | Control Year | | Experimental Year | |
| | Control II Professor | Metacognitive II Professor | Control II Professor | Metacognitive II Professor |
| Control I Professor | 0 Metacognitive Semesters | | 0 Metacognitive Semesters | 1 Metacognitive Semester |
| Metacognitive I Professor | | | 1 Metacognitive Semester | 2 Metacognitive Semesters |

professor" that students had during fall semester, the "General Chemistry II professor" that the students had during spring semester, and the total "number of semesters of metacognitive training" that students received. These effects are summarized below:

1. Semester Effect: two semesters of General Chemistry II final exam scores were compared, with final exam scores during the control spring 2015 semester coded as "0" and the scores from the experimental spring 2017 semester coded as "1." This effect measures the change in student final exam performance between the two semesters.
2. General Chemistry I Professor Effect: the General Chemistry II students were coded as "1" if they took the first semester of general chemistry from the metacognitive I professor. Students coming from any other General Chemistry I course were coded as "0" and were considered to have received chemistry instruction from one of the control I general chemistry first semester professors. This effect measures the general change in student final exam performance based on the professor from whom they took General Chemistry I, independent of metacognitive training.

3. General Chemistry II Professor Effect: second semester students were coded as "0" if they took General Chemistry II from the control II professor and "1" if taking second semester chemistry from the metacognitive II professor. This effect measures the change in student final exam performance based on the professor from whom they took General Chemistry II, independent of metacognitive training.
4. Number of Semesters of Metacognitive Training Effect: As summarized previously, during the experimental year, students had the potential of having zero, one, or two full semesters of metacognitive training. During the control year all students had zero semesters of metacognitive training regardless of the professor teaching the course. As such, all students taking General Chemistry II during the control year were coded as "0," while students during the experimental semester were coded as "0," "1," or "2" based on the number of semesters of metacognitive training they received over the full year of general chemistry. This effect measures the change in student performance on the final exam as a result of each semester that they experienced metacognitive training.

In regard to these variables, it should be noted that metacognitive training was only incorporated during the experimental 2016-2017 school year, while the metacognitive I and metacognitive II professors taught during both the control and experimental years of general chemistry. As such, the "professor" effects measure the differences in student performance based on general teaching style and demographics of these courses, while the "number of semesters of metacognitive training" separates out the specific effect of each semester of metacognitive training received in general chemistry on final exam score from the effect of the professor providing this training.

As one final step in the analysis, we were interested in the impact of metacognitive training on final exam score based on students' end of second semester ability. To do so, students were first grouped based on each unique combination of the above four variables. These combinations have been designated letters A through H, with their corresponding total semesters of metacognitive training, in Table 6.2.

## 6.3 Results and Discussion

To measure the effect of the year-long implementation of metacognitive training, a regression was implemented, comparing students' scores on the General Chemistry II final exam to each of the four independent variables: students' semester taking General Chemistry II, their General Chemistry I professor, their General Chemistry II professor, and the total number of semesters of metacognitive training. After accounting for all other factors, this regression isolated the effect of each semester of metacognitive training on students' second semester final exam scores. Results of the overall regression demonstrated that there was no significant difference in General Chemistry II final exam

Table 6.2 Analysis groups and number of semesters of metacognitive training

| General Chemistry I | General Chemistry II | | | |
|---|---|---|---|---|
| | Control Year | | Experimental Year | |
| | Control II Professor | Metacognitive II Professor | Control II Professor | Metacognitive II Professor |
| Control I Professor | A, 0 Semesters | B, 0 Semesters | E, 0 Semesters | F, 1 Semester |
| Metacognitive I Professor | C, 0 Semesters | D, 0 Semesters | G, 1 Semester | H, 2 Semesters |

score based on the number of semesters of metacognitive training that students received (*p* = 0.50). In the regression regarding overall student performance, only the General Chemistry I professor effect was statistically significant: students taking first semester chemistry from the metacognitive I professor did 2.7% better on average than those coming from the control I professors' courses (see Table 6.3).

After the initial regression was performed, students were grouped by their unique combination of the independent variables and were assigned a performance quartile based on their score on the General Chemistry II final exam. This allowed for a measurement of the effect of the total number of semesters of metacognitive training based on students' end-of-semester performance on the General Chemistry II final exam. As can be seen in Table 6.3, the semester effect for the top quartile of students was statistically significant, with the top 25% of students performing on average 2.0% better on the final exam during the experimental semester than the control semester (*p* = 0.019) However, no other effect, including the number of semesters of metacognitive training, were found to be statistically significant.

For the middle two quartiles, three effects were found to be statistically

Table 6.3 Regression results of the cumulative effect of metacognitive training

| | Overall | Bottom 25% | 3rd Quartile | 2nd Quartile | Top 25% |
|---|---|---|---|---|---|
| Semester | --- | --- | $2.7^{d}$ | $3.4^{d}$ | $2.0^{b}$ |
| General Chemistry II Professor | --- | $-2.9^{b}$ | $-1.1^{b}$ | $-1.3^{b}$ | --- |
| General Chemistry I Professor | $2.7^{b}$ | $3.7^{c}$ | $3.5^{d}$ | $3.3^{d}$ | --- |
| Number of Semesters of Metacognitive Training | --- | $\mathbf{2.7}^{b}$ | --- | --- | --- |

*N* = 895. [a]Nonsignificant values excluded; $^{b}p < 0.05$, $^{c}p < 0.01$. $^{d}p < 0.001$.

significant: the semester, General Chemistry I professor, and General Chemistry II effects. In summarizing these effects, students in both of these quartiles did approximately 3% better on the General Chemistry II final exam in the experimental semester than during the control semester. The middle two quartiles also tended to perform better on the final exam, by about 3% on average, if they took general chemistry from the metacognitive I professor than from the control I professors. Finally, students taking General Chemistry II from the metacognitive II professor tended to perform approximately 1% worse on the ACS General Chemistry II final exam than students taking the second semester course from the control II professor. In regard to this last effect, it should be noted that the difference in performance between professors is the result of teaching style and demographics and is not a measure of the metacognitive training performed in the course. Once more, students in the middle two quartiles did not significantly change in final exam performance as a result the number of semesters of metacognitive training received over the year of general chemistry (see Table 6.3).

For the bottom quartile of students, three effects were found to be statistically significant. In summary, the bottom quartile of students tended to do nearly 4% better if they took first semester chemistry from the metacognitive I professor, when compared to students taking General Chemistry I from any of the control I professors ($p = 0.003$). Additionally, these students performed almost 3% worse on the final exam if they took second semester chemistry from the metacognitive II professor's course ($p = 0.012$). Once more, these two effects represent the result of teaching style and demographics on student performance and does not represent any effect of the metacognitive training itself. Finally, students in the bottom quartile based on final exam performance were found to improve by 2.7% on the ACS General Chemistry II exam for each semester of metacognitive training received across the year of general chemistry ($p = 0.045$, see Table 6.3). In other words, the metacognitive training implemented across the year of general chemistry instruction specifically targeted the bottom 25% of students of general chemistry. As these students could potentially receive two full semesters of metacognitive training across the year of general chemistry, the bottom quartile of students could experience a 5.4% increase in final exam average if they received two full semesters of metacognitive training.

This effect can be observed by comparing the final exam averages for the bottom quartile of students from one semester to the next. To observe the effect, we compared the bottom quartile's final exam score for students taking general chemistry only from the metacognitive I and metacognitive II professors. Results indicate that these students averaged 47.4% on the ACS final exam during the control year and 53.8% on the final exam during the experimental year. During this latter year, these students received two full semesters of metacognitive training and were provided with a homework system including greater structure to the flipped classroom. These changes combined to yield a 6.4% increase in bottom quartile final exam average from one year to the next.

Results of the regression first indicate that all students, except those in the bottom quartile, performed better on the final exam during the experimental than during the control semester. During General Chemistry I and II in the experimental year, greater structure was implemented for the flipped classroom, including embedding videos in required homework assignments that were accompanied by tutorial questions. Homework assignments became more frequent, reflecting a preview and review pattern based on the study cycle. Finally, weekly quizzes and practice tests were incorporated across both semesters as regular measures of student ability. Regression results of the semester effect demonstrate that these general changes significantly improved student performance on the ACS General Chemistry II final exam for students in the top three quartiles. However, these changes had no significant effect on final exam performance among the lowest-performing students in the class.

Regarding the professor effects, students taking the first semester of general chemistry from the metacognitive I professor tended to do better on the General Chemistry II final exam than those coming from the control I professors. This trend was noted for all quartiles except the highest performing students. Regression results also demonstrate that students taking the second semester of general chemistry from the metacognitive II professor tended to do worse than those taking chemistry from the control II professor. This effect was also significant for all but the highest performing students in the course. As noted previously, the metacognitive II professor had less experience teaching general chemistry than the control II professor. Though this professor observed the control II professor's lectures, it is possible that this difference in teaching experience could have negatively impacted student performance, with the lowest quartile of students most significantly impacted.

Finally, the number of semesters of metacognitive training students received represent the unique effect of the year of metacognitive training on students' second semester final exam performance. This effect, then, represents the cumulative effect of student score prediction, ability feedback, and the creation of study plans on final exam performance over as many as two total semesters of general chemistry. It should be noted that this effect was only statistically significant for students in the bottom quartile on the final exam, demonstrating that only the lowest performing General Chemistry II students statistically improved in final exam performance due to the year-long implementation of metacognitive training.

Results of Chapters 5 and 6 demonstrate that metacognitive training significantly improves student chemistry performance, with the greatest improvement in student performance experienced by the bottom 25% of students. These results also demonstrate that metacognitive training has a cumulative effect: for each semester of metacognitive training received, students in the bottom quartile improve by nearly 3% on the final exam. With two possible semesters of metacognitive training, the result is a 5.4% increase in final exam average for the lowest achieving students in the General Chemistry II course. These results indicate that metacognitive training across the year of general chemistry specifically targets the bottom quartile of students, improving their performance on the final exam.

# CHAPTER 7

# THE EFFECT OF ONLINE METACOGNITIVE TRAINING ON GENERAL CHEMISTRY STUDENT COURSE TRAJECTORY

In Chapters 5 and 6 of this dissertation, metacognitive training was demonstrated to improve student exam performance. In particular, a system of regular assessment, student ability prediction, ability feedback, and the creation of study plans was found to significantly improve student performance on the ACS general chemistry final exams, especially targeting the lowest quartile of students in improving their chemistry ability.

Though the system was demonstrated to significantly improve the bottom quartile of students' ability in both General Chemistry I and II, we were interested in quantifying the individual effects of each aspect of the metacognitive training on students' improvement in the course. We were especially interested in the degree to which student metacognitive awareness, as measured by end-of-semester assessment score prediction ability, and students' efforts toward improving metacognitive control, as measured by their completion of study plans, each individually affected students' performance across the semester. Finally, we were interested in exploring the degree to which this metacognitive training influenced student trajectory across the course of the semester.

It was hypothesized that students who had high metacognitive awareness by the end of the semester would be able to accurately predict their score on an end-of-semester assessment. Additionally, as students regularly completed study plans, we hypothesized that they would improve in their metacognitive control across the semester. As students improved in their metacognitive ability and awareness, it was hypothesized that they would then improve in their overall metacognitive ability, resulting in an improved course trajectory over the semester. In comparison, students with poor metacognitive ability would regress over the course of the semester.

With these in mind, this chapter will address three research questions: 1) how can student performance trajectory across the course be measured, 2) to what degree does metacognitive awareness, as measured by end-of-semester prediction ability, influence student course trajectory, and 3) to what degree does student efforts toward improving their metacognitive control, as measured by the number of study plans created across the semester, influence student course trajectory.

## 7.1 Methods and Framework

All of the data analyses were performed on the data collected in the General Chemistry I, fall 2016 semester. Additionally, these data only represent the students taking first semester chemistry from the metacognitive I professor. To determine course trajectory, change in student performance was measured on five major assessments during the semester: the introductory quiz; midterms 1, 2, and 3; and the 2009 form of the ACS General Chemistry I final exam. The intro quiz, which was simply an assessment of topics learned in the first week of the class, was selected because it represented the first opportunity for students to predict their scores and measure their chemistry ability on an assessment. Additionally, student percentile on this quiz was found to be highly correlated to student percentile on the final exam ($R^2 = 0.33$, see Appendix D for a list of questions from this quiz).

Because percent averages on the assessments varied widely, students were assigned a percentile based on their assessment performance relative to other students in the course. Afterward, student's change in percentile across these five assessments were assessed by generating a "percentile change slope." This slope conceptually represents students' average change in course percentile from one major assessment to the next across the semester. As such, it represents student trajectory in their general chemistry ability compared to their peers, with positive slopes representing students who improved in their chemistry performance over the semester, in comparison to the other students in the course, with each assessment.

To measure the effect of students' end-of-semester metacognitive awareness on course trajectory, we first analyzed students' ability to predict their scores at the end of the semester. Recall that students predicted their scores before taking each quiz or practice test, with "prediction accuracy" scores calculated as the difference between their predicted and their actual assessment score. For these values, a score of zero represented a perfect prediction and a positive value represented score over prediction. As mentioned in Chapter 5, students' average prediction ability on the first quiz was very poor, but on the final quiz in the semester prediction accuracy averages was nearly perfect for each quartile. Upon further review, the variation determined in students' score prediction for the first and last quizzes were both very wide: the first quiz had a prediction score standard deviation of 19.3%, with the final quiz of the semester having a prediction score standard deviation (SD) of 16.0%.

We hypothesized that students whose final quiz predictions were near their actual quiz score had become highly metacognitively aware of their chemistry ability by the end of the semester. In comparison, those students who could not accurately predict their score on this final quiz did not achieve high metacognitive awareness by semester’s end. Within this framework, it was hypothesized that students that had achieved high metacognitive awareness by semester’s end would be more likely to improve their course trajectory across the semester. As such, we first compared students’ final quiz prediction accuracy to their course trajectory slope to determine if a correlation existed between these factors.

In addition, we were interested in comparing the effect student’s level of metacognitive control to their course trajectory. Though metacognitive control could not be directly measured, it was hypothesized that regular completion of study plans across the semester would facilitate student improvement in metacognitive control. Within this framework, we determined the number of study plans that students completed during the semester and compared this value to their course trajectory slope. It was hypothesized that regular completion of study plans would represent an active effort to improve metacognitive control and would also correlate to an improvement in student percentile over the course of the semester.

Additionally, we reviewed the assignment types that each study plan was embedded within, since some of these study plans were incorporated into students’ weekly homework assignments and others were assigned as an isolated homework assignment with no chemistry content questions. Afterward, these study plans were divided into two groups based on the type of assignment the study plans were integrated within. This analysis enabled the determination of the effect of each study plan delivery method on a student's course trajectory.

Finally, two regressions were developed as part of this analysis. The first regressed final quiz prediction ability and the number of study plans that students completed against their course percentile trajectory slope. This regression represented the degree to which each variable uniquely contributed to students' trajectory over the semester. Second, a regression was created that compared student course trajectory to the total number of homework assignments they completed in the semester, their final quiz score prediction ability, and the number of study plans they created. This allowed us to separate out the effect of student effort over the semester in course trajectory that was independent of metacognitive training.

## 7.2 Experimental Details

### 7.2.1 Course Trajectory

We were first interested in determining a way to quantify student course trajectory across the semester, with a goal of numerically evaluating if a student's chemistry ability improved, declined, or remained stable across the semester. To measure students' course trajectory across the semester, five assessments were selected: the first quiz, each of the three midterms, and the final exam. Note that only students who completed each of the five assessments and the final quiz were considered in the analysis.

For each of these assessments, students were assigned a percentile rank based on their score on each assessment, which compared their relative score to the scores of the other students in the class. It should be noted that individuals receiving identical scores on a given assessment were given the same intermediate percentile rank. To interpret these percentile values, a percentile of nearly zero represents a student scoring at the bottom of the class, the 100th percentile represents a score at the top of the class, and the 50th percentile representing an average assessment score. Afterwards, a slope value was generated for each student with the independent variable of the slope representing the five major assessments and the dependent variable representing student percentile on each assessment. This slope, then, quantified how students changed in chemistry performance across the semester in comparison to their peers. In particular, this slope represents a particular student's average change in percentile rank from one assessment to the next, with positive slopes representing students that improve during the course of the semester.

To visualize this effect, Table 7.1 provides the percentile rank of three different students on each of the five course assessments. After, the table provides the course trajectory slope that was calculated for these students. In interpreting these data, student 1 can be seen to start just below the class average on the first quiz and midterm and falling to scoring near the bottom of the class by the final exam, resulting in a strongly negative slope value. Student 2 began a little better than the class average and improved somewhat in percentile rank near the end of the semester. As this student remained generally stable in percentile, the slope calculated for this individual was slightly positive but close to zero. Finally, student 3 demonstrated very low performance on the first quiz but steadily progressed in performance to scoring above average on the last two assessments of the semester, resulting in a highly positive slope value. These slope values can be interpreted

Table 7.1 Student percentile ranks over time and corresponding course trajectory slopes

| | First Quiz | Midterm 1 | Midterm 2 | Midterm 3 | Final Exam | Slope |
|---|---|---|---|---|---|---|
| Student 1 | 44.0 | 27.6 | 27.6 | 21.2 | 2.3 | -9.0 |
| Student 2 | 63.4 | 62.7 | 63.4 | 70.7 | 67.3 | 1.6 |
| Student 3 | 24.9 | 43.3 | 50.7 | 65.0 | 67.3 | 10.6 |

as follows: student 1 declined, on average, by 9% of the total number of students in the course from one assessment to the next; student 3 inclined, on average, by 10.6% of students after each assessment, and student 2 only changed by 1.6% of the course from one assessment to another.

### 7.2.2 Study Plans

Students completed several study plans throughout the semester, responding to questions regarding their performance on weekly quizzes and practice tests and what topics they would focus their studying upon. Across the semester, they were given 18 total opportunities to generate study plans within the homework system. Eight of these study plans were embedded in the general homework assignments that students completed as part of the general homework system. In comparison, ten study plans questions were given to students in separate study plan assignments with no additional homework questions incorporated in the assignment. We were interested first in how frequently students utilized each of these study plan assignments and to what degree completion of each study plan type correlated to a change in a student’s percentile across the semester. It was hypothesized that greater study plan utilization would result in high metacognitive control by the end of the semester, correlating to an improvement in course trajectory during the semester.

Overall results indicate that students' utilization of these study plans was high, with students completing on average 15.4 of the 18 total study assignments. However, upon further review, it was found that completion of the study plans embedded in daily homework was very high, with students completing on average 7.7 of the 8 plans (*SD* = 0.67). In comparison, students completed an average of only 7.6 out of 10 of the study plans given as separate homework assignments (*SD* = 2.0), with students completing anywhere from 2 to all 10 of these study plans. With this large difference in completion rates, regressions were later developed to each represent the degree to which completion of study plans within each of these categories correlated to student course trajectory.

### 7.2.3 Score Prediction Ability

One goal of the semester was to improve student's metacognitive awareness in order to overcome the Dunning-Kruger effect among poor performing students. In Chapter 5, it was demonstrated that by the final quiz in the semester, students in all four quartiles accurately predicted their assessment score. Though the average prediction ability was improved for all quartiles, the standard deviation of these predictions remained very wide, indicating that some students were better at predicting their final quiz scores than others. It was hypothesized that students with very accurate score predictions on the final quiz had achieved high metacognitive ability by the end of the semester, while students who did not predict their score accurately did not achieve high metacognitive ability by semester's end. As such, we were interested in the degree to which students' metacognitive ability at the end of the semester, as measured by their final quiz prediction ability, correlated to their course trajectory. After, a combined regression was generated, comparing students' percentile slope to their end-of-semester prediction ability and the number of study plans they completed over the semester.

### 7.2.4 Accounting for Other Factors

Finally, as a control, course trajectory slope was regressed against the total number of homework assignments that were completed during the semester. Additionally, students' number of completed homework assignments, their final quiz prediction accuracy, and their number of study plans were each regressed against course trajectory. These final comparisons controlled for changes in student course trajectory that were the result of regular completion of assignments, enabling an isolation of unique the effect of regular metacognitive training.

## 7.3 Results and Discussion

### 7.3.1 Course Trajectory

To calculate students' trajectory in the course, they were first designated a percentile rank for each of five major assessments across the semester. Afterwards, a slope was generated for each student, representing the average change in students' course percentile over these five major assessments. These slope values, then, represent the average class percentile change for each student from one major assessment to another in the semester. The most negative slope value was approximately -11, indicating that the student declined in course percentile by an average of 11% of students from one major assessment to the next over the course of the semester. In comparison, the most positive slope value was approximately 16, indicating that this student inclined in course percentile by 16% after each major course assessment.

Though some students had very positive or very negative trajectories over the semester, the majority of students generally remained in the nearly same percentile rank across the semester. In fact, 67% of all students were found to have slope values between -5 and 5, indicating that the majority of students change in their class rank on average by less than 5% of the total class from one assessment to another (Figure 7.1). These results demonstrate that most of the students taking general chemistry at the University of Utah do not substantially change in their chemistry ability relative to other students in the class across the semester. In other words, students that do well at the beginning of the semester tend to continue to do well to the end of the semester, while students that struggle at the beginning of the semester will struggle through to the final exam.

Results of Chapters 5 and 6 demonstrated that metacognitive training can improve student performance in the course, so it was hypothesized that metacognitive training

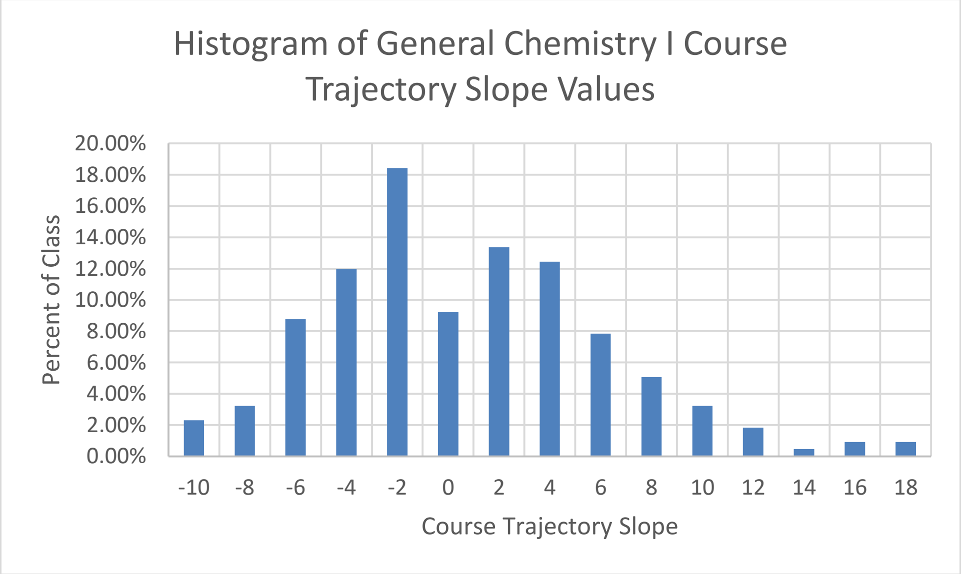


Figure 7.1 Histogram of trajectory slope values for General Chemistry I (*N* = 217)

could positively impact student trajectory over the semester. In particular, we were interested in the degree to which student trajectory in the course was influenced by their utilization of the metacognitive training across the semester.

### 7.3.2 Semester Trajectory Correlations

First, students' trajectory slope values were regressed against the number of study plans that they completed and their prediction ability on the final quiz of the semester. In regard to study plans, it was found that the number of study plans that students completed significantly correlated to student trajectory slope values ($p = 0.007$). However, when study plans were further grouped into categories based on the type of assignment they were assigned within—either embedded in regular homework assignments or given to students as a separate study plan assignment—only the completion of plans within the study plan assignments significantly correlated to student course trajectory ($p = 0.237$ and $p = 0.007$ respectively). These results indicate that the completion of study plans within separate assignments significantly correlated to student course trajectory.

Next, students' prediction scores on the final quiz were regressed against students' course trajectory slope, with results indicating a generally negative correlation between score prediction and course trajectory. This result indicates that students who under-predict their score on the final quiz tended to improve over the course of the semester more than those who overpredicted their final quiz score. However, this result was just outside of statistical significance ($p = 0.065$).

Upon further review, it was found that students who accurately predicted their final quiz score tended to have a more positive course trajectory when compared to students who were more extreme in their score predictions. In other words, any score prediction, either positive or negative, that was far from students' actual score was correlated to poorer student trajectory. With this in mind, the absolute value of each students' prediction score for the final quiz was generated, with a value of 0 representing perfect prediction and any value greater than zero representing how far a student's prediction (positive or negative) was from their actual score. This trend significantly correlated with course trajectory, with students who accurately predicted their final quiz score generally having a more positive course trajectory than students whose score predictions were far from their actual score ($p$ = 0.032). It should also be noted that students' predictions ranged from perfect score predictions to as much as 50% away from their actual score. In general, however, students tended to predict their final quiz score fairly closely to their actual score, with students' prediction score averaging 13.1% away from their actual final quiz scores ($SD$ = 10.4).

### 7.3.3 Regression Results

The results of the section 7.3.2 indicate that students' trajectory in the course, as measured by their percentile trajectory slope, significantly correlated to the number of study assignments that they completed and their final quiz prediction ability. With these in mind, a regression was developed to determine the individual effects of students' completion of study assignments and their final quiz prediction accuracy against students' semester percentile trajectory. To do so, the following were incorporated in the regression:

1. Dependent Variable: Student Course Trajectory: the average change in student

percentile from one major assessment to the next, calculated as their percentile trajectory slope.

2. Independent Variable 1: Number of Study Plans: the number of study plans completed in study plan assignments over the course of the semester, excluding any study plans completed during the regular flipped classroom homework assignments
3. Independent Variable 2: Student Prediction Accuracy: students' final quiz score prediction accuracy calculated as the absolute value of the difference between students' predicted and actual scores on the final quiz of the semester.

Results of the regression indicate that variation in each independent variable significantly correlated to variation in student trajectory slope values. In particular, course percentile slope was calculated to increase by 0.54 percentile for each study plan that students completed within study assignments ($p = 0.005$). Prediction accuracy had the opposite effect: prediction scores close to zero on the final quiz tended to correlate to better course trajectory ($p = 0.027$). In particular, course percentile slope drops by 0.086 slope percentile with each percent that a student's final quiz prediction deviates from their actual score.

In interpreting these values, students had the possibility of completing a total of 10 study plans in the study plan assignments, with their percentile slope increasing by 0.54 for each study plan created. In other words, a student who completed all ten study assignments would have, on average, a course trajectory slope that was 5.4 percentile (10*0.54) higher than a student who did not complete any study plans. In other words, a student completing all ten study assignments would improve, on average, by 5.4 percent more of the total students in the course following each assessment than a student who did not completing any of the study plans.

Similarly, recall that students' final quiz predictions ranged from perfect predictions to as much as 50 percent away from their actual score. Regression results indicate a drop of 0.086 percentile in the course for each percent that a student's prediction was away from their actual final quiz percent. Considering that the worst prediction on the final quiz was 50% away from the student's actual score, this student was predicted to have a course percentile slope that was 4.3 (50*0.086) percentile lower, on average, than the student with a perfect score prediction. As a result, a student with the worst possible prediction declines in course percentile by 4.3% more of the students in the course after each assessment than a student who predicted the final quiz score perfectly.

Finally, there was an apparent cumulative effect of both student completion of study plans and their end-of-semester prediction ability on student trajectory. In particular, as students complete more study plans, with prediction accuracy levels closer to zero, their course trajectory improves.

To visualize the effects of each of these variables on course trajectory, a three-dimensional scatterplot was generated using Matlab (see Figure 7.2).[32] This scatterplot compared students' course trajectory slope to their final quiz prediction and the number of study plans they generated. Additionally, the program was used to generate a "plane of best fit," representing the cumulative effect of these two independent variables on students' course trajectory slope. For ease of interpretation, the plane has been color-coded, with yellow and orange representing positive course trajectory, blue and purple

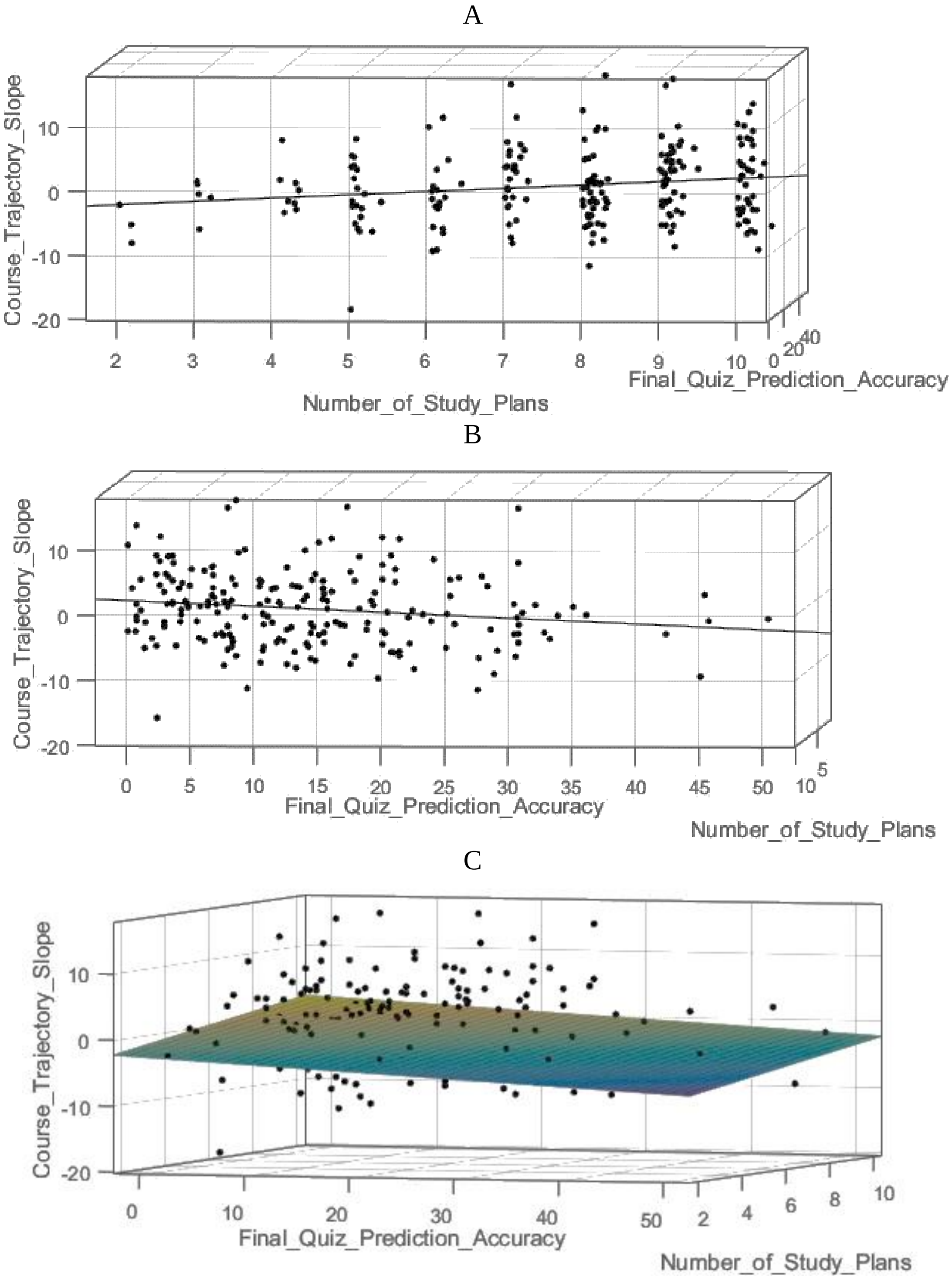


Figure 7.2 Three perspectives of a 3D scatterplot (*N* = 217).
A: This view demonstrates the positive correlation between the completion of study assignment study plans on student course trajectory. B: This view demonstrates the negative correlation between students' prediction on the final quiz and their course trajectory. C: This view demonstrates the cumulative effect of final quiz prediction ability and study plans on course trajectory.

representing negative course trajectory, and green representing course trajectory slopes near zero. Three different angles of this graph have been shown for clarity, with two graphs demonstrating the individual effects of independent variables and the third showing the cumulative effect of both variables.

### 7.3.4 Statistical Mediation

To understand the findings of the next section, 7.3.4, it is important to first understand the concept of “statistical mediation.” In statistics, a variable is said to be mediated by another variable when a statistically significant finding that is actually caused by another factor that was not accounted for in the regression.[30] Consider the results of Chapter 6: though the present text demonstrates a different result, early findings before writing this chapter indicated that metacognitive training improved final exam performance for all students and not just for the bottom quartile. In this initial iteration of the regression, the General Chemistry I professor effect was not included, and results seemed to indicate that for each semester of metacognitive training received, all students’ final exam average score improved by 3%. However, when the General Chemistry I professor effect was added to the regression, the statistical significance of the metacognitive training disappeared for all quartiles except the bottom quartile. In other words, these later results indicated that the initial finding regarding the metacognitive training was *mediated* by the fact that all students who received metacognitive training also took General Chemistry I from the metacognitive I professor. In other words, the initial effect attributed to metacognitive training was actually the result of the effect of the professor from whom students took first semester chemistry.

### 7.3.5 Course Assessment Completion and Student Trajectory

Students’ trajectory in the course, as measured by their percentile trajectory slope, was compared to the number of homework assignments that they completed over the course of the semester. This comparison allowed us to control for the effect of completion of assignments that may have influenced the regression results of the metacognitive training. Results indicated that the total number of homework assignments that students completed significantly correlated to their trajectory during the semester ($p = 0.012$). This indicates the logical result that as students complete more course assessments, they tend to improve in their ability over the semester, while students who complete fewer assessments tend to regress.

After, a final regression was performed, comparing students’ course trajectory slope to three independent variables: their number of study plans completed within study assignments, students’ final quiz score prediction accuracy, and the total number of homework assignments the students completed over the semester. It was hypothesized that the degree to which student completed homework assignments represents how hard students were working in the course. As such, including this effect in the regression allowed us to separate the effect of metacognitive training from the general effect of student effort in the course.

After the addition of the course assessments to the metacognitive training regression, results indicate that student effort, as measured by the number of homework assignments they completed, no longer significantly correlated to course trajectory ($p = 0.165$). In other words, the general effect of student effort across the semester decreased to a nonsignificant value when metacognitive training is included in the regression. This indicates that the effect of metacognitive training, as represented by score prediction ability and the completion of study plans, more significantly influenced student course trajectory than student effort, as represented by their completion of course assignments. In other words, when metacognitive training is incorporated in a course, this training over the semester will influence student trajectory to a greater degree than students' completion of homework assignments.

In this regression, final quiz score prediction ability still significantly correlated to course trajectory, even after accounting for student effort on homework assignments ($p = 0.021$). However, that the effect of completing study plans, after accounting for student effort on homework assignments, was just outside of the range of statistical significance ($p = 0.083$). In other words, after accounting for of the number of homework assignments that students completed during the semester, the completion of study plans no longer predicted student changes to course trajectory within a 95% confidence interval.

In the mediation example given previously, the effect metacognitive training was completely mediated by the effect of students' General Chemistry I professor. The result was that the effect of the General Chemistry I professor remained statistically significant while the effect of metacognitive training was eliminated for all but the bottom quartile of students. In comparison, the results of the present regression indicate a reduction in the significance of both the study plan and number of homework assignments effects to nonsignificant $p$ values. The reduction in the significance of both variables indicates that these variables are partially mediated by each other. In other words, these variables, in part, are measuring the same effect, which may be termed as student course effort. After separating out this effect of student effort in the course, the study plan effect had a $p$

value of 0.08, indicating that this variable may represent a unique effect on student trajectory as a result of increased metacognitive ability, but was just outside of statistical significance. These combined results indicate that students who are regularly completing course assessments will also be more likely to complete the study assignment study plans based on their relative amount of effort in the semester. However, after accounting for student effort, the metacognitive training as a result of completing study plans may also have a unique effect on improving course trajectory, though this effect is just outside of a 95% confidence interval.

Finally, even after considering the effect of student effort throughout the semester, student score prediction ability at the end of the semester is significantly correlated to semester trajectory. These results indicate that high levels of metacognitive awareness, as measured by score prediction ability, can help students to have a positive course trajectory. Additionally, student completion of study plans can assist them in improving their ability across the semester, though this effect seems to be the combined effect of their general effort across the semester and their efforts to improve their metacognitive awareness over the semester. It should be noted that R2 values were also calculated for each effect and Venn Diagrams were generated to represent the individual effects of prediction ability, study plan completion, homework completion, and the general effect of effort. For the sake of brevity, these have been included in Appendix E.

### 7.3.6 Discussion

Results of this analysis indicate that students who completed the majority of the study plans within the study assignments and who accurately predicted their score on the final quiz tended to have positive course trajectories. In comparison, those students who completed relatively few study assignment plans and who predicted their final quiz score poorly tended to have negative course trajectories. In other words, as students complete more study plans and become proficient in assessment score prediction, they tend to improve in their chemistry performance over the course of the semester, inclining in their course percentile when compared to students who did not complete as many study plans or who could not predict assessment scores accurately at the end of the semester.

The completion of study plans was hypothesized to assist students in improving their metacognitive control, though regression results also indicate that the observed study plan effect is also mediated by students' effort across the semester. In comparison, accurate score prediction was hypothesized to represent high end-of-semester metacognitive awareness. As such, these results indicate that as students improve in their metacognitive ability, combined with regular efforts across the semester, their course trajectory correspondingly improves.

The results of Chapters 5 and 6 demonstrate that the training performed in the metacognitive section of the first and second semester of General Chemistry Improved student performance on assessments when compared to the control section. These results help to clarify how this process occurs: as students work to complete the study plans and learn to accurately predict their assessment scores, their metacognitive ability improves. They are then able to influence their course trajectory and improve in their course assessment percentile as the semester progresses. In comparison, students whose utilization of the study plans was poor and who did not learn to accurately predict their assessment scores by the end of the semester generally regressed in course percentile as a result of being overtaken by other students who utilize the metacognitive training. In other words, improvement in course trajectory tends to occur among students who are working hard and who improving their metacognitive ability through regularly completing study plans and learning to predict their assessment scores accurately.

These results are important for students who may begin the general chemistry course at a lower ability than desired, as well as for teachers who would like to know how to help these students improve in their ability over the semester. First, results of the intro quiz indicate that a student’s incoming chemistry ability can be assessed, with some degree of success, within the first week of the semester. Second, these results indicate that most students tend to remain stable in their chemistry performance across the semester. Third, these results indicate that these students’ trajectory in the course can be improved through regular metacognitive training. The system in this study involved the implementation of regular assessments, the requirement of score and ability prediction, feedback of prediction accuracy and ability by topic, and the development of regular study plans. Results of this project indicate that students who regularly utilize this system, especially regarding frequent creation of study plans and accurate end-of-semester prediction ability, can improve student trajectory in the course and help students to be more successful in general chemistry.

# CHAPTER 8

# THE LONG-TERM EFFECTS OF THIS RESEARCH EFFORT

The purpose of this research project was to improve student performance and pass rates in general chemistry without lowering academic standards. With this goal in mind, several changes were made in the general chemistry program over the past few years at the University of Utah, as summarized in Table 8.1 Note that changes in parentheses describe actions that influenced student enrollment or performance but were not directly part of this research effort.

To track changes in general chemistry student performance, we analyzed changes in four student measures: course enrollment, number of students passing the course, course percent pass rates, and final exam average. Data were obtained for General Chemistry I and General Chemistry II for the school years from 2013 to the present. Additionally, organic chemistry I fall semester pass rates were obtained to determine the long-term effects of the changes made in the general chemistry curriculum. Note that long-term graphs of pass rates including previous terms are included in Appendix F.

Table 8.1 Timeline of research changes

| Semester | Summary | Details |
|---|---|---|
| Fall 2013 | Standardization of General Chemistry | Control Year: (All sections of general chemistry standardized with a common, flipped classroom curriculum) |
| Fall 2014, Spring 2015 | Prerequisites and discussion attendance | General Chemistry I math and chemistry prerequisites established, discussion attendance required |
| Fall 2015 | General Chemistry I practice tests and metacognitive training | Implementation of practice test system, including ability feedback. Metacognitive Section: addition of score and ability prediction and study plan creation. |
| Spring 2016 | General Chemistry II practice tests and metacognitive training, increased flipped classroom structure | Implementation of practice tests, with feedback of ability by topic.<br>Metacognitive Section: addition of score and ability prediction and study plan creation.<br>Flipped Classroom: greater structure within the homework system during General Chemistry II |
| Fall 2016, Spring 2017 | Increased metacognitive training including weekly quizzes and practice tests, flipped classroom structure, and the study cycle | Implementation of practice tests and weekly quizzes, with feedback of ability by topic. Metacognitive Section: addition of score and ability prediction and study plan creation. Flipped Classroom: greater structure within the homework system during General Chemistry I according to the study cycle. (Development of online sections of general chemistry) |
| Fall 2017 | Study assignments | Addition of targeted study assignments to the flipped classroom and regular metacognitive training<br>(University-wide increase in student enrollment) |

## 8.1 Changes to General Chemistry I

### 8.1.1 Changes to General Chemistry I Enrollment

During the fall 2013 semester, a baseline of 997 students enrolled in General Chemistry I. Before the fall 2014 semester, a series of course prerequisites were implemented, one of which students must accomplish to enroll in general chemistry. As a result, enrollment declined over the next two years, reaching 872 students in the fall 2015 semester of general chemistry. However, during the fall 2016 semester, the department of chemistry developed an online section of General Chemistry I, a change that was not part of this research project, resulting in an increase in fall general chemistry enrollment levels. During the fall 2017 semester, a push for higher enrollment at the University of Utah resulted in university-wide freshmen enrollment increasing by 5.1% and sophomore enrollment by 7.2% compared to the previous fall semester. As a result, general chemistry enrollment during the fall 2017 semester reached the largest levels seen since the fall 2012 semester, with 1102 students enrolling in general chemistry (see Figure 8.1).

### 8.1.2 Number of Students Passing General Chemistry I

Acting as a baseline for this study, 722 of the 997 students enrolled in general chemistry passed the course during the fall 2013 semester. Implementation of course prerequisites in 2014 resulted in a sharp decline in students passing the course this semester compared to the previous year: 662 students passed General Chemistry I this semester. In spite of the decrease in enrollment in subsequent semester, the number of students passing the course value remained relatively stable during the fall 2015 semester, with 657 students passing general chemistry. In 2015, practice tests were

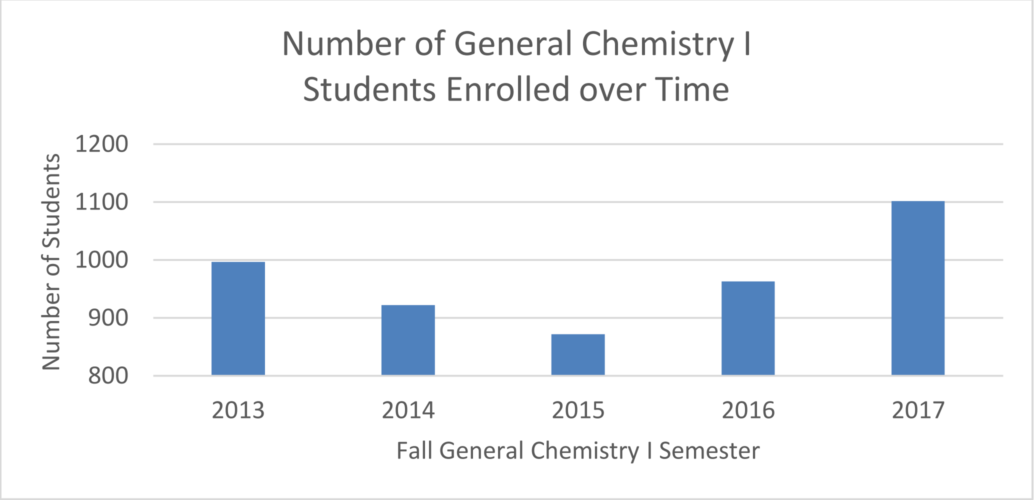


Figure 8.1 Fall General Chemistry I enrollment over time. The decrease in enrollment is attributed to the prerequisites, with the subsequent increase the result of the addition of online sections and increased university enrollment.

implemented in all course sections, and metacognitive training was added to the practice tests in one section.

Likely, these changes to course structure helped to maintain the number of students passing the course in spite of the enrollment drop. During the fall 2016 semester, the chemistry department developed an online section of general chemistry, greater structure was added to the flipped classroom in the online homework, weekly quizzes were incorporated in all sections, and weekly metacognitive training was implemented in one course section. Enrollment rose this semester when compared to fall 2015, but this effect was likely the result of increased enrollment and not due to changes to course curriculum (see section 8.1.3). However, during the fall 2017 semester, full metacognitive training and optional study assignments were implemented in all course sections. These changes, in conjunction with the large increase in General Chemistry I enrollment, resulted in the highest number of students passing general chemistry since

Dr. Atwood arrived at the University of Utah: 850 students passed the first semester chemistry course in fall 2017 (see Figure 8.2).

### 8.1.3 General Chemistry I Pass Rate

During the fall 2013 semester, a baseline of 72.4% students passed General Chemistry I. Curiously, though prerequisites were implemented and discussion attendance became required during the fall 2014 semester, the pass rate fell slightly when compared to the previous semester. However, in response to the incorporation of practice tests in the general chemistry program, the pass rate rose dramatically, in the fall 2015 semester. The following fall semester, 2016, involved the incorporation of weekly quizzes to the class structure, with greater structure added to the flipped classroom, yet pass rate decreased to 72.0%, a pass rate just below the fall 2013 baseline value. However, with the incorporation of metacognitive training and optional study

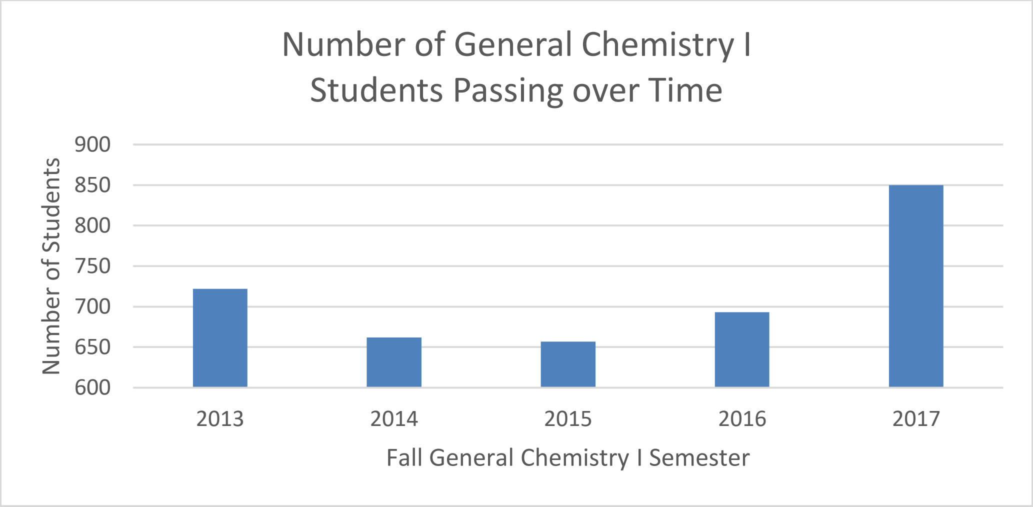


Figure 8.2 Fall General Chemistry I number of students enrolling over time. The drop in 2014 is the result of implementing prerequisites, and the large increase in 2017 is likely the cumulative effect of increased enrollment and improved course structure.

assignments in all courses, the pass rate in fall 2017 reached a higher level than seen during any semester since Dr. Atwood's arrival at the University of Utah: 77.1% of students passed General Chemistry I, just shy of a 5% improvement in General Chemistry I pass rate when compared to the fall 2013 baseline value (see Figure 8.3).

### 8.1.4 General Chemistry I Final Exam Scores

During the fall 2013 semester, the overall average on the ACS General Chemistry I final exam was 62.4%, with the lowest quartile scoring 38.7% on average. In response to the addition of prerequisites and the requirement of discussion attendance, fall 2014 witnessed a significant improvement in final exam score for all quartiles. In particular, the bottom quartile increased in final exam average to 45.4%, with the class overall averaging 66.7%. Final exam averages continued to increase over the next two semesters, climbing to an average of 69.0% for the class overall, with the bottom quartile averaging

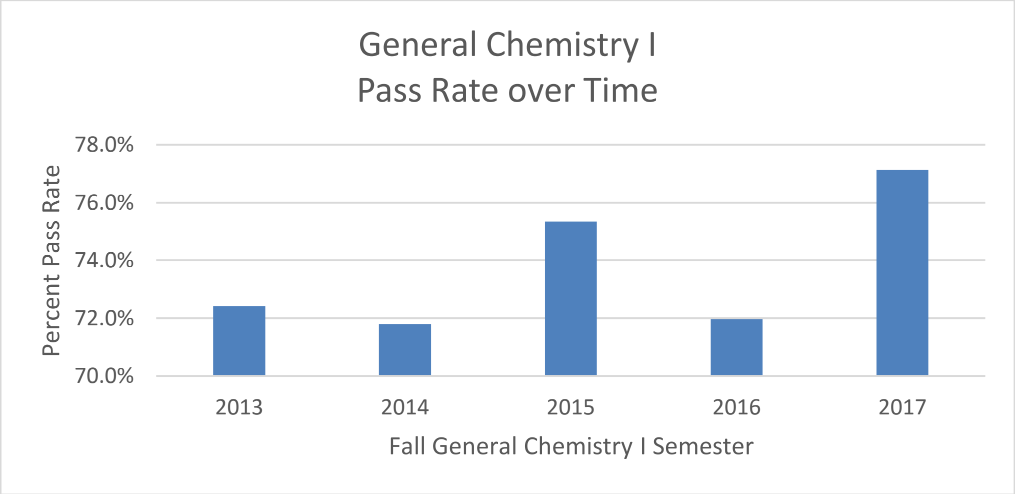


Figure 8.3 Fall General Chemistry I pass rate over time (*N* = 4856). Prerequisites and the requirement of discussion was implemented in 2014, practice tests in 2015, weekly quizzes in 2016, and metacognitive training in all sections in 2017.

46.7%, during the fall 2016 semester. This continual rise in final exam scores is likely the result of the addition of the practice tests, weekly quizzes, and flipped classroom structure during these semesters. Curiously, during the fall 2017 semester, final exam averages fell to levels observed during the fall 2014 semester (see Figure 8.4). One possible explanation for this could be a departmental change during this most recent semester, where the course section times taught by two of the professors were switched when compared to previous semesters. The number of students each course, student demographics, professor teaching styles, and the professors' general performance differed widely between the two sections. As such, swapping the times these two professors taught the course may provide some explanation for why the final exam averages fell during the fall 2017 semester.

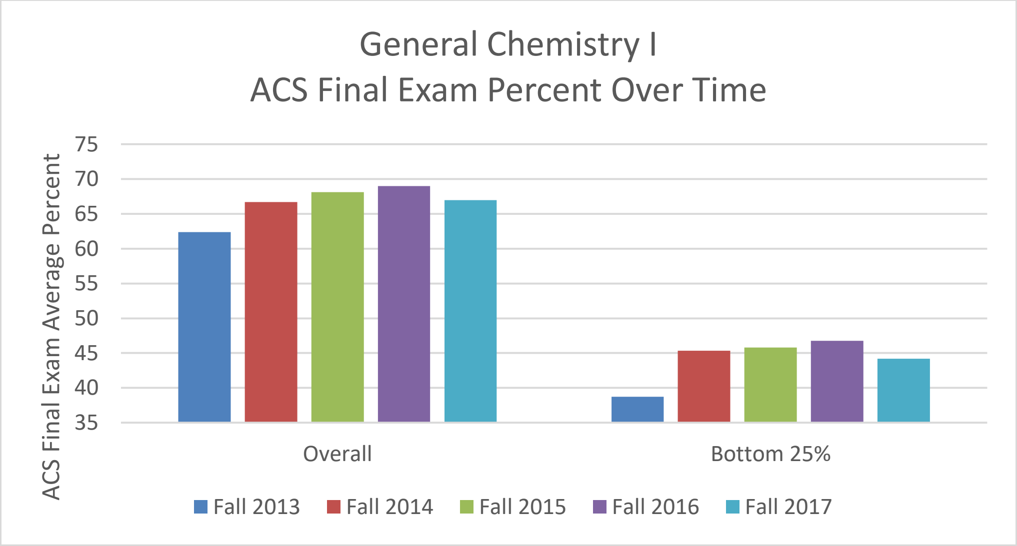


Figure 8.4 Overall and bottom quartile General Chemistry I final exam averages over time ($N$ = 3868). The most recent decrease in exam average may be the result of a change in the times that two professors taught the courses.

## 8.2 Changes to General Chemistry II

### 8.2.1 Changes to General Chemistry II Enrollment

Enrollment in General Chemistry II is influenced by more factors than during first semester general chemistry. In addition to the same factors that influenced General Chemistry I enrollment—the addition of general chemistry prerequisites in 2014, the development of online general chemistry sections in the 2016-2017 school year, and the change in university enrollment in 2017—enrollment in General Chemistry II is further affected by pass rates of General Chemistry I during the previous fall semester.

With this in mind a baseline of 593 students enrolled during the spring 2014 semester of second semester general chemistry. During the spring 2015 and 2016 semesters, enrollment fell both semesters, likely the result of the implementation of prerequisites during the fall 2014 semester, with General Chemistry II enrollment bottoming out at 472 students in spring 2016. However, enrollment climbed dramatically during the following semester: 604 students enrolled in the course. Finally, during the present, spring 2018, semester, currently 631 students are enrolled in second semester general chemistry, with more students enrolling in General Chemistry II this current semester than any previous semester since Dr. Atwood's arrival at the university (see Figure 8.5). Certainly, the changes to enrollment were influenced by the implementation of prerequisites, the addition of an online section in both General Chemistry I and II, and changes to university wide enrollment. However, the sharp increase in enrollment from spring 2016 to the current semester was likely also influenced by course structural elements incorporated in the general chemistry curriculum that improved student pass rate in General Chemistry I.

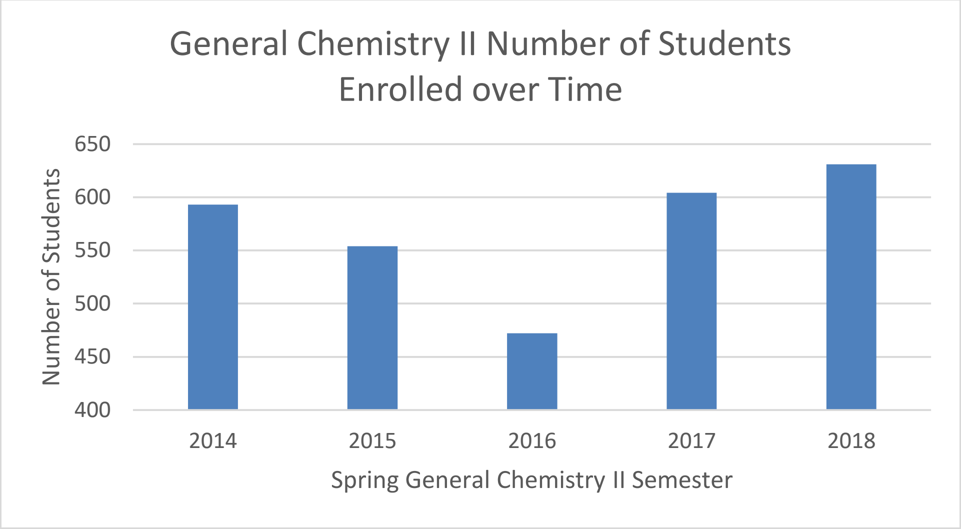


Figure 8.5 General Chemistry II enrollment over time. Mirroring the first semester, the decline witnessed is likely the result of the implementation of prerequisites, with the increase in 2017 and 2018 the combined result of increased number of students passing General Chemistry I and improved course structure.

### 8.2.2 Number of Students Passing General Chemistry II

As a baseline, 454 students passed General Chemistry II spring 2014 semester. The next two spring semesters witnessed a steady decline in the number of students passing General Chemistry II, dropping to only 402 students passing in 2016. This drop in the number of students passing the course likely is most influenced by the sharp decline in students enrolling in General Chemistry II during these semesters as a result of the incorporation of General Chemistry I prerequisites. The following semester, the number of students passing General Chemistry II increased significantly, with more students passing General Chemistry II than any other semester since Dr. Atwood’s arrival. During this semester, 510 students passed General Chemistry II, a 14.3% increase in students passing the course from the baseline value (see Figure 8.6). This number is influenced by implementation of the online courses during the 2016-2017 school year,

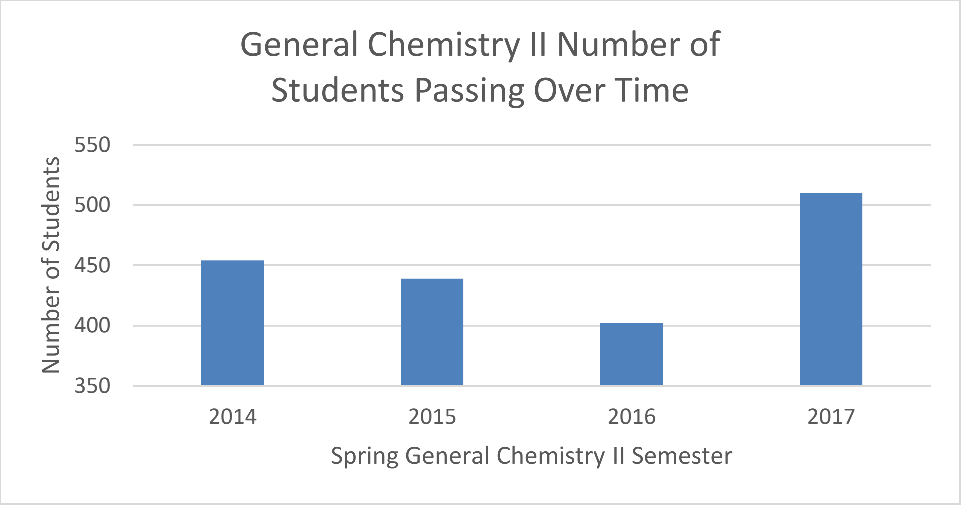


Figure 8.6 Number of students passing General Chemistry II over time. The decrease in students passing the course mirrors the decrease in students passing General Chemistry I as a result of the prerequisites. The increase in 2017 is the result of increased enrollment and improved course structure.

which generally increased general chemistry student enrollment. However, the increase in the number of students passing the course the fall 2016 semester in General Chemistry I, in conjunction with the incorporation of practice tests across the year of general chemistry and greater structure of the flipped classroom structure in the second semester, likely contributed to this sharp increase in the number of students passing second semester general chemistry.

### 8.2.3 General Chemistry II Pass Rate

During the spring 2014 semester, 77.0% of all students in General Chemistry II passed the course. Pass rate during the following spring semester increased somewhat, likely the result of the implementation of prerequisites in fall 2014 and requirement of discussion attendance during the 2014-2015 year of chemistry. However, during the spring 2016 semester, pass rate in General Chemistry II increased dramatically, with

85.2% of students passing General Chemistry II. The implementation of practice tests during both the fall and spring semesters of this school year in general chemistry likely influenced this increase. Additionally, the spring 2016 semester in General Chemistry II was the first semester to implement greater structure for the flipped classroom within the homework system. This would indicate that these research efforts incorporated into general chemistry resulted in a substantial increase in student pass rate in General Chemistry II during this semester when compared to previous semesters. Finally, during the spring 2017 semester, pass rate remained essentially unchanged from the previous semester, with 84.8% of students passing the class once more (see Figure 8.7). These results indicate a stable 8% increase in pass rate in General Chemistry II compared to the baseline pass rate as a result of these research efforts.

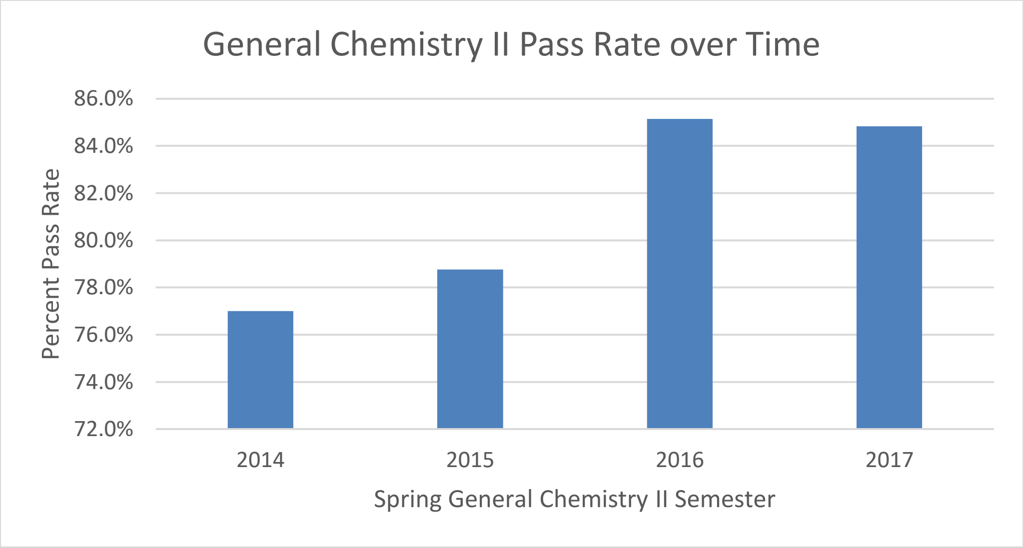


Figure 8.7 General Chemistry II pass rates over time (*N* = 2223). Prerequisites and required discussion attendance occurred in 2015, flipped classroom structure and practice tests in 2016, and weekly quizzes in 2017.

### 8.2.4 General Chemistry II Final Exam Scores

During the spring 2014 semester, students averaged 60.2% overall on the ACS General Chemistry II final exam percent, with the bottom quartile averaging 41.5%. During spring 2015, final exam scores rose, likely due to the implementation of prerequisites and the requirement of discussion attendance: the class overall averaged 63.9% and the lowest 25% of students scoring 46.0% on average. At the end of the spring 2016 semester, final exam scores reached their peak levels, with the class overall averaging 68.8% and the bottom quartile scoring 50.7% on average. This improvement was likely due to the incorporation of practice tests during both semesters of general chemistry and the increase in flipped classroom structure implemented in second semester chemistry. During the recent spring semester, 2017, second semester final exam averages decreased very slightly, with the bottom 25% of students averaging 49.2%, and the class overall averaging 68.0% (see Figure 8.8). These results demonstrate that the recent increase in student performance in General Chemistry II generally remained stable from one semester to the next.

## 8.3 Year-Long Enrollment and Pass Rates

As seen previously, fall General Chemistry I enrollment decreased during the fall 2014 and 2015 semesters as a result of the incorporation of course prerequisites for General Chemistry I. However, students had the ability of taking General Chemistry I after passing one of several university courses, such as college algebra and prep for general chemistry. As such, it was hypothesized that the implementation of these course prerequisites may have shifted student enrollment into the trailing semesters of general

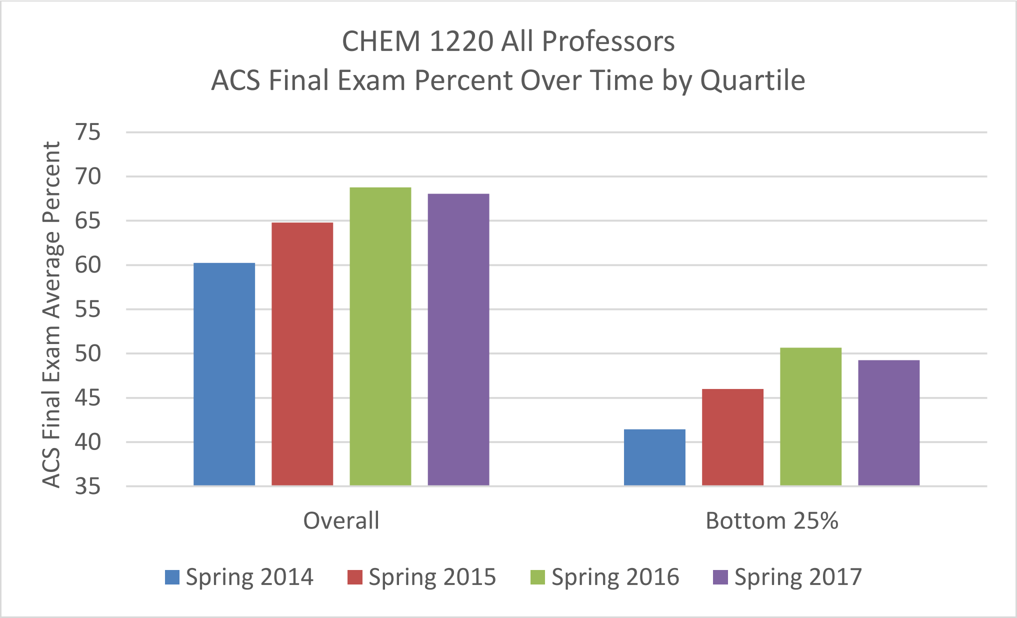


Figure 8.8 Overall and bottom quartile General Chemistry II final exam averages (*N* = 1748). Note the general improvement in final exam averages over time and the consistent performance from 2016 to 2017.

chemistry, resulting in lower fall enrollment without decreasing overall enrollment.

To test this hypothesis, full-year enrollment and pass rates in General Chemistry I and II were obtained from OBIA. These numbers represent the total numbers of students that enrolled in and passed General Chemistry I during the fall, spring, or summer semester of a particular school year. Correspondingly, for second semester the totals represent the number of enrolled and passing in General Chemistry II during the spring, summer, or fall semesters of a particular calendar year.

It should be noted that the standardization of curriculum, as well as the research changes mentioned in this dissertation, were only mandatory during the main semester sequence of the course, including a common syllabus, slides, and midterms, as well as the flipped classroom, quizzes and practice tests, and metacognitive training. In comparison, professors during the trailing semesters were given more freedom of interpretation during their course. Some course aspects were required in all courses during the main sequence and trailing semesters, including prerequisites, discussion attendance, the use of clickers, and testing using the ACS final exam. However, other course aspects, such as the use of the flipped classroom, the implementation of practice tests, weekly quizzes or any type of online homework system, or the implementation of metacognitive training, were only incorporated into some of the trailing semester courses.

### 8.3.1 Enrollment and Pass Rates in General Chemistry I

During the fall, spring, and summer semesters of the 2013-2014 school year, a total of 1693 students enrolled in general chemistry, with the majority of these students enrolling during the fall semester. General Chemistry I course prerequisites were implemented during the fall 2014 semester, resulting in a significant decrease in enrollment in the course during fall semester as seen previously. However, the spring and summer enrollments for General Chemistry I during this school year correspondingly increased, resulting in essentially the same year-long enrollment as the previous year: 1694 students enrolled in General Chemistry In 2014-2015. During the subsequent year, student enrollment decreased by 3% to reach 1641 students, but during the following year, 2016-2017, student yearly enrollment increased dramatically to 1735. Notably, though fall General Chemistry I enrollment decreased by 120 students from 2014 to 2016, the yearly drop in enrollment in this time period was much smaller. These results indicate that the implementation of prerequisites did not significantly decrease student General Chemistry I enrollment.

Of the 1693 students enrolled during the 2013-2014 school year, 1207 students passed General Chemistry I from the fall through the summer semesters. Curiously, the exact same number of students passed General Chemistry I during the 2014-2015 school year as the previous year. Though General Chemistry I enrollment dropped during the following school year when compared to the previous year, the total number of students passing the course actually increased to 1265 compared to the year before. Finally, during the most recent academic school year, from fall 2016 to summer 2017, the number of students passing remained nearly identical as the previous year, with 1256 students passing General Chemistry I from the fall to the summer semester (see Figure 8.9).

### 8.3.2 Enrollment and Pass Rates in General Chemistry II

For this set of data, each “year” was defined as beginning with the spring semester of the school year and concluded with the fall semester of the following school year. In other words, each year’s total value represents the number of students enrolled and passing General Chemistry II during the spring, summer, and fall semesters of a given calendar year. During each calendar year, the majority of students enrolled in General Chemistry II during the spring semester.

As a baseline, 1070 enrolled in General Chemistry II during the 2014 calendar year. The following two years witnessed a steady decline in student enrollment into General Chemistry II, reaching a study-wide low of 926 enrolling in second semester general chemistry during the 2016 calendar year. However, more students enrolled in General Chemistry II the following year than at any point in the study: 1146 students enrolled in second semester chemistry during the spring, summer, or fall semesters of 2017. This final increase in enrollment was likely due to the combined effect of the

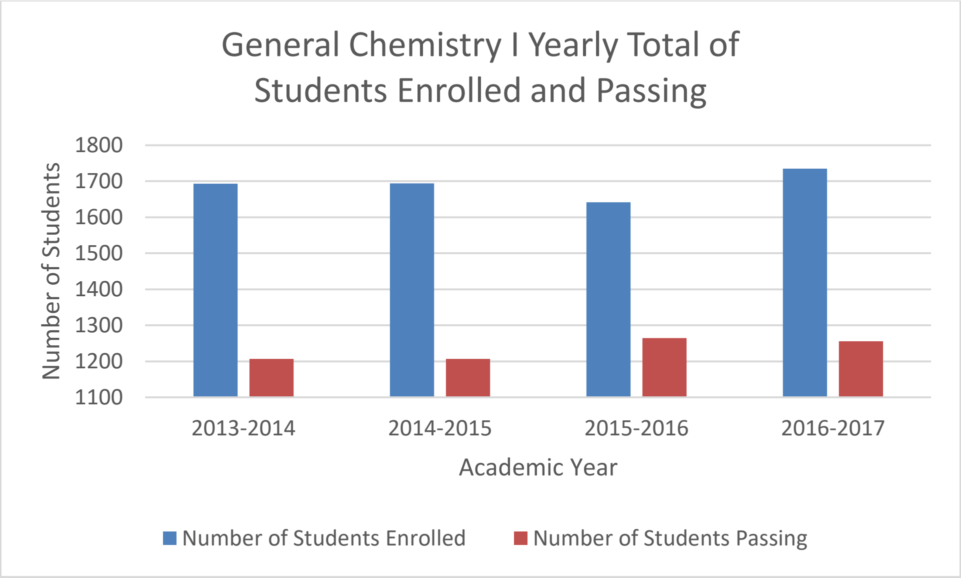


Figure 8.9 Yearly total of students enrolled in and passing General Chemistry I. Note that enrollment did not significantly decrease in 2014 in spite of the implementation of prerequisites.

incorporation of online course sections and higher numbers of students passing General Chemistry I during the 2016-2017 school year.

During the 2014 calendar year, a total of 884 of the 1070 students who enrolled in General Chemistry II passed the course. Afterward, the number of students passing General Chemistry II fell consistently over the next two calendar years, reaching a low of 775 students passing second semester General Chemistry In 2016. However, paralleling the increase in enrollment during the final calendar year, the number of students passing second semester General Chemistry Increased to 941 total students during 2017. This value represents the highest number of students passing General Chemistry II in a calendar year since the arrival of Dr. Atwood, with 6.4% more students passing than the baseline value (see Figure 8.10).

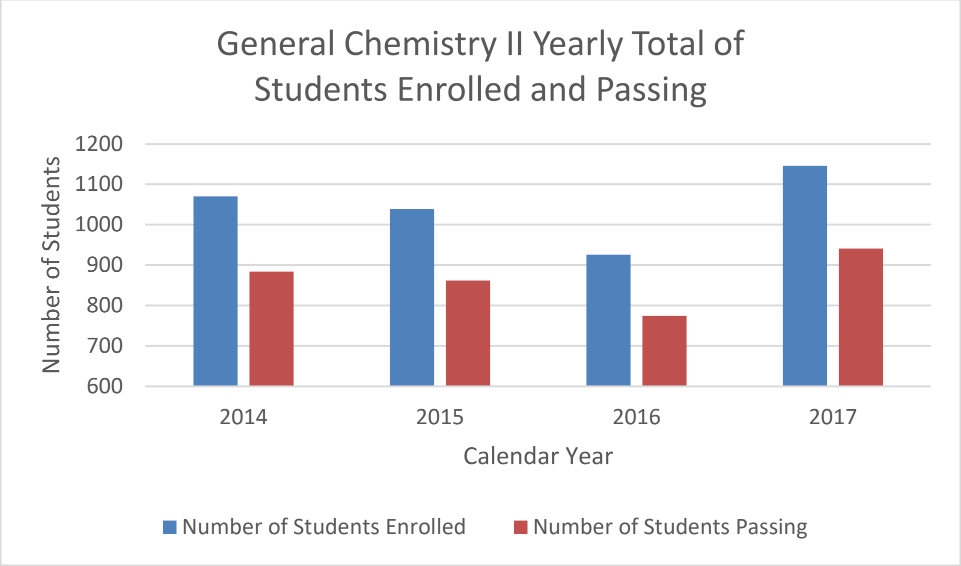


Figure 8.10 Yearly total of students enrolled in and passing General Chemistry II.

## 8.4 Enrollment and Pass Rates in Organic Chemistry

As one final analysis, we were interested in the long-term effects of the research efforts on student trajectory after general chemistry. In particular, we were interested if the enrollment and pass rates in the fall semester of organic chemistry I paralleled the trends observed in general chemistry. As the fall 2013 semester was considered the baseline semester for General Chemistry I, and as organic chemistry is usually taken one calendar year following the first semester of general chemistry, the fall 2014 semester was considered the corresponding baseline semester for organic chemistry 1. During this semester, a total of 338 of the 434 enrolled students passed the first semester organic chemistry. Enrollment in organic chemistry steadily decreased over the next two years, paralleling the trends observed in general chemistry, reaching a study low of 403 students enrolling in the course in fall 2016. During these two semesters, the number of students passing the course decreased slightly in 2015 but increased to the baseline value during the fall 2016 semester despite the decrease in enrollment. In fall 2017, the most recent organic chemistry semester, enrollment increased significantly, with 475 students taking the course during fall 2017. Of these students, 380 passed the course, more than any students since the beginning of this study (see Figure 8.11).

During the baseline semester, 77.9% of all students passed organic chemistry I, with the percent pass rate falling to 75.3% the subsequent school year. However, during the fall 2016 semester, pass rate in organic chemistry rose to 83.6%, falling only slightly to 80.0% during the fall 2017 semester (see Figure 8.12). These results indicate that enrollment rates, and total number of students passing organic chemistry I, were positively influenced by the trends established in general chemistry over the years.

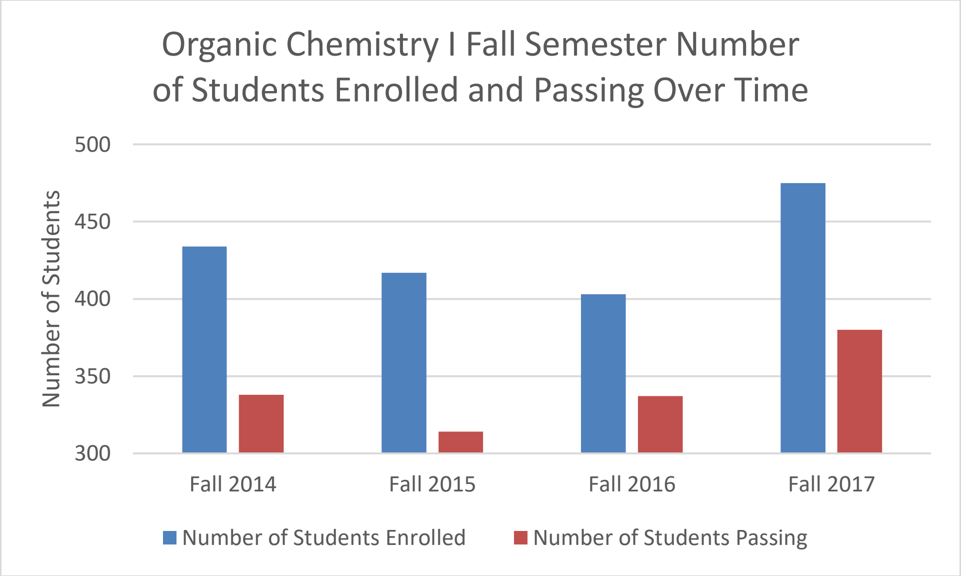


Figure 8.11 Number of students enrolled in and passing organic chemistry I over time.

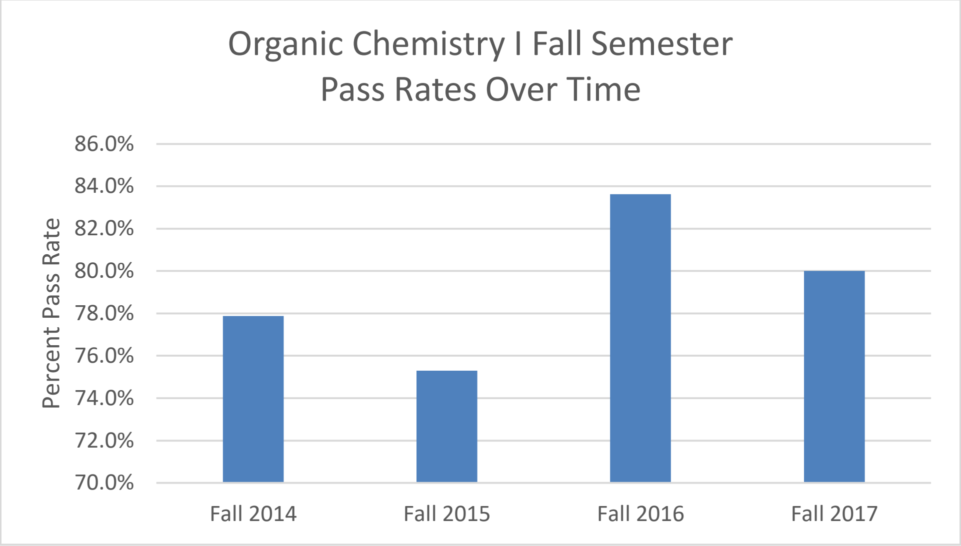


Figure 8.12 Pass rate in organic chemistry I over time (*N* = 2197)

## 8.5 Conclusion

The purpose of this research project was to improve student performance and pass rate in general chemistry without lowering academic standards. To accomplish these goals, four primary changes were made to the general chemistry curriculum as part of the research project: development of General Chemistry I prerequisites; requirement of discussion attendance; implementing greater structure for the flipped classroom already in place; and the implementation of practice tests, weekly quizzes, and metacognitive training for the general chemistry classes. In addition to these efforts, two additional changes influenced student enrollment in general chemistry that were not part of this research: the implementation of online sections of first and second semester general chemistry and the increase in freshmen and sophomore enrollment on a university-wide level.

In both General Chemistry I and II, main semester enrollment decreased after the implementation of prerequisites. However, yearly enrollment totals in General Chemistry

I do not demonstrate the steady decline observed during the fall semester, instead remaining fairly stable over this period of time. These results indicate that the incorporation of prerequisites generally shifted students to taking the course into later semesters instead of keeping them out of general chemistry. Also, with the addition of online sections of each course and the increase in university-wide enrollment, the enrollment in fall semester General Chemistry I reached levels not seen since the outset of this study. In conjunction with improved pass rates in General Chemistry I, enrollment levels in General Chemistry II have also been higher in the past two years than in any previous year during this research project.

In General Chemistry I, the number of students passing the course did decrease in fall 2014, paralleling the decrease in enrollment as a result of the addition of the course prerequisites. However, the number of students passing the course the subsequent semester remained unchanged in spite of lower enrollment values, indicating a course structure that helped more students succeed in the course during the fall 2015 semester. Paralleling the changes to enrollment during the fall 2016 and fall 2017 semesters, the number of students passing General Chemistry I steadily increased through to the most recent fall semester. During this semester, more students passed General Chemistry I than any previous semester since the initiation of this program.

In General Chemistry II, the number of students passing the course steadily decreased during the spring 2015 and 2016 semesters, paralleling the decrease in enrollment during these years. However, the following spring semester, 2017, witnessed a dramatic increase in the number of students passing the course when compared to the previous semester, reaching levels not seen in the general chemistry program since Dr.

Atwood’s arrival. Yearly numbers of students passing the course parallel these totals, with more students passing General Chemistry I in the most recent semesters than during the baseline year. Additionally, more students passed General Chemistry II during the most recent calendar year than in any other year under Dr. Atwood’s leadership.

Though the percent of students passing general chemistry has gone up and down over the past fall semesters, pass rates in the most recent semester of General Chemistry I reached the highest level observed during this research program. In General Chemistry II, pass rates steadily increased over the course of the program, plateauing in recent semesters at pass rates 8% higher than the baseline values. Additionally, final exam scores have consistently improved over the years, though the most recent semesters have witnessed a slight decrease in student performance. These results indicate that pass rates in general chemistry have increased, and the students passing general chemistry are doing so with greater proficiency in the course topics than observed at the outset of this project. This improved proficiency might explain why organic chemistry pass rates are higher in recent semesters than they have been in previous semesters.

Enrollment, numbers of students passing, pass rates, and final exam scores have gone up and down over the course of this research. However, results of the most recent years of general and organic chemistry demonstrate that more students are enrolling and passing these courses with higher pass rates than has been observed in any previous semester in this research effort. In conjunction with improved final exam scores, these results indicate that students have reached higher levels of proficiency in general chemistry than seen in any semester since the outset of this program.

# CHAPTER 9

# CONCLUSION

The purpose of this study was to improve the general chemistry pass rate over time without lowering academic standards for students. In general, the study encompassed four major efforts toward this goal: 1) identifying factors that were highly predictive of students failing general chemistry, 2) adjusting course structure to overcome these challenges, 3) incorporating curriculum changes to bring the course in alignment with the latest chemical education research, and 4) implementing metacognitive training to improve student performance in the course.

## 9.1 Identifying and Resolving Early Challenges to Passing General Chemistry

Early chemical education research indicated that student math ability was highly correlated to pass rates in general chemistry.[4,5] Paralleling these findings, we worked with OBIA and found that ACT and SAT math scores were highly predictive of general chemistry pass rates at the University of Utah. In particular, students with ACT math scores below 25 were found to have a pass rates that were 15.4% lower than students scoring 25 or above on the test. However, it was found that students with low incoming math test scores, after passing college algebra or the prep for general chemistry course, performed much better in general chemistry than those who went straight into general chemistry with poor math test scores.

In a subsequent study, we hoped to determine factors that predicted poor pass rate after students enrolled in the general chemistry. This study analyzed the results of General Chemistry II, where it was found that General Chemistry I grades were highly predictive of second semester pass rates: students who did poorly in General Chemistry I were most at-risk of failing second semester general chemistry. As part of this analysis, students were retroactively grouped into three categories based on their incoming first semester chemistry grade: “above average,” “average,” and “below average.” After, their performance in various aspects of the course were compared to their pass rates in each General Chemistry I grade category.

During this and previous semesters, TA-led discussion sections were part of the course structure but were not required. Upon review of students’ discussion attendance, it was found that students who were average or at-risk according to their General Chemistry I grade were among the least likely to attend discussion. However, students in these categories who regularly attended discussion had dramatically higher pass rates than those who rarely or never attended discussion.

With these results in mind, two early changes were made to the general chemistry course structure to improve student performance in the course. First, math and chemistry prerequisites were set in place: to enroll in General Chemistry I, students needed to 1) have a minimum math ACT or SAT test score, 2) pass college algebra or a higher math course at the university, or 3) pass the prep for general chemistry course. Additionally, the grading rubric was changed for general chemistry to include attendance in discussion sections.

In response to these curriculum adjustments, three major changes were noted in the general chemistry classroom. First, attendance in discussion rose dramatically. Second, the median score on the ACS general chemistry exam improved by nearly 3%. Third, enrollment during the fall semester of the course decreased over the next two years. However, yearly enrollment results indicated that spring and summer enrollment correspondingly increased, resulting in only a 3% drop in yearly general chemistry enrollment in this period of time.

## 9.2 Providing Greater Structure for the Flipped Classroom

After these changes were made, further efforts were made to the general chemistry curriculum to bring the course in line with the latest research in chemical education. Previous to the outset of this research project, the flipped classroom was implemented in the general chemistry courses at the University of Utah. However, survey results indicated that student utilization of the flipped classroom videos and the course textbook were poor, hampering their ability to be successful within the course structure.

To meet these needs, greater structure was developed for the flipped classroom system at the university. In particular, video links were embedded into tutorial homework questions, providing greater structure and accountability for students to utilize the flipped classroom videos created. During this semester, practice tests were also incorporated, providing students with limited opportunities toward metacognitive training. These changes were made in the second semester general chemistry course in the spring 2015 semester, and median scores on the final exam rose by 4.3% during this semester when compared to the spring 2014 semester. Additionally, when compared to the previous spring semester, pass rate rose 6.4%, reaching a study-high General Chemistry II pass rate of 85.2%.

## 9.3 Early Efforts toward Implementing Metacognitive Training

After review of educational research, it was hypothesized that poor metacognitive ability was hampering student success in general chemistry. In particular, previous research indicated that low performing students consistently overestimated their chemistry ability according to the Dunning-Kruger effect.[18-20] This overestimation was hypothesized to contribute to poor exam preparation among general chemistry students at the University of Utah, resulting in lower exam scores.

Within this framework, a system of metacognitive training was devised to provide students with opportunities for self-evaluation of their chemistry ability before taking course exams. Students were provided with practice tests to assess their chemistry ability before each course exam. During these practice tests, they predicted their ability and their score on each practice test, post-dicted their ability and score after taking the assessment, received feedback of their ability for each practice test topic, and developed study plans on how they would improve their chemistry ability.

On the first practice test of the semester, score prediction results were in line with the Dunning-Kruger effect: poor-performing students dramatically overestimated their assessment scores. However, by the end of the semester, students in all quartiles were accurately predicting their scores on all final exam practice tests. These results indicate the novel finding within chemistry education research that a system of metacognitive training within a system of practice tests can overcome the Dunning-Kruger effect. In regard to assessment performance, results of the bottom quartile were encouraging, with the bottom quartile improving 3% as a result of the metacognitive training when compared to a control section of the course. However, this improvement was not statistically significant.

These results demonstrated that metacognitive training as part of practice tests could overcome the Dunning-Kruger effect, helping even the poorest of students to accurately predict their chemistry ability on course assessments. Additionally, results of the final exam demonstrated an encouraging, though not statistically significant, improvement in exam average score among students in the bottom quartile on the assessments. These results indicated that metacognitive training had the potential of improving student performance, but the training was not sufficient for the results to be statistically significant. Notably, pass rates during this semester increased by 3.5% when compared to the previous semester, likely the result of the incorporation of practice tests in each section of the course.

## 9.4 The Effect of Regular Metacognitive Training in General Chemistry I

The results of the initial implementation of metacognitive training indicated that exam performance for the bottom quartile of students could potentially be improved through a system of self-reflection of chemistry ability. However, this effect was not strong enough to reach statistical significance. As such, it was hypothesized that more regular metacognitive training could help poor achieving students to significantly improve in exam performance. With this in mind, a system was devised to provide students with weekly opportunities for self-reflection, paralleling the metacognitive training established as part of the practice tests.

In this system, weekly quizzes were added to the existing system of metacognitive training through practice tests. For each weekly quiz, students predicted their score and their chemistry ability. After taking the quiz, students post-dicted their score and ability once more. They then received feedback regarding their prediction accuracy as well as their ability on each chemistry topic covered on the quiz. After, students completed a study plan in the following homework assignment, indicating their ability by each chemistry topic. The addition of weekly quizzes to the practice tests was meant to provide students with frequent opportunities throughout the semester for metacognitive self-reflection.

As with the previous year, the Dunning-Kruger effect was observed within student predictions on the first quiz: high achieving students accurately predicted their first quiz score, while low achieving students significantly overestimated their score on this assessment. However, by the final quiz of the semester, results indicate that the Dunning-Kruger effect was generally overcome, with students in all quartiles actually underestimating their ability on this assessment.

In regard to assessment scores, the study compared two General Chemistry I course sections: the control and the metacognitive section. On the first quiz, the average scores of the two sections were not statistically significantly different from each other, indicating that students in both sections had statistically similar average incoming ability. However, as the semester progressed, students in the metacognitive section, on average, significantly outperformed the students in the control section on each midterm and on the ACS final exam. These initial results indicated that regular metacognitive training helped students to significantly and consistently improve their performance on course exams when compared to the control section.

A subsequent analysis compared the results of the experimental semester to those of a previous, control semester to account for general differences between the course section. Final exam results indicated that, after accounting for differences in teaching style and course demographics, metacognitive training improved all students' average final exam score by 4.2%. Afterward, students were grouped into quartiles based on their final exam performance, with the bottom three quartiles demonstrating a statistically significant improvement in final exam averages. In particular, the second highest quartile improved in final exam average by 1.4%, the third highest by 5.2%, and the bottom quartile of students improved, on average, by 10.3% on the final exam as a result of metacognitive training in the course. These results demonstrate that a system that supports regular self-reflection of students' chemistry ability helps them to more accurately predict their assessment scores and to improve their exam scores. In particular, students in the bottom quartile were able to accurately predict their assessment scores by the end of semester and experienced the greatest improvement in exam scores as a result of this metacognitive training. These results indicate that metacognitive training targets the lowest performing students in the course, overcoming the Dunning-Kruger effect and helping them to improve in their chemistry ability.

## 9.5 The Cumulative Effect of a Year of Metacognitive Training in General Chemistry

The system of metacognitive training was implemented across the full year of general chemistry, with one section of each semester acting as the metacognitive section and the other sections acting as the control. The result of this system was that students could receive zero, one, or two full semesters of metacognitive training during the year depending on the instructors from whom they took each semester of general chemistry. Once more, the results of the final exam were compared to a previous semester to separate out the effect of general differences in teaching style and course demographics between the two sections. Results of this study indicated that metacognitive training did not significantly change overall students' average final exam score. However, upon grouping students by final exam quartile, it was found that metacognitive training significantly improved the bottom quartile of students' performance on the final exam by 2.7% with each semester of metacognitive received. Across the year of general chemistry, these students could have received as many as two full semesters of metacognitive training. As a result, students in the bottom quartile who received a full year of metacognitive training improved on average by 5.4% on the final exam relative to students who did not receive any metacognitive training during the year of general chemistry.

## 9.6 The Effect of Metacognitive Training on General Chemistry Course Trajectory

The results of the fall 2016 semester, which was the first General Chemistry I semester that metacognitive training was implemented as part of weekly quizzes, were then revisited. For this study, data from the metacognitive section was specifically analyzed to measure the effect of metacognitive training on student trajectory in the course. In particular, students' completion of study plans and their end-of-semester score prediction ability was compared to their trajectory over the semester. It was hypothesized students who regularly completed study plans and who accurately predicted their assessment score at the end of the semester would increase in metacognitive ability over the course of the semester. As a result, it was predicted that these students generally would improve in their course trajectory, as measured by an increase in course percentile over the semester, relative to students who did not.

In this study, students were assigned a percentile rank based on their score on each of five major assessments during the semester. After, a slope was generated that represented their student trajectory over the semester. In particular, this slope demonstrated students’ average change in class percentile from one major assessment to the next, with positive values representing students who were improving in their percentile over the course of the semester. Results indicated that students who regularly completed study plans and who accurately predicted their score on the final quiz of the semester generally had positive course trajectory. In comparison, students who did not regularly complete these study plans or who did not predict their final quiz score accurately tended to regress over the semester, decreasing in their course percentile. Further analysis of this data indicated that the effect of the completion of study plans was, in part, mediated by the effort that students put forth in the semester as measured by regular completion of homework assignments. This demonstrates that the effect of study plans is, in part, a measure of students’ general effort across the semester. However, results also indicate that study plans may have uniquely influenced student course trajectory after accounting for student effort, though this effect was outside of a 95% confidence interval.

General results of this study indicate that students who regularly completed the study plans and who learn to predict their assessment scores accurately tended to improve over the course of the semester. Both of these effects were hypothesized to improve students' metacognitive ability over the semester, with results indicating that students who utilized both aspects of this system of metacognitive training generally improved in their trajectory over the semester. These results, then, provide some explanation for the results of the year of metacognitive training: the students who regularly took part in the system of metacognitive training became more self-aware of their chemistry ability, learned how to improve their chemistry ability, and improved in their course trajectory as a result.

## 9.7 General Changes to Student Enrollment and Performance in General Chemistry

We then analyzed the long-term effects of these study efforts on four major course aspects: semester and yearly enrollment, semester and yearly number of students passing the course, semester pass rates, and semester final exam scores.

During this study, three specific factors influenced student General Chemistry I enrollment over the years: the implementation of course prerequisites, the addition of online sections of general chemistry, and a general increase in undergraduate enrollment at the University of Utah. Of these three, only the first was influenced by changes made as part of this research effort. Additionally, General Chemistry II enrollment was further impacted by pass rates in General Chemistry I. Results of the fall and spring semesters indicated that main semester enrollment generally declined for two years as a result of the implementation of course prerequisites. However, with the addition of online sections and with the university enrollment increasing, the number of students enrolled in general chemistry during the fall and spring semesters have increased the highest levels witnessed at the university since the outset of this research project.

In comparison to these results, yearly enrollment in General Chemistry I only slightly decreased as a result of the implementation of the prerequisites, with a significant number of students enrolling in the trailing spring and summer semesters after having fulfilled the course prerequisites. Curiously, while general chemistry enrollment remained stable, yearly enrollment in General Chemistry II declined in the two years following the implementation of prerequisites. However, in both General Chemistry I and II, the yearly total of students enrolling in General Chemistry Increased in the most recent year, reaching the highest levels witnessed in the program since the arrival of Dr. Atwood.

The number of students passing General Chemistry I and II is influenced, first, by the number of students enrolling in the course. However, course structural changes, such as the increased structure of the flipped classroom or the addition of practice tests or weekly quizzes, also influences this total. Paralleling the changes to the number of students enrolling in General Chemistry I, the number of students passing General Chemistry I fell during the fall 2014 semester as a result of the implementation of prerequisites for the course. However, in spite of decrease in enrollment during the subsequent year, the number of students passing general chemistry the following year, fall 2015, remained essentially stable when compared to the previous year. During this semester, practice tests were implemented in the course, helping a greater percentage of students to pass the course, resulting in a stable number of students passing the course when compared to the previous semester. Over the next two years, the number of students passing General Chemistry I increased dramatically. This increase could be explained, in part, by the high student enrollment during the fall 2017 semester. However, metacognitive training was incorporated in all sections of general chemistry during this semester, contributing to the number of students passing the course. The result was that far more students passed first semester chemistry in fall 2017 than any previous semester in this research effort.

In General Chemistry II, the number of students passing the course during the spring semesters steadily declined for two years, paralleling the decrease in enrollment observed as a result of the implementation of course prerequisites. However, more students passed General Chemistry II in the spring 2017 semester than any other semester under Dr. Atwood. This increase was certainly influenced by the rise in General Chemistry I enrollment due to the addition of online sections of the course but was likely influenced by the high General Chemistry II pass rates resulting from metacognitive training and improved flipped classroom structure.

While the number of students passing General Chemistry I in the fall semester significantly declined the two years following the incorporation of the course prerequisites, yearly numbers of students passing the course actually increased over this period. Additionally, during the two most recent academic years, the numbers of students passing General Chemistry I across the academic year reached a study-wide high. In comparison, the number of students passing General Chemistry II in a given calendar year decreased the two years following the implementation of course prerequisites, but more students passed General Chemistry II during the most recent calendar year, 2017, than any previous year of this study.

Fall semester pass rates have gone up and down in the course of the study. For example, pass rates remained relatively stable during the fall 2013 and fall 2014 semesters in spite of the addition of prerequisites and the requirement of discussion. However, fall 2015 witnessed a significant increase in the percent of students passing the course compared to the previous year, likely due to the incorporation of practice tests in the semester. Curiously, pass rates during the following semester, fall 2016, fell once more in spite of the increased structure in the flipped classroom and the addition of weekly quizzes. However, with the addition of metacognitive training in all sections of the General Chemistry I course, pass rates in General Chemistry I reached the highest levels observed since the outset of this project, with 77.1% of students passing the course.

In comparison to the General Chemistry I results, pass rate in General Chemistry II steadily climbed from 2014 to 2016, with pass rates during the spring 2016 semester reaching the highest levels observed in General Chemistry II since Dr. Atwood was hired. During the subsequent semester, the pass rate remained virtually unchanged, indicating a stable change to General Chemistry II pass rate.

Over this period of time, overall final exam scores, and scores for the bottom quartile of students, steadily climbed from fall 2013 to fall 2016 in General Chemistry I and from spring 2014 to spring 2016 for General Chemistry II. During the fall 2017 semester, final exam scores decreased somewhat in General Chemistry I, likely the result of a reordering of the times that two professors taught general chemistry. In comparison, final exam scores in General Chemistry II plateaued, with the students scoring very similarly during the spring 2017 semester when compared to the previous year.

## 9.8 Concluding Remarks

The purpose of this study was to improve student performance and pass rates in general chemistry. Initially, factors that significantly contributed to student fail rate were identified, including poor math ability and poor attendance at discussion. Course prerequisites were implemented and the course grading scale was adjusted to meet these needs. The flipped classroom had previously been implemented in this period of time, but survey results indicated that relatively few students utilized the video and textbook resources. As such, the tutorial video questions were then embedded in homework assignments mimicking the study cycle, providing greater structure for the flipped classroom system. Each change resulted in improved ACS final exam performance.

We finally hypothesized that poor metacognitive ability among the lowest achieving students was hampering their success. A system of metacognitive practice tests was implemented, resulting in an elimination of the Dunning-Kruger effect among the students. However, the system was not effective enough to result in statistical improvements in exam scores for any quartile. As a result, the system was expanded to include metacognitive training as part of weekly quizzes, a change that improved student performance on each exam over the system. After accounting for differences in teaching style and course demographics, the bottom three quartiles were all found to improve on the final exam as a result of this system, with the bottom quartile experiencing a 10% improvement in final exam average. During the second semester of a year-long effort to incorporate metacognitive training in general chemistry, the bottom quartile was found to improve once more in final exam score as a result of metacognitive training. In fact, for each semester of metacognitive training that students received, the bottom quartile of students improved in final exam score by 2.7%. The result is a potential final exam average increase of 5.4% as a result of a year of metacognitive training.

Upon further analysis of the fall 2016 semester results, metacognitive training was found to impact students' percentile trajectory in the course. In particular, as students completed more study plans and were able to predict their scores accurately by the end of the semester, they improved in course trajectory. Though some of this effect was mediated by the general effort that students put forth in the class, these results indicate that students who took advantage of the regular metacognitive training improved in their course performance over the semester of general chemistry.

At the conclusion of this study, student enrollment, the number of students passing the course, and student pass rates are at their highest level since the outset of this research project. Though there has been some decline in final exam scores in the most recent semester of General Chemistry I, averages on the ACS final exams, including averages of the bottom quartile, have generally increased over the course of this study. The purpose of this research project was to improve student performance in and pass rate of general chemistry. The results summarized in this dissertation indicate that the research efforts of this study have been successful, improving student performance in each course measure in both General Chemistry I and II.

# APPENDIX A

## SAMPLE ONLINE FLIPPED CLASSROOM

## HOMEWORK QUESTION

The following question represents a sample flipped classroom question within the Madra Learning Homework System, with the correct answers underlined. Note that this assignment has more videos than the typical assignment.

Topic: Protons, Neutrons, and Electrons: Khan Academy

To prepare for this question, watch the following videos:1st Video Link, 2nd Video Link, 3rd Video Link, 4th Video Link. Additionally, please read the following chapter through the section 'Atomic Mass vs Mass Number: Chapter Link.

Several isotopes of lead exist in nature: lead-204, lead-206, lead-207, and lead-208. What must be the difference between each of these?

a. The number of protons
b. The number of neutrons
c. Either the number of protons or neutrons
d. Neither the number of protons nor the number of neutrons

Lead-206 is the end product of the radioactive decay of uranium. Which of the following are included in the number '206?' Select all that apply.

a. Protons
b. Neutrons
c. Electrons
d. None of the above

How many protons are in lead-206? 82

How many neutrons are in lead-206? 124

Complete the following atomic symbol for lead-206. What would you enter for X? $^{A}_{Z}X$
Pb

What would you enter for A? 206

What would you enter for Z? 82

Iron can exist as several different ions, with the most common of these being $Fe^{2+}$, $Fe^{3+}$, and $Fe^{6+}$. What must be the difference between these? Select all that apply.

a. The number of protons
b. The number of neutrons
c. The number of electrons
d. None of the above

Which are used to calculate the charge of $Fe^{2+}$? Select all that apply.
a. The number of protons
b. The number of neutrons
c. The number of electrons

How many protons are in $^{55}Fe^{2+}$? 26

How many neutrons are in $^{55}Fe^{2+}$? 29

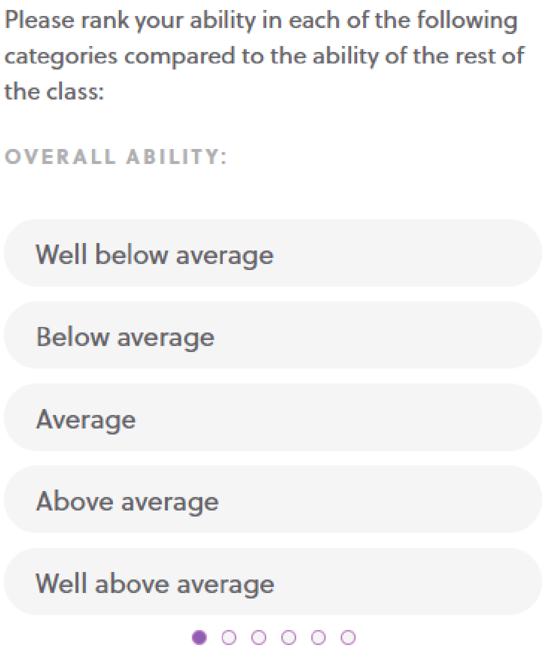
How many electrons are in $^{55}Fe^{2+}$? 24

# APPENDIX B

## SCREENSHOTS OF THE METACOGNITIVE TRAINING CYCLE

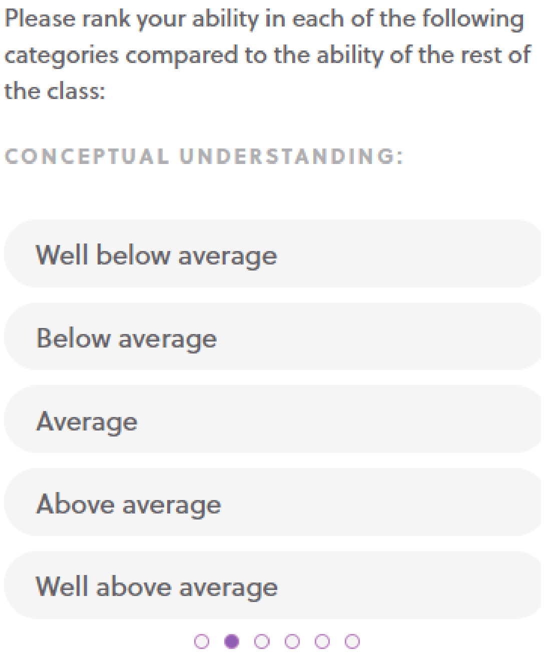
The below screenshots were taken of a simulated student’s results on a practice test regarding intermolecular forces, stoichiometry, and aqueous reactions. These screenshots are provided to guide to reader on the full method of implementation of each step of the metacognitive training in the study.

Steps 1 and 2: Score and ability analysis: Students predict their score and ability by question types before taking the taking the assessment. After completing the assessment, they answer each of these questions once more.



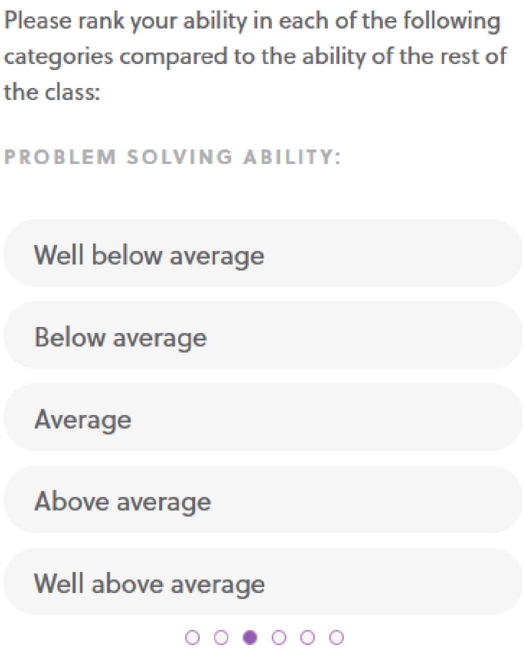
Please rank your ability in each of the following categories compared to the ability of the rest of the class:

OVERALL ABILITY:



Please rank your ability in each of the following categories compared to the ability of the rest of the class:

CONCEPTUAL UNDERSTANDING:



Please rank your ability in each of the following categories compared to the ability of the rest of the class:

PROBLEM SOLVING ABILITY:

Well below average

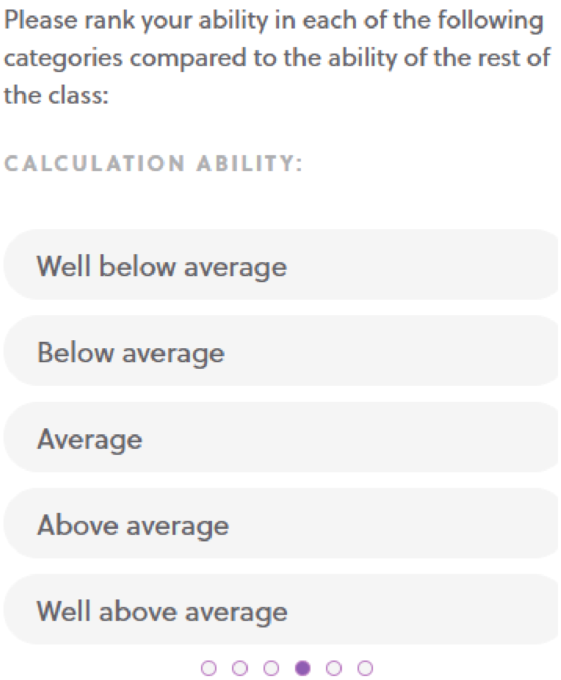
Below average

Average

Above average

Well above average

○ ○ ● ○ ○ ○



Please rank your ability in each of the following categories compared to the ability of the rest of the class:

CALCULATION ABILITY:

Well below average

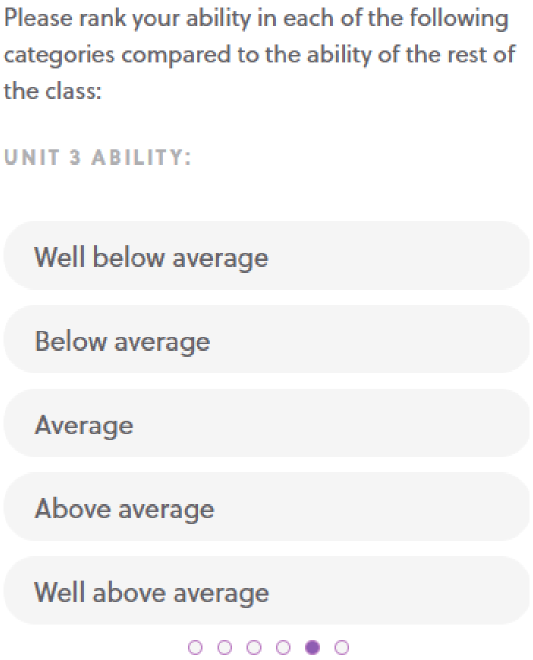
Below average

Average

Above average

Well above average

○ ○ ○ ● ○ ○



Please rank your ability in each of the following categories compared to the ability of the rest of the class:

UNIT 3 ABILITY:

Well below average

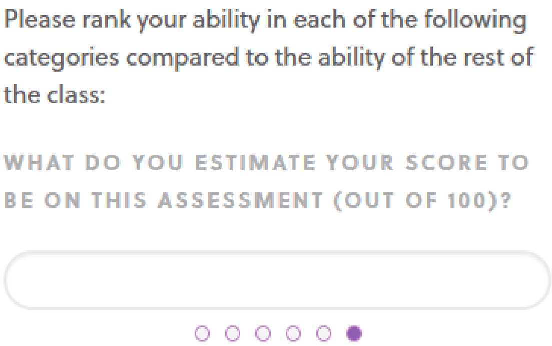
Below average

Average

Above average

Well above average

○ ○ ○ ○ ● ○

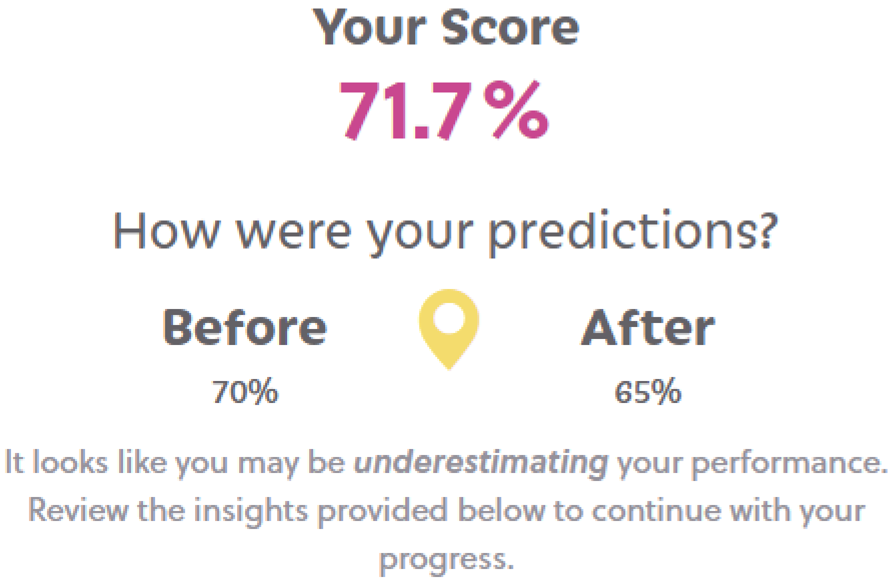


Please rank your ability in each of the following categories compared to the ability of the rest of the class:

WHAT DO YOU ESTIMATE YOUR SCORE TO BE ON THIS ASSESSMENT (OUT OF 100)?

○ ○ ○ ○ ○ ●

Step 3: The Feedback Process
Step 3a: Score and Prediction Accuracy Feedback: students receive feedback on their assessment score and prediction and post-diction accuracy:



**Your Score**

**71.7%**

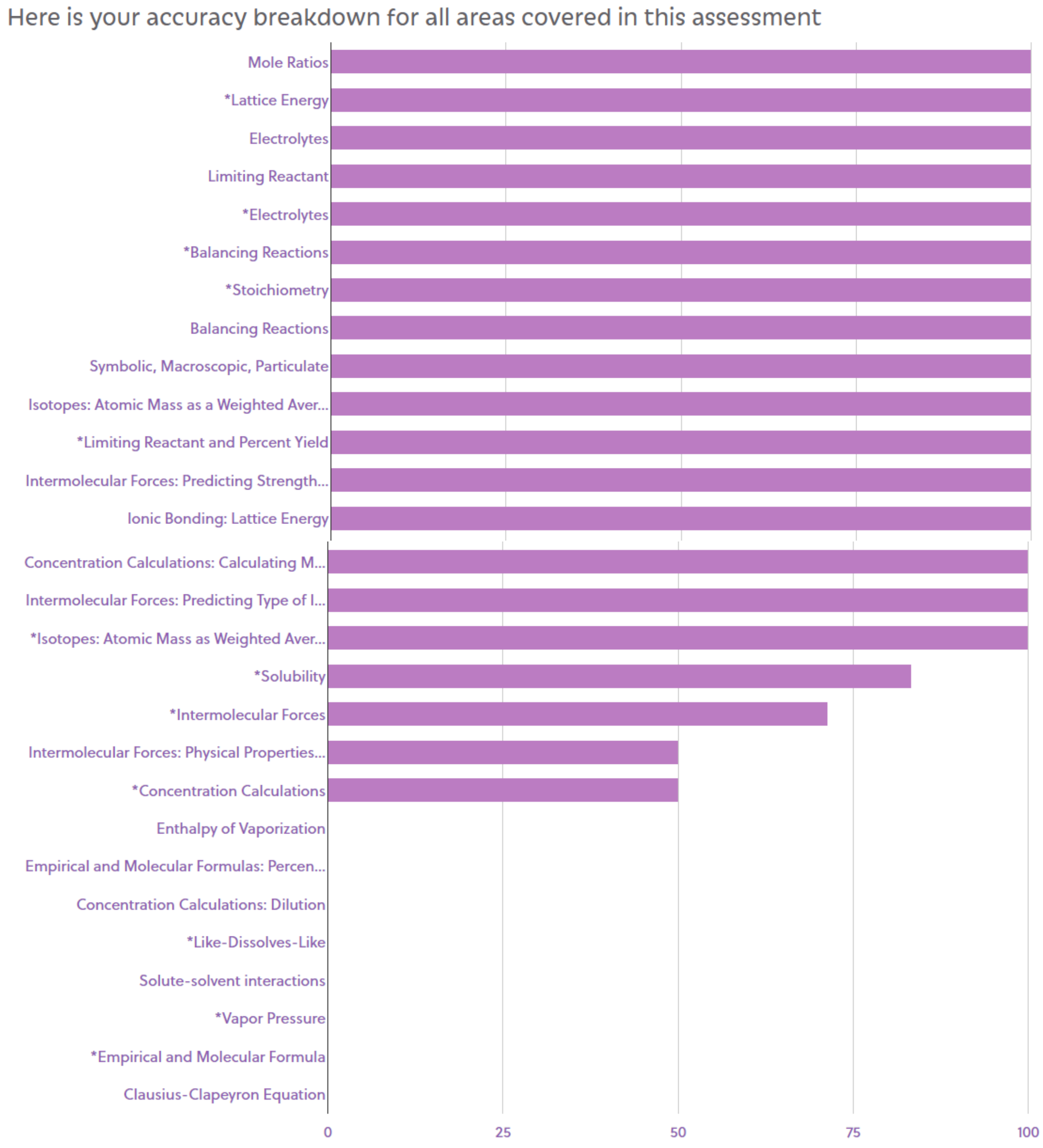
How were your predictions?

**Before**
70%

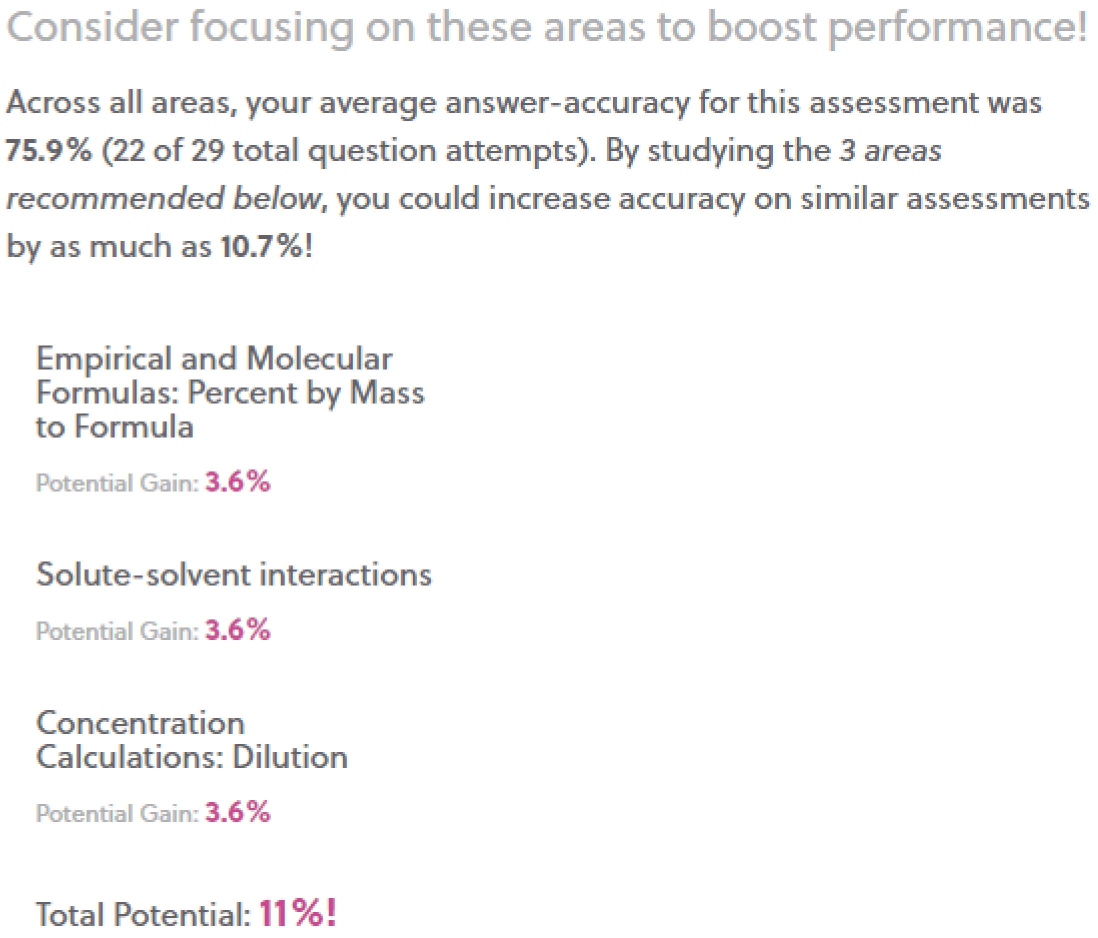
**After**
65%

It looks like you may be ***underestimating*** your performance. Review the insights provided below to continue with your progress.

Step 3b: Feedback by Assessment Topic: Students receive feedback on their ability by assessment topic. Note: the topics with ‘asterisks’ represent major topics, and those without represent specific topics within those major topics. Also note that ‘accuracy’ is calculated by the number of attempts within a question category that they got correct. If a question had multiple parts, the accuracy was calculated considering for each part of each question as a separate ‘attempt’ within a topic.



Step 3c: Students receive feedback on key areas they can focus their studying upon to improve their ability:



## Consider focusing on these areas to boost performance!

Across all areas, your average answer-accuracy for this assessment was **75.9%** (22 of 29 total question attempts). By studying the *3 areas recommended below*, you could increase accuracy on similar assessments by as much as **10.7%**!

Empirical and Molecular Formulas: Percent by Mass to Formula

Potential Gain: **3.6%**

Solute-solvent interactions

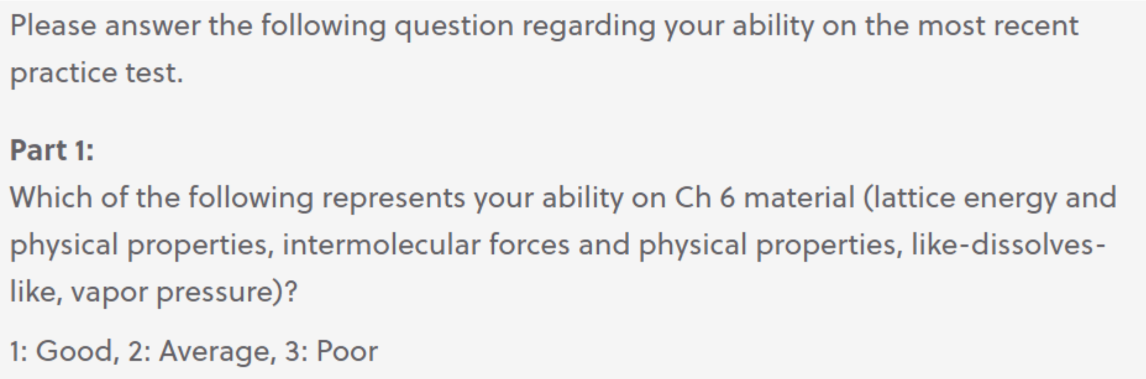
Potential Gain: **3.6%**

Concentration Calculations: Dilution

Potential Gain: **3.6%**

Total Potential: **11%!**

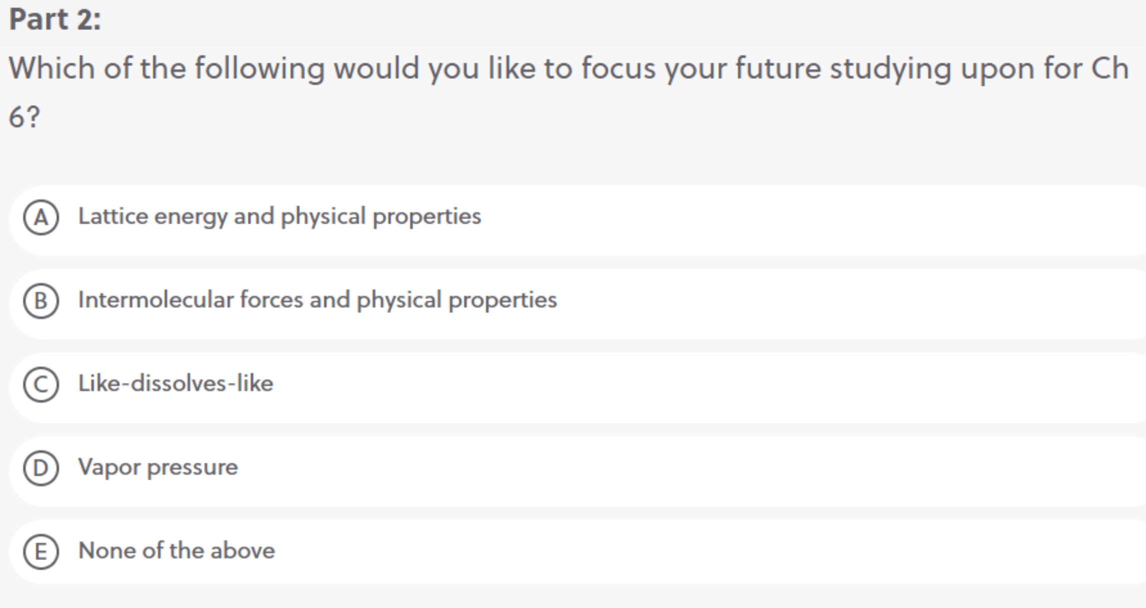
Step 4: Students create study plans by noting areas of weakness and strength and selecting topics to focus future studying upon. Note, following a quiz, students were only given questions similar to parts 1, 3, and 5.



Please answer the following question regarding your ability on the most recent practice test.

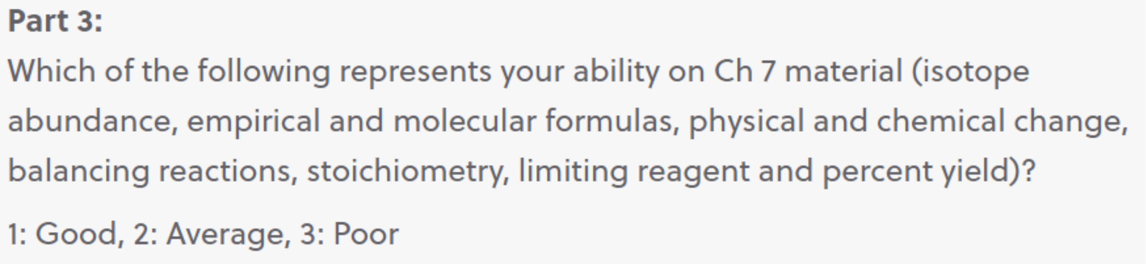
**Part 1:**
Which of the following represents your ability on Ch 6 material (lattice energy and physical properties, intermolecular forces and physical properties, like-dissolves-like, vapor pressure)?

1: Good, 2: Average, 3: Poor



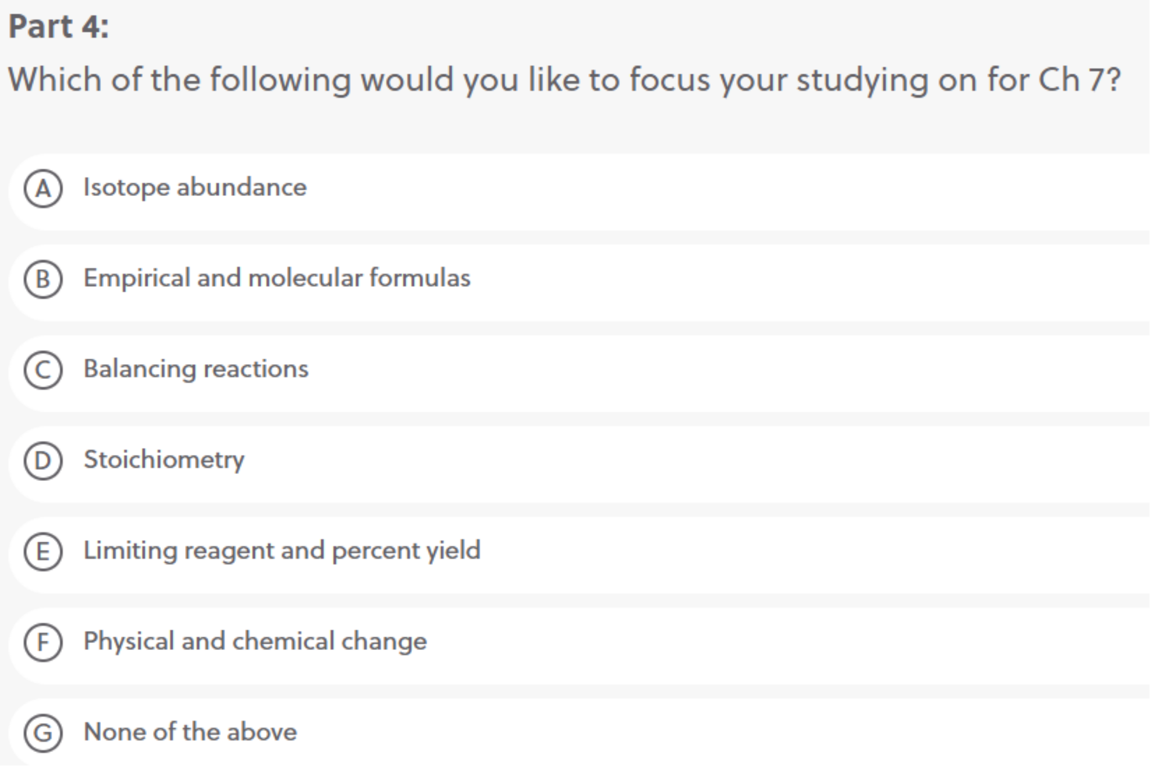
**Part 2:**
Which of the following would you like to focus your future studying upon for Ch 6?

- Ⓐ Lattice energy and physical properties
- Ⓑ Intermolecular forces and physical properties
- Ⓒ Like-dissolves-like
- Ⓓ Vapor pressure
- Ⓔ None of the above



**Part 3:**
Which of the following represents your ability on Ch 7 material (isotope abundance, empirical and molecular formulas, physical and chemical change, balancing reactions, stoichiometry, limiting reagent and percent yield)?

1: Good, 2: Average, 3: Poor



**Part 4:**

Which of the following would you like to focus your studying on for Ch 7?

Ⓐ Isotope abundance

Ⓑ Empirical and molecular formulas

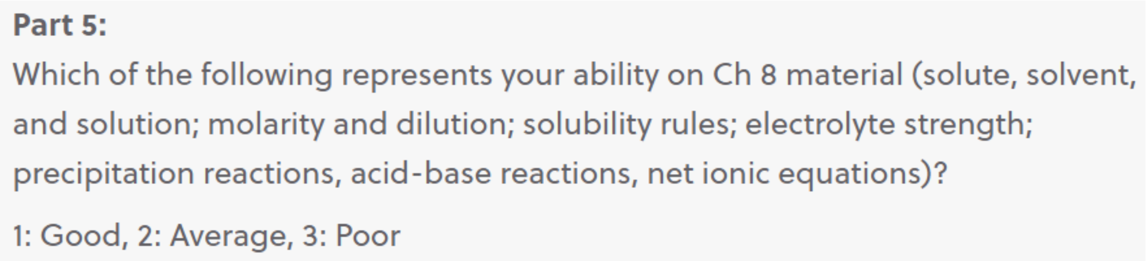
Ⓒ Balancing reactions

Ⓓ Stoichiometry

Ⓔ Limiting reagent and percent yield

Ⓕ Physical and chemical change

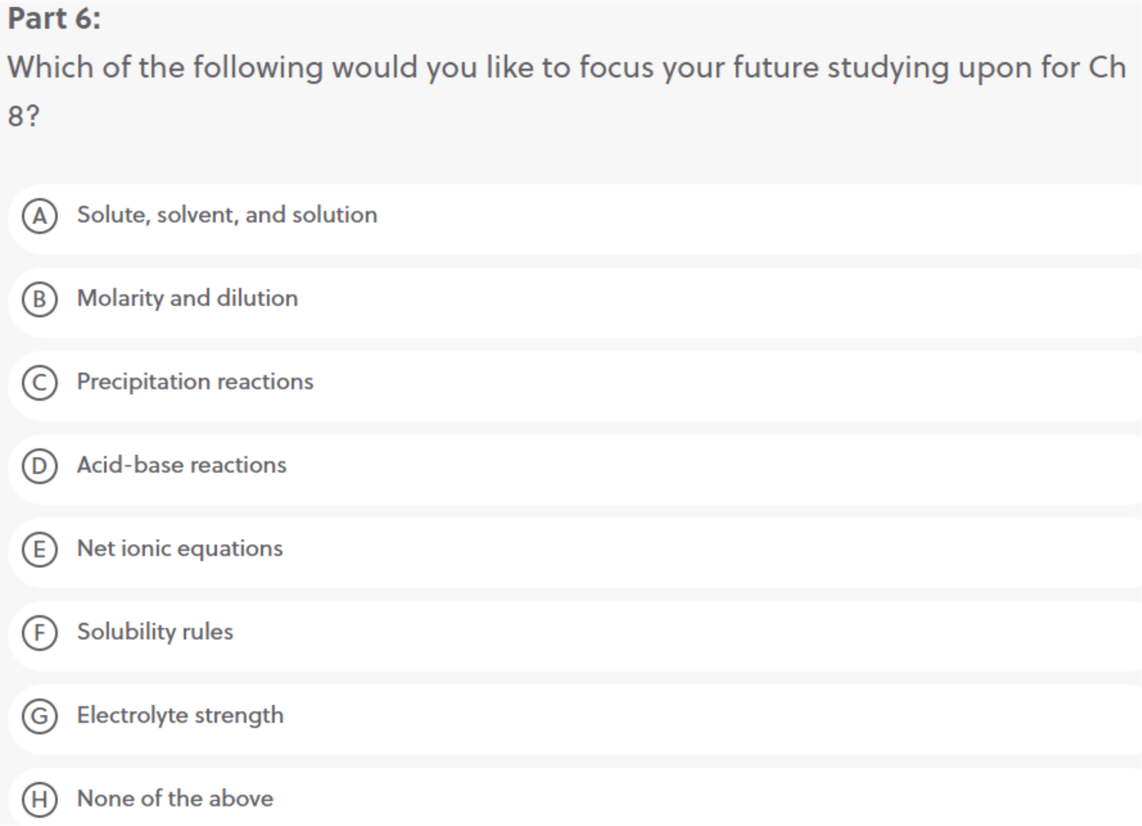
Ⓖ None of the above



**Part 5:**

Which of the following represents your ability on Ch 8 material (solute, solvent, and solution; molarity and dilution; solubility rules; electrolyte strength; precipitation reactions, acid-base reactions, net ionic equations)?

1: Good, 2: Average, 3: Poor



**Part 6:**

Which of the following would you like to focus your future studying upon for Ch 8?

Ⓐ Solute, solvent, and solution

Ⓑ Molarity and dilution

Ⓒ Precipitation reactions

Ⓓ Acid-base reactions

Ⓔ Net ionic equations

Ⓕ Solubility rules

Ⓖ Electrolyte strength

Ⓗ None of the above

# APPENDIX C

## CALCULATING INTERACTION SLOPE VALUES:

## A DIFFERENCE OF DIFFERENCES

The below information demonstrates how the interaction can be calculated overall and for each quartile. In each case, the difference in average ACS final exam performance was calculated for each semester from the control to the metacognitive professor. After, the interaction was calculated by finding the difference in these ‘difference’ values. For example, overall the metacognitive professor did 0.4% better than the control professor in the control semester, while doing 4.6% better in the experimental semester. The difference of 4.6% and 0.4% produces the value of 4.2% for the overall interaction. Within rounding error, these values are identical to the unstandardized B values of the interaction regression given in the text.

| | | ACS Final Exam Average | | | |
|---|---|---|---|---|---|
| | | Control Semester | Experimental Semester | Difference | Interaction |
| Overall | Control Professor | 66.8% | 67.1% | 1.5% | 4.3% |
| | Metacognitive Professor | 68.3% | 72.9% | 5.8% | |
| Bottom Quartile | Control Professor | 45.5% | 43.3% | 0.6% | 10.2% |
| | Metacognitive Professor | 46.1% | 54.1% | 10.8% | |
| 3rd Quartile | Control Professor | 62.2% | 63.1% | 1.0% | 5.2% |
| | Metacognitive Professor | 63.2% | 69.3% | 6.2% | |
| 2nd Quartile | Control Professor | 72.9% | 74.9% | 2.9% | 1.4% |
| | Metacognitive Professor | 75.8% | 79.2% | 4.3% | |
| Top Quartile | Control Professor | 85.9% | 87.0% | 2.5% | -0.3% |
| | Metacognitive Professor | 88.4% | 89.2% | 2.2% | |

# APPENDIX D

## QUESTIONS FROM THE INTRO QUIZ FOR FALL 2016

The following questions are from the intro quiz during the fall 2016 semester, which correlated to final exam performance with an $R^2$ of 0.33. These questions were completed within the Madra Learning Homework System, with the correct answers underlined.

1. Three individual students using different methods to determine the mass and volume of three different copper rods calculate the copper density as indicated in the table below. Note: the true density of copper is 9.04 g/mL.

| **Density Measurement** | **Student A** | **Student B** | **Student C** |
|---|---|---|---|
| 1 | 8.75 g/cm$^3$ | 9.04 g/cm$^3$ | 9.05 g/cm$^3$ |
| 2 | 8.62 g/cm$^3$ | 9.03 g/cm$^3$ | 8.88 g/cm$^3$ |
| 3 | 10.96 g/cm$^3$ | 9.01 g/cm$^3$ | 8.85 g/cm$^3$ |
| Average | 9.44 g/cm$^3$ | 9.03 g/cm$^3$ | 8.93 g/cm$^3$ |

Which student that has the most accurate density measurement? Student B

Which student that has the most precise density measurement? Student B

2. How many protons, neutrons, and electrons are present in one ION of $^{48}\text{Sc}^{3+}$?
Protons: 21, Neutrons: 27, Electrons: 18

3. A sheet of paper is burned in the presence of oxygen, producing carbon dioxide and water vapor. If the sheet of paper weighed 4.5 g and the masses of carbon dioxide and water vapor produced were 5.6 g and 2.7 g respectively, what mass of oxygen gas was consumed? (Assume there were no other reactants of products).
3.8 g

4. Sweet iced tea has a density of 1.37 g/cm$^3$. When made at the local Kentucky Fried Chicken (KFC store) in a large container the tea has a mass of 35.2 pounds. How many **gallons** of tea were made at the store? Note: 1 gallon = 3.79 L, 1 pound = 0.454 kg, 1 ml = 1 cm$^3$ 3.08 gallons

5. In the lab, you dissolve several different ionic compounds into water. You determine that in every case, when an ionic compound is dissolved in water, the water conducts electricity. Which represents the most appropriate combination of scientific theory and natural law you might make to explain this observation?

| | Theory | Natural Law |
|---|---|---|
| a. | Dissolving sodium chloride, one type of ionic compound, causes water to conduct electricity. | This is because charged particles are now dissolved in the water to conduct the electricity. |
| b. | Adding sodium chloride, one type of ionic compound, causes water to conduct electricity. | This is because charged particles are now dissolved in the water to conduct the electricity. |
| c. | This is because charged particles are now dissolved in the water to conduct the electricity. | Adding sodium chloride, one type of ionic compound, causes water to conduct electricity. |
| d. | This is because charged particles are now dissolved in the water to conduct the electricity. | Dissolving any ionic compound to water causes the water to conduct electricity. |
| e. | Adding sodium chloride, one type of ionic compound, causes water to conduct electricity. | Dissolving any ionic compound to water causes the water to conduct electricity. |
| f. | Dissolving any ionic compound to water causes the water to conduct electricity. | Adding sodium chloride, one type of ionic compound, causes water to conduct electricity. |

6. How many grams are contained in 3.94 x $10^{24}$ atoms of silicon? 184

7. Which set of numbers results in an answer of 4.9 to correct number of significant figures?
   a. 2.99/0.61
   b. 12.80-7.90
   c. 3.60*1.36
   d. 2+2.9

8. Identify which of the following masses are equal to one another:
   1000 µg 100 mg 10 g 0.1 g 0.001 mg 0.1 kg
   0.1 g and 100 mg

9. Which of the following are both true about the indicated scientists?

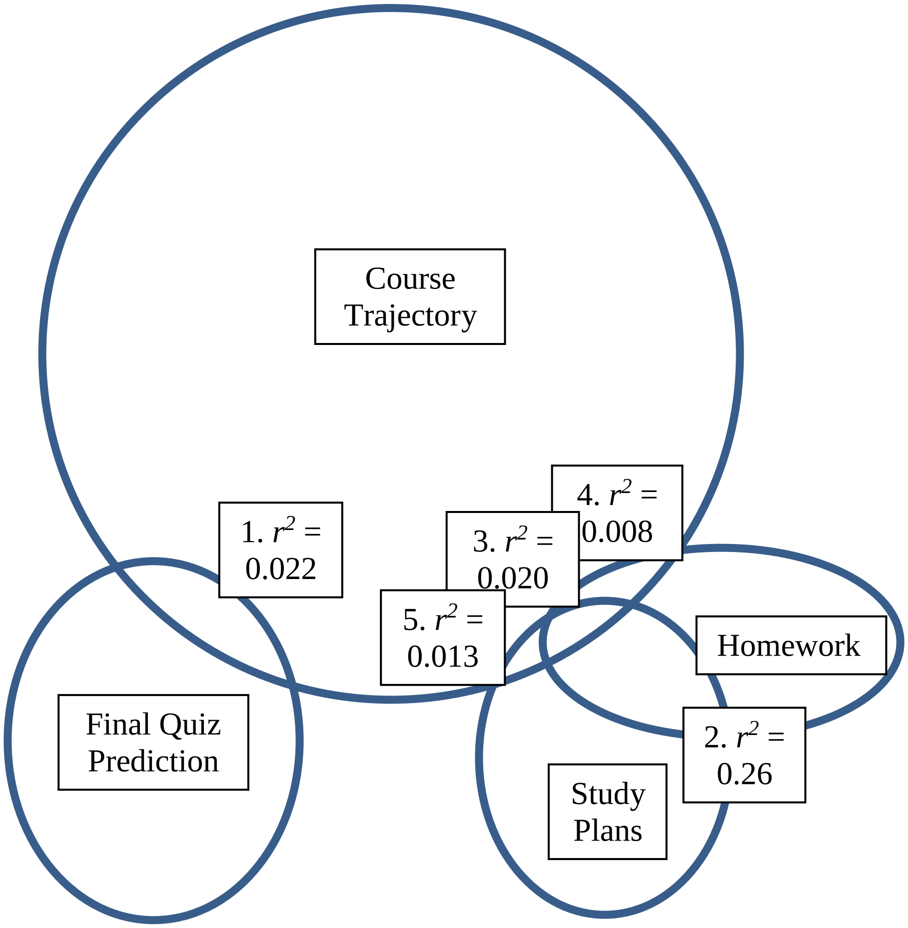
| | J.J. Thompson | Rutherford |
|---|---|---|
| a. | Described protons of an atom as being embedded in a negatively charged sphere like raisins in a ‘plum pudding.’ | Described the atom as having a small, concentrated negative charge in the center, which he called the nucleus |
| b. | Described electrons of an atom as bearing a negative charge based on how they were deflected in an electric field. | Used alpha particles to determine that the center of an atom is a concentrated positive charged area. |
| c. | Used a beam of alpha particles and observed they were deflected in an electric field, leading to a conclusion that they were positively charged. | Shot a cathode ray tube at gold foil and observed the electrons shot back, indicating that the atom is surrounded by a negatively charged electron cloud. |
| d. | Described the atom as having a small, concentrated positive charge in the center, which he called the nucleus | Described electrons as being imbedded in a positively charged sphere like raisins in a ‘plum pudding’ |

10. In your new job with a contractor you are asked to determine the cost of buying carpet to cover a convention center floor that is 3678 feet long and 807 feet wide. The carpet which your customer has chosen is made by a European company and sells for $\$4.75/\text{m}^2$. How much will the carpet cost for this convention center?
$1.31*10^6$

# APPENDIX E

## VENN DIAGRAM TO VISUALIZE THE EFFECTS ON TRAJECTORY

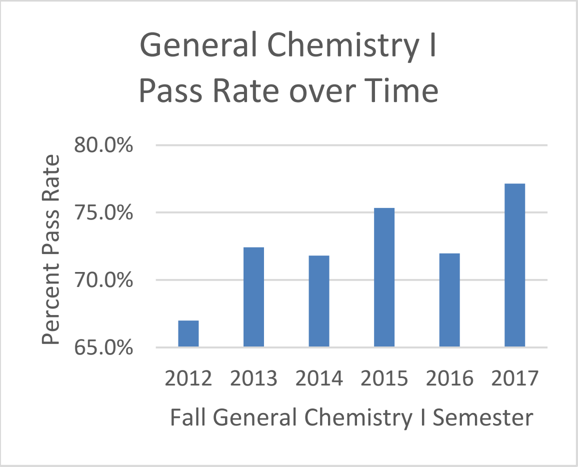
The following Venn Diagram represents the individual effects upon student course trajectory. The circles represent all possible variation within each factor, and the overlap represents the degree that each factor explains the variation: the overlap between 1: final quiz prediction and course trajectory; 2: study plans and homework, representing course effort; 3: course effort and trajectory; and the unique overlap of 4: homework and trajectory and 5: study plans and trajectory. Note:0.003 of the total $r^2$ is not below and represents the overlap between final quiz prediction and either study plans or homework.

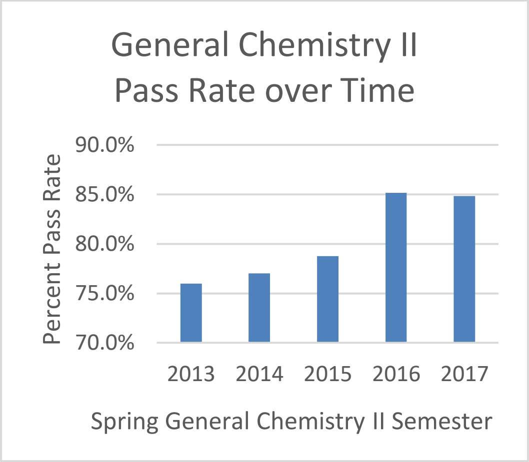


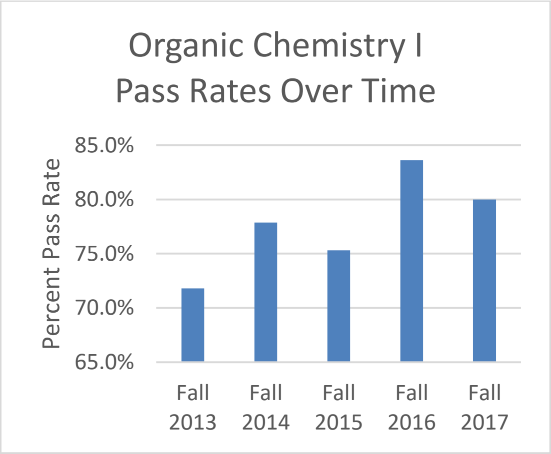
# APPENDIX F

## SIX-YEAR CHANGES TO COURSE PASS RATE

Six-year changes to general chemistry pass rates in general chemistry and five-year changes to organic chemistry pass rates. These changes include the effects of all changes to the general chemistry program under the leadership of Dr. Atwood. Chapter 8 does not consider the earliest year for each graph as this year occurred before the outset of this specific research program.







# REFERENCES

1) Tobias, S. *They're Not Dumb They're Different: Stalking the Second Tier*, Research Corporation: Tucson, AZ, 1994, pp 26.

2) Casselman, B. L.; Ohlsen, B. R.; Atwood, C. H. *Quim. Nova*, **2017**, *40*, 456−464.

3) Casselman, B. L., Atwood, C. H., *J. Chem. Ed.* **2017**, *94*, 1811-1821.

4) Coley, N. R. *J. Chem. Ed*., **1973**, *50*, 613-615.

5) Ozsogomonyan, A.; Loftus, D. *J. Chem. Ed*. **1979**, *56*, 173-175.

6) Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*; Lawrence Erlbaum Associates, Inc.: New York, NY, 1977.

7) Christiansen, M. A. *J. Chem. Ed.* **2014**, *91*, 1845−1850.

8) Ryan, M. D., Reid, S. A. *J. Chem. Ed.* **2016**, *93*, 13-23.

9) Weaver, G. C.; Sturtevant, H. G. *J. Chem. Ed.* **2015**, *92*, 1437−1448.

10) Shattuck, J. C. *J. Chem. Ed.* **2016**, *93,* 1984-1992.

11) Hibbard, L.; Sung, S.; Wells, B. *J. Chem. Ed.* **2016**, *93*, 24-30.

12) Mooring, S. R.; Mitchell, C. E.; Burrows, N. L. *J. Chem. Ed.* **2016**, *93*, 1972-1983.

13) Cook, E.; Kennedy, E.; McGuire, S. Y. *J. Chem. Ed*. **2013**, *90*, 961-967.

14) Zumdahl, S. S., Zumdahl, S. A. *Chemistry: An Atoms First Approach*, 1st ed.; Cengage Learning: Boston, 2011.

15) Madra Learning. https://www.madralearning.com/ (accessed Feb 2018).

16) Dunlosky, J.; Metcalfe, J. *Metacognition*, Sage Publications: Thousand Oaks, CA, 2009.

17) Rickey, D.; Stacy, A. M. *J. Chem. Ed*., **2000**, *77*, 915-920.

18) Kruger, J.; Dunning, D. *Journal of Personality and Social Psychology*, **1999**, *77*, 1121-1134

19) Pazicni, S.; Bauer, C. F. *Chem. Ed. Res. Pract.* **2014**, *15*, 24-34.

20) Hawker, M. J.; Dysleski, L.; Rickey, D. *J. Chem. Ed*. **2016**, *93*, 832−840.

21) Parker Siburt, C. J.; Bissell, A. N.; MacPhail, R. A. *J. Chem. Ed.* **2011**, 88, 1489−1495.

22) Solaz-Portoles, J. J. *Quim. Nova*. **2010**, 33, 1403−1407.

23) Cooper, M. M.; Sandi-Urena, S. *J. Chem. Ed*. **2009**, 86, 240−245.

24) Cooper, M. M.; Sandi-Urena, S.; Stevens, R. *J. Chem. Ed. Res. Pract.* **2008**, 9, 18−24.

25) Knaus, K. J.; Murphy, K. L.; Holme, T. A. *J. Chem. Ed*. **2009**, *86*, 827-832.

26) Pyburn, D. T.; Pazicni, S.; Benassi, V. A.; Tappin, E. M. *J. Chem. Ed*. **2014**, *91*, 2045-2057.

27) Koriat, A.; Sheffer, L.; Ma'ayan, H. *J. Exp. Psy.* **2002**, *131*, 147-162.

28) Lichtenstein, S.; Fischhoff, B. *Org. Beh. Hum. Perf.* **1980**, *26*, 149-171.

29) Nietfeld, J. L.; Cao, L.; Osborne, J. W. *Metacog. Learn.* **2006**, *1*, 159-179.

30) Cohen, J.; Cohen, P.; West, S. G.; Aiken, L. S. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, 3rd ed.; Routledge: New York, 2003, pp 355-357.

31) Cook, E.; Kennedy, E.; McGuire, S. Y. *J. Chem. Ed*. **2013**, *90*, 961-967.

32) MATLAB 8.0 and Statistics Toolbox 8.1, The MathWorks, Inc., Natick, Massachusetts, United States.