Analysis of the Relationships among Goal Orientation, Error Orientation, Online Homework Behaviours, and Learning in Organic Chemistry

by

Jaclyn J. Stewart

M.Sc. (Wood Science), University of British Columbia, 2005
B.Sc. (Chemistry), University of British Columbia, 2002

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the
Educational Psychology Program
Faculty of Education

Simon Fraser University

Summer 2014

Approval

Name: **Jaclyn J. Stewart**

Degree: Doctor of Philosophy

Title: ***Analysis of the Relationships among Goal Orientation, Error Orientation, Online Homework Behaviours, and Learning in Organic Chemistry***

Examining Committee: **Chair:** Elina Birmingham
Assistant Professor

Philip Winne
Senior Supervisor
Professor

John Nesbit
Supervisor
Professor

Kevin O'Neill
Internal/External Examiner
Associate Professor
Faculty of Education

Michelle Richards-Babb
External Examiner
Associate Professor
Department of Chemistry
West Virginia University

Date Defended/Approved: July 23, 2014

Ethics Statement

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

a. human research ethics approval from the Simon Fraser University Office of Research Ethics,

or

b. advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University;

or has conducted the research

c. as a co-investigator, collaborator or research assistant in a research project approved in advance,

or

d. as a member of a course approved in advance for minimal risk human research, by the Office of Research Ethics.

A copy of the approval letter has been filed at the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Simon Fraser University Library
Burnaby, British Columbia, Canada

update Spring 2010

Abstract

I explored goal orientation, error orientation, organic chemistry achievement, problem-solving confidence, and online homework behaviours among university-level organic chemistry learners using online homework software. Goal orientation is a well-studied, multi-construct theory explaining people's specific reasons for engaging in learning behaviours. The concept of error orientation, which I adopted from the organizational behaviour domain, describes one's attitudes and beliefs about making errors. Since students learning to solve organic chemistry problems routinely make errors, I explored if students could be characterized by their goal and error orientations to predict how they interact with an online homework system. Prior research shows postsecondary learners most strongly endorse task-approach and task-avoidance goals, although goal orientations most related to achievement were task-approach and other-approach goals. Students reported having highest levels of learning from errors, error communication, and thinking about errors. Thinking about errors had the strongest relationship with achievement. Learning from errors, error risk taking, and thinking about errors had small, but detectable, positive correlations with task-approach goals.

As is common in undergraduate science courses, students in this study practiced problem solving using online homework software that provided immediate feedback. Information about students' hint viewing, giving up, question attempts, and question scores were mined from the online homework database. Using these data, I constructed a measure of learning from errors to investigate how students' goal orientation and error orientation relate to online learning behaviours. Behaviours differed considerably between high- and low-performing students. Viewing penalty-free hints was not related to achievement, confidence, or number of attempts. Cluster analysis grouped learners by behaviours, and clusters differed in achievement.

These findings could be used to better customize online learning environments for learners with different profiles. Such customizations could improve learning, which, in turn, could enhance students' experience in organic chemistry and improve attitude toward science in general.

Keywords: goal orientation; error orientation; organic chemistry; online homework; confidence

This dissertation is dedicated to my students,
who I now understand a little bit better.

Acknowledgements

I would like to express my deepest appreciation for my dissertation supervisor, Dr. Philip Winne. Phil, thank you for your patience and advice. Our lengthy discussions taught me how to consider the theoretical, practical, and empirical aspects of education research. Your questions and suggestions helped me produce work of which I am proud. You expertly guided me along the path I defined for myself. Every meeting taught me something, showed me what I do not know, and increased my confidence. Everything seemed doable after our meetings!

I would also like to thank my committee member, Dr. John Nesbit, for helping to shape my research plan and providing feedback on this dissertation. I had positive experiences in all of my doctoral courses and the faculty members I interacted with at Simon Fraser University were deeply invested in my growth.

Sapling Learning graciously provided some of the detailed data I used in this study. Thank you Dr. James Caras, Josh Biberdorf, and Dr. Phil Volkers for your support and time during the past several years.

I am grateful for my colleagues in the Department of Chemistry at the University of British Columbia. It is motivating to be surrounded by colleagues who place a very high value and great importance on teaching. I would like to thank the community of educators in the Faculty of Science for their support and encouragement.

To my family, thank you for teaching me to value education and for supporting me unconditionally. I am lucky to have grown up in an environment in which attending university was not only an option, but highly encouraged.

Mark, I am incredibly lucky to have you in my life. Thank you for your endless encouragement, and for parsing my xml data. If you can put up with me during the last few years of a doctoral degree, I think it will be smooth sailing from here!

Table of Contents

List of Tables

List of Figures

List of Acronyms

Term	initial components of the term
AGQ	achievement goal questionnaire
CFA	confirmatory factor analysis
DF	degrees of freedom
EM	expectation-maximization
EOQ	error orientation questionnaire
FIPPA	Freedom of Information and Privacy Act
HSD	honest significant difference
GFI	goodness-of-fit index
RMSEA	root mean square error of approximation
STEM	science, technology, engineering, and mathematics
TLI	Tucker-Lewis index
UBC	The University of British Columbia

Chapter 1.

Introduction

Organic chemistry is the study of the structure and reactivity of organic compounds, which are molecules that include the element carbon. Gasoline, enzymes, and caffeine are examples of organic compounds. Professional organic chemists work in many industries; they synthesize new compounds in the pharmaceutical industry and they develop new consumer products such as lens coatings and cosmetics ("Different Types of Chemistry," 2009). Successful professional organic chemists apply creativity, knowledge, and critical thinking to solve problems (American Chemical Society, 2014).

Typically, organic chemistry is taught thoroughly for the first time at the second-year undergraduate level, although students can be exposed to the subject as early as high school. For university students who are not chemistry majors, organic chemistry serves as an important building block for other disciplines, such as pharmacy, medicine, nutrition, and biology. Several fundamental concepts from organic chemistry are important to these fields, including the acid-base behaviour of chemical species, the three-dimensional nature of chemical structure, the relationship of molecular structure to material properties, and chemical reactivity and stability. A specific example is protein folding, which is driven by interactions between parts of the protein (e.g. hydrophobic interactions) and creates a particular three-dimensional shape that has a specific biological function. Just like the underpinnings of chemistry are physics concepts, the underpinnings of many biology phenomena are chemical concepts and are thus important for a wide variety of university learners.

Many believe that learning and teaching organic chemistry is uniquely challenging (Lafarge, Morge, & Méheut, 2014). Seymour and Hewitt (1997) argue that organic chemistry acts as a "filter" in science because of its difficulty. Organic chemistry courses often have a poor reputation on campuses because they are perceived as “weeder” courses designed to fail students (Bradley, Ulrich, & Jones, 2002; Mahal, 2013). Negative experiences in chemistry courses also demotivate students from pursuing entrance to medical school (Barr, Matsui, Wanat, & Gonzalez, 2010). When students have difficulty succeeding in organic chemistry, they may become disinterested in science careers (Seymour & Hewitt, 1997). Attracting students to and retaining students in science, technology, engineering, and mathematics (STEM) fields is important for continued innovation, research, and development. An equally important goal is increasing the level of scientific literacy in the general population to encourage data-driven decision-making. Since organic chemistry is a foundational subject for several disciplines and careers, improvement in organic chemistry instruction is likely to help many students and have a positive impact on STEM fields and society.

Some challenges in teaching and learning organic chemistry are shared with other disciplines. For instance, organic chemistry learners may not know or apply appropriate study strategies (Szu et al., 2011). In the cognitive domain, learners need to link new information to ideas they already know to create well-organized knowledge structures that can be accessed for problem solving (Chi, Feltovich, & Glaser, 1981; Gabel, 1999). When learners memorize disconnected facts, it is less likely they will view the topic as experts do, as a system of knowledge built on and organized by models (Lafarge et al., 2014).

Many science disciplines challenge learners to differentiate and work with three different “levels” or “representations” of matter, described as the macroscopic, microscopic/atomic, and symbolic (Gabel, Samuel, & Hunn, 1987; Johnstone, 1991; 1993). Developing an understanding of abstract microscopic concepts through the lenses of symbolic representations and macroscopic observations is challenging for students because they only indirectly experience the concepts (Gabel, 1999) and educators frequently move inexplicably from one representation to the other (Johnstone, 1991). An example is boiling water, described symbolically as $H_2O(l) + \text{heat} \rightarrow H_2O(g)$, macroscopically viewed as steam and bubbles in the water, and microscopically imagined as water molecules moving farther apart from one another. To facilitate connections between the three levels, some posit chemistry examples should be based on the physical world (Gabel, 1999). However, this is not possible because organic chemical reactions, and most other concepts, are not directly observable in everyday life.

Organic chemistry educators have identified specific difficult concepts in a study by Duis (2011), and the list included reaction mechanisms, acid-base chemistry, synthesis, stereochemistry, and resonance. Chemistry education research has uncovered that many students have misconceptions, which are persistent and sometimes result from traditional, explanation-based instructional methods (Gabel, 1999). In the case of boiling water, many students have a misconception that causes them to confuse physical (phase) change and chemical change (forming different chemicals) (Kind, 2004). When asked what is inside the bubbles of boiling water, 20% percent of new chemistry graduate students at a major university reported they contain air or oxygen, and 5% reported they contained a mixture of oxygen and hydrogen gas instead of steam, i.e. gaseous water (Bodner, 1991). Instructional recommendations to avoid or overcome misconceptions include providing examples and non-examples to help students view the features of a particular concept (Henderleiter, Smart, & Anderson, 2001). The language used to discuss chemistry and the tendency of instructors to try to simplify concepts can also lead to misconceptions when learners construct their understanding of a concept (Bodner, 1991).

Studies have shown that several cognitive and non-cognitive variables are related to performance in organic chemistry. Like in other disciplines, prior knowledge is a significant predictor of learning in organic chemistry (Lawson, 1983; Seery, 2009). One study found that up to 45% of the variance in organic chemistry performance could be accounted for by performance in the prerequisite general chemistry course (Rixse & Pickering, 1985). This suggests that the characteristics of general chemistry learners, including prior knowledge, motivation, spatial ability, and mastery learning orientation, lead to different levels of achievement in later organic chemistry classes. A student who passes general chemistry with a low grade has a weak understanding to begin with, and may have difficulty incorporating new concepts. There is often a break of four or more months between the end of general chemistry and the start of organic chemistry, over which time considerable forgetting may occur. However, even if some of this knowledge is forgotten it may be relearned relatively easily (Arzi, Ben-Zvi, & Ganiel, 1986).

Organic chemistry requires students to visualize structures in three-dimensions, as well as to consider the orientation and movement of these structures in space. Spatial ability (visualization in three dimensions) has been found to correlate to organic chemistry achievement (Pribyl & Bodner, 1987) and spatial ability training has been found to improve organic chemistry achievement (Small & Morton, 1983). However, some studies have failed to find a relationship between spatial ability and specific topics in organic chemistry (Krylova, 1997; Turner & Lindsay, 2003). Training students to use physical or electronic molecular models has been shown to enhance their ability to make connections between chemical representations (H. K. Wu, Krajcik, & Soloway, 2001) and deal with abstract structural properties (Copolo & Hounshell, 1995).

The non-cognitive variables of attitude, perception of usefulness, confidence, interest, motivation, anxiety, and use of learning strategies have been studied to determine their effect on learning organic chemistry. These constructs may influence organic chemistry learning independently from cognitive variables (Turner & Lindsay, 2003). Using multivariate analysis, Garcia and colleagues (1993) found that prior achievement, motivation, and use of learning strategies significantly predicted student achievement in organic chemistry and overshadowed gender and ethnic differences. In a study examining gender differences in cognitive and non-cognitive predictors of organic chemistry achievement, 39% of the variance in organic chemistry achievement was explained by second-semester general chemistry grade and ACT math score (Turner & Lindsay, 2003). Other ACT subscores, spatial ability, confidence, anxiety, usefulness, and effectance motivation (desire for challenges) did not further explain the variance in organic chemistry performance. In another class, the researchers found that effectance motivation, ACT math score, and general chemistry grade explained 55% of the variance in organic chemistry achievement. Of the non-cognitive variables studied, confidence independently accounted for 9%−26% of variance in organic chemistry achievement. These predictors were found to be mostly stronger for men than women.

In Turner's 2003 study, confidence was defined as "students' trust in their abilities to learn and perform well on tasks in chemistry" (p. 564), which is actually more similar to self-efficacy, the "beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (Bandura, 1997, p. 7). I will expand on this research and examine how confidence relates to online homework behaviours and learning. Specifically, I will explore confidence as a “task-specific metacognitive experience” (Stankov, Lee, Luo, & Hogan, 2012, p. 747), and is how certain a person is that their response to a task is correct. Recent research has shown that confidence may behave as a general trait and that it is a better predictor of math and English achievement than self-efficacy (Stankov et al., 2012).

Using the Motivated Strategies Toward Learning Questionnaire, Lynch and Trujillo (2011) identified gender differences in the associations between intrinsic motivation and performance (positive association in males), extrinsic motivation and performance (negative association in females). Males more highly valued the learning tasks and had higher self-efficacy. The only significant correlations between learning strategies and performance were found in females in one of the two classes in the study. A negative correlation was found between time spent studying and grade and a positive correlation was found between effort regulation and grade.

The relationship between achievement and mastery goals of pre-medical students was explored to see if norm-referenced exam grading had consequences on students’ goals and learning (Sommet, Pulfrey, & Butera, 2013). The researchers showed, through observation and experimentation, that norm-referenced grading negatively impacts self-efficacy, which resulted in reduced learning (Sommet et al., 2013). Although organic chemistry learners have both mastery and extrinsic goal orientations (Horowitz, 2009), we do not yet know how goal orientation is related to achievement in organic chemistry learners.

Chemistry instructional methods are being developed that target students’ cognition, metacognition, and motivation (Crippen, Schraw, & Brooks, 2005). Classroom-based pedagogies such as peer-led team learning (PLTL) (Tien, Roth, & Kampmeier, 2002), in which senior students facilitate small-group problem solving, aim to help students better construct knowledge. Process-oriented guided inquiry learning (POGIL) is a form of inductive teaching in which small groups of students assigned to roles work through gradually more abstract and challenging problems (“Process Oriented Guided Inquiry,” n.d.). Chemistry pedagogues have embraced social-constructivist theories of learning and attempt to incorporate small group interactions in the classroom whenever possible.

Outside the classroom, educators are providing support to students before and after class. Paper-based and online homework are commonly intended to encourage students to practice skills and develop understanding of concepts. Since practice and feedback play such an important role in learning, it follows that using a computer system that can provide immediate, specific, detailed feedback may enhance learning over traditional problem sets with solutions. Learners are often unable to accurately judge what they know (Hacker, 2008; Dunlosky & Lipko, 2007) and it is possible that working with a software system could improve their calibration.

Computer-assisted practice and feedback is not new. A symposium report from 1986 suggested that "computer education materials" could provide students with the opportunity to "explore variations on the theme" presented in lecture (Kingsbury, 1986). The argument given for computer-assisted practice was to reduce time-pressure in exams because students would be able to solve exercises more quickly (Kingsbury, 1986). For many decades, the main motivation for computer-assisted practice has been to provide students with personalized feedback, which is especially challenging in large-enrollment courses (Chamala, Ciochina, & Grossman, 2006; Penn, Nedeff, & Gozdzik, 2000). Software programs that provide chemistry practice and feedback to students are commonly used in post-secondary education. Instructors use them to motivate students to practice and to provide them with more personalized feedback than an instructor could provide him or herself.

The purpose of feedback is to "reduce discrepancies between current understandings/performance and a desired goal" (Hattie & Timperley, 2007). Feedback may work at the task level, process level, self-regulation level, or self level (Hattie & Timperley, 2007). Most feedback in online homework systems used in science education is aimed at the task level. It identifies correct/incorrect content and offers suggestions of what to fix in the student's response. Some feedback in this channel is aimed at the process level in the form of suggestions for what the learner can do to improve answers, such as re-reading an appropriate section of the textbook. Task-level feedback is only effective if processing or self-regulation are enhanced; feedback about self does not enhance performance (Hattie & Timperley, 2007).

Feedback may change a learner's affect and cognitive processes (Hattie & Timperley, 2007). Through meta-analysis, average effect sizes of –0.04 to 1.10 have been found for various types of feedback (Hattie & Timperley, 2007). Four meta-analyses of computer-assisted instructional feedback of 161 studies found an average effect size of 0.52 (Hattie & Timperley, 2007). Generally, specific task feedback is more effective than feedback providing praise or punishment (Hattie & Timperley, 2007). Feedback in the form of extrinsic awards is negatively correlated with motivation (Deci, Koestner, & Ryan, 1999). Researchers suggest this is because extrinsic rewards reduce self-motivation and/or self-regulation (Deci et al., 1999). Despite the potential for feedback to help students learn, some may not attend to it enough to improve their performance (A. Mason & Singh, 2010a).

One of the biggest differences between engaging with homework online or on paper is that, online, students can incorporate feedback and immediately try a question again. This allows students to self-regulate their learning, update their perception of task criteria and engage in conceptual change. These processes may affect students' self-efficacy and motivation. Some students may not engage in these processes. In a study of a general chemistry class, up to 35% of students completing online homework admit to guessing, which indicates they did not process the questions and/or feedback deeply (Richards-Babb, Drelick, Henry, & Robertson-Honecker, 2011). I suspect that there are cognitive and metacognitive variables that relate to the extent students learn from errors when practicing problem solving.

Quasi-experiments have indicated that organic chemistry students doing their homework online perform better on exams than students doing their homework on paper (Penn et al., 2000; Richards-Babb et al., 2011). In physics, a quasi-experiment compared online and paper-based homework and yielded non-statistically detectable differences on course tests and concept tests (Bonham, Beichner, & Deardorff, 2001). The researchers also found that students in the online homework treatment reported spending more time on task. The authors concluded that students did not learn less with online homework than paper-based homework, and that students responded positively to online homework. Comparisons between online homework and other forms of practice are only useful for convincing practitioners that their use does not negatively affect student learning outcomes. The results can be used to reassure educators that if they were to offer online homework to their students they would end up with similar results than if they had only used paper-based homework. However, these types of studies do not provide information about which specific student behaviours relate to learning. This information is essential to designing better programs and to capitalize on the potential of online homework software.

Research Overview

Since post-secondary institutions are increasing their resolve to strategically employ educational technology, we need to go beyond quasi-experiments and develop a deeper understanding of how students interact with educational software. In this chapter and in Chapter 2, I describe several shortcomings of the existing literature about online homework in science education. For example, despite the availability and importance of feedback, it is not clear that all students learn in the same way from computer-provided feedback during problem solving. Since online homework is ubiquitous in science education, instructional decisions should be informed by research that examines how individual differences and goal states affect student learning from these software programs. In this research, I investigate how students with various goal orientations and error orientations learn from their mistakes via computer-provided feedback. I use education data mining techniques (Romero & Ventura, 2010) to explore students' interactions with an organic chemistry online homework program. I characterize the extent to which students learn from their errors during online practice, investigate relationships between achievement and confidence, and explore associations between two psychological constructs.

The first psychological construct, goal orientation, is highly-researched, often using a form of the Achievement Goal Orientation Questionnaire. Achievement goals are "the purpose for which a person engages in achievement behaviour" (Elliot & Thrash, 2001) as opposed to the objective or goal a person is trying to meet (Anderman & Maehr, 1994). The latest version of the Achievement Goal Orientation Questionnaire is a 3 × 2 model, which refer to two valences (approaching success and avoidance of failure) and three referents (task, self, and other) (Elliot, Murayama, & Pekrun, 2011). Since learning organic chemistry requires practicing exercises and problems and achievement is typically measured with traditional exams with many questions, the 3 × 2 measure has items that relate to how learners might think about their motivational goals at the question level. It is plausible that organic chemistry learners differ in their goals for successfully completing a task, and in the extent to which they want to succeed compared to their own previous achievement or that of others. Thus, I determine the extent to which organic chemistry learners endorse the various achievement goals and how these levels relate to online homework behaviours and achievement.

The other psychological construct I employ is error orientation, which is adapted from the organizational behaviour domain, and is measured using the Error Orientation Questionnaire (Rybowiak, Garst, Frese, & Batinic, 1999). I selected this tool since it is a self-report measure that attempts to directly quantify several components of students' views toward errors, namely error competence, learning from errors, error risk taking, error strain, error anticipation, covering up errors, error communication, error motivation, and thinking about errors. When learning organic chemistry by practicing on paper or online, students make many errors that ideally help them learn and guide their future learning. No previous studies have investigated the connection between learning error orientation and achievement, and these results will add to our understanding of how various learners develop their problem solving skill using software. This will also lay a foundation for future instructional interventions that could manipulate error orientation.

I collected data in a large organic chemistry class that used online homework as a method of providing feedback. Specifically, I examined behaviours including question attempts, hint use, and question abandonment. Students and questions range in the distribution of attempts and hints used. Students displayed maladaptive learning behaviours such as giving up. The extent that motivational characteristics (goal and error orientation) impact learning behaviours will be used to describe individuals and types of learners. Future studies could explore the optimal way to situation these types of learners in online homework environments.

Research Questions and Hypotheses

1. How do organic chemistry students' levels of task, self, and other-approach and avoidance goals relate to achievement, online homework behaviours, and confidence?

- Do the data from the 3 × 2 Achievement Goal Questionnaire fit the responses from these organic chemistry students?
- Are students' endorsements of achievement goals influenced by major course events, such as a midterm examination?
- Since many participants are pre-medical students, I hypothesize that they will have high levels of task- and other-approach orientations, and that self-approach will have the strongest association with achievement.

2. What levels of error competence, learning from errors, error risk taking, error strain, error anticipation, covering up errors, error communication, error motivation, and thinking about errors do organic chemistry students possess, and how do these relate to learning?

- What evidence is there of validity for inferring students' view of errors as measured by the adapted Error Orientation Questionnaire?
- Are students' views toward errors influenced by major course events, such as a midterm examination?
- Do higher-achieving students have higher levels of learning from errors and thinking about errors?
- Do lower-achieving students have higher levels of error strain and covering up errors?
- What relationships exist between achievement goals and error orientation?

3. How do students approach their online homework, and what is the impact of this on learning?

- What is the extent of the variation in behaviours for different questions, and what are the characteristics of individual questions?
- Are students who are more confident in their ability more certain of their performance predictions?
- Do students who make higher confidence judgments during online homework sessions achieve higher examination grades than those who make lower judgments?
- How do organic chemistry students varying in task, self, and other-approach and avoidance goals learn from the errors they make while solving online homework problems?
- Do students with higher levels of performance goal orientation make use of hints more often?
- Are students who view errors as valuable to learning less likely to give up or view hints, compared to students who do not see errors as valuable for learning?

- How do goal orientation and error orientation collectively impact online homework learning and achievement?
- Are there distinct profiles of learners based on how they interact with their online homework?
- How do these groups of students perform on course examinations and online homework?
- How confident and certain are these learners, and to what extent do they learn from online homework errors?
- What profiles of students exist with respect to achievement, confidence, and learning from errors?

Chapter 2.

Literature Review

In this chapter, I briefly review the literature about online homework in science education, goal orientation, error orientation, problem solving confidence, and achievement in organic chemistry. I also summarize the literature at the intersection of these constructs and learning processes. I have narrowed the scope of this review to research with adults unless otherwise indicated.

Online Homework in Post-Secondary Science Education

In post-secondary science education, online homework refers to students' interaction with software that facilitates drill-and-practice or presents scaffolded tutorials. In this dissertation, I define online homework as question-and-answer tasks that students work on outside of class time, which are presented through web-based software. Homework on computers could be done offline, but this is rare because of the logistical advantages of storing and communicating information about students' interactions with the software through the Internet. Online homework does not refer to educational technologies such as online discussions, presentation of course material in written or video form, simulations and animations, or computerized testing. In contrast, I define paper-based homework as students working exercises and problems from textbooks or problem sets provided by their instructor. The problems may be distributed online (i.e. as a file downloaded from the course website), but since student responses are not submitted to a software program for checking, they are not considered online homework. In the science education literature over the last two decades, paper-based homework is often referred to as "traditional homework".

Instructors who offer students online homework do so to enhance the number and quality of opportunities for practice and immediate feedback. In skills-based courses such as organic chemistry, practice is critical to learning. Online homework may encourage students to take “more responsibility for their learning” (Cole & Todd, 2003), enhance their engagement (M. B. Butler & Zerr, 2005), and reduce faculty grading load (Morrissey, Kashy, & Tsai, 1995). In many cases, faculty who teach hundreds of students would simply not be able to grade the homework assignments for each student. In this dissertation, participants completed low-stakes online practice for their graded homework. Ungraded paper-based problem sets with solutions were also offered for students to work on outside of class. Homework is different from quizzes or tests as it is distinctly formative in nature. Due to the formative nature of online homework, feedback is usually provided after each question is attempted, but feedback could be reserved for after a set of questions. The online homework setting can be considered a learning environment in which student characteristics, technology characteristics, and the education context interact over time (Khanlarian & Singh, 2013). Online assignments are often used as part of a “blended learning” approach to university course design, which combines face-to-face and online learning activities.

In many cases, the homework tasks could be identical in online or paper homework settings. That is, students could be asked to calculate a value and either write it down or submit it in a textbox in the software. In chemistry, students could be asked to draw a molecule when given its systematic name. If the software presents the same information to students, one would not expect there to be large differences in learning outcomes. In paper-based homework students may receive feedback from a grader, or evaluate their own work with solutions provided by their instructor. They could work a problem multiple times, but the only feedback given is usually the correct solution. It is possible that an instructor or tutor provides feedback that allows the student to correct errors or move toward the correction solution, as in the tutoring literature. In contrast, software can respond in real-time to student responses and requests. For example, software can be built to provide hints and immediate cognitive or motivational feedback. Since immediate feedback allows students to find out their task success right away, they could attempt questions multiple times. It is these features that account for learning differences between paper-based and online homework. Unfortunately, much of the science education literature comparing paper-based to online homework does not systematically attempt to attribute the learning differences to specific design features.

Learning goals in post-secondary science courses span higher-order and lower-order cognitive skills, affective change, and learning-to-learn. Deliberate practice is effective for developing expert-like thinking in many disciplines (Ericsson, 2006), and practice with cognitive and motivational coaching from a tutor can have a “two-sigma” effect compared to conventional classroom instruction (Bloom, 1984; Lepper & Woolverton, 2002). However, a more recent analysis suggests this effect may be as low as .79 sigma (VanLehn, 2011). Current research aims to understand the software-learner interactions that produce the most learning. Showing great potential are cognitive tutors that use cognitive models representing the intended competence instruction aims to achieve (J. R. Anderson, Corbett, & Koedinger, 1995). Certainly, software programs have the potential to cue students’ attention, self-regulation, and conceptual change (Cheng, Thacker, Cardenas, & Crouch, 2004). Online homework has the potential to be more effective than paper-based homework, which has been shown to have a positive effect on learning (Cooper, Robinson, & Patall, 2006). However, none of the organic chemistry software programs currently on the market allow students to experience gradually increasingly challenging tasks with personal, detailed feedback that mimics tutoring or deliberate practice. They are best described as “drill-and-practice” programs that replace the need for a human grader.

Online homework programs in the sciences typically include a variety of exercises and problems, feedback on correct and incorrect responses, hints, and tutorials. Student responses may be text, numerical, or multiple-choice. More detailed, longer, textual answers make grading and providing feedback difficult, but innovations aim to automate grading of textual responses (Litherland, Carmichael, & Martínez-García, 2013). Chemists commonly communicate with two-dimensional drawings that represent the three-dimensional nature of molecules. Thus, organic chemistry poses an additional challenge to online homework developers and instructional designers: an input based on two-dimensional chemical drawings is essential for students to respond as they would on an exam. Most chemistry software programs before the early 2000s allowed numerical and text-based entry but did not have drawing capabilities for chemistry (Chamala et al., 2006). Some of the first online homework systems were simple web pages that allowed students to respond to multiple-choice questions by selecting a radio button (Hall, Butler, & Kestner, 1999). The first program to use a chemical drawing tool as input on the instructor side was WE_LEARN (Penn et al., 2000). Following that, programs that allowed students to respond with drawings were developed by textbook publishers and educators, such as EPOCH/ACE, Synthesis Explorer, and Sapling Learning (Parker & Loudon, 2013). Some platforms are chemistry-specific, while others offer content spanning a range of disciplines. Also, at least one open-source software program exists, called WeBWorK. Participants in this dissertation research used Sapling Learning, a proprietary system created by a company that has been acquired by Macmillan New Ventures.

Effect of Online Homework on Learning

The articles that compare paper-based and online homework lack detail about learning objectives tested and specific homework tasks. Many did not use the same tasks in the online homework as in the paper-based homework, nor did they indicate the extent to which tasks were aligned with course goals, learning activities and other assessments. This literature does not provide evidence of a clear benefit for online homework over paper-based homework. Few studies have attempted to discover what, if any, unique contribution *online* homework makes to student learning. As Clark (1994) argues, many studies that compare computer-based instruction to classroom instruction fail to properly separate out the effects of instructional method and the features of the media. This is common in the comparative studies in the chemistry education literature. Compared to in-class quizzes, the use of software has been found to detectably improve performance on course tests in general chemistry (Richards-Babb et al., 2011), in introductory accounting (Dillard-Eggers & Wooten, 2011; Khanlarian & Singh, 2013), and in college algebra (Burch & Kuo, 2010). However, many experimental and quasi-experimental studies have not detected a difference in performance in outcome measures between groups who completed online homework versus traditional homework in physics (Bonham et al., 2001; Demirci, 2010; El-Labban, 2003; Fynewever, 2008; Huesgen, 2012; D. Mason, 2009). Burch and Kuo (2010) reported that students retained more of what they learned in classes offering online homework compared to traditional homework assignments in college algebra. A study with general chemistry learners did not find a difference in test scores between groups who were given immediate feedback via web-based problem sets compared to responding to the same problem sets on paper (Cole & Todd, 2003).

When improved learning has been observed from online homework, it has been attributed to hints, feedback, and the process of reworking problems. The biggest difference between paper-based practice and online practice is the immediacy and specificity of feedback. Correlational studies indicate performance in online homework over the course of an academic term positively relates to test performance in general chemistry (Eichler & Peeples, 2013), organic chemistry (Parker & Loudon, 2013), and engineering (thermodynamics) (Taraban, Anderson, Hayes, & Sharma, 2005). Some best practices have emerged from this research, which include providing a grade incentive to motivate students to do the problems and to select programs that provide high-quality feedback. Additionally, mathematics educators have suggested using online homework in conjunction with paper-based practice, as the paper-based practice requires students to “write” mathematics using statements and graphs (Jungic, Kent, & Menz, 2012). This recommendation is adhered to in many organic chemistry classes that use online homework; students are provided paper-based problem sets and practice exams to better simulate exam conditions. If online homework does not generally make a difference to student learning, instructor time savings may still be a worthwhile benefit. Additionally, many systems make it easy for the instructor to evaluate where the class stands on certain concepts and skills, which requires more resources in a context that uses paper-based homework. Such evaluation may help the instructor be more responsive to the needs of a class. Because the preceding studies did not examine the nature of students’ interactions with the system nor measure theoretically relevant student individual differences, it is possible that working in an online homework environment enhances learning of some concepts and skills, in some types of learners.

In addition to examining learning and performance outcomes, researchers have quantified the effect of online homework on students’ effort, perceived engagement, time on task, and perception of helpfulness to learning. For example, accounting students with low need for cognition exerted more effort in online homework than those with high need for cognition (Peng, 2009). The results from this study suggest that student motivational tendencies can influence how students engage with an online homework system.

In a study that did not detect a performance difference between paper-based homework and online homework in college algebra, students reported they studied about the same amount of time with the online homework program as they would have if the homework was paper-based (Hauk & Segalla, 2005). However, Allain and Williams (2006) compared four sections of an astronomy course that used different types of homework. The four types were 1) online graded with immediate feedback, 2) paper-based ungraded homework, 3) graded online homework for the first half and ungraded for the second half, and 4) ungraded for the first half and online graded for the second half. The authors argue that the comparison between 1 and 2 is most relevant to instructors of large classes, who do not have the resources for human grading. In all four classes, homework was worth 10% of students' grades. Between the groups using online graded and paper-based ungraded homework, there was no detectable difference in scores on identical tests, but the authors found that students who used the online homework reported spending more total time outside of class than those who had paper-based homework. The group who had online homework in the second half of the course performed the best, but the group was smaller and the tests used to measure the difference were not identical.

Physics students also reported spending more time using an earlier (1996) version of a homework software program (Thoennessen & Harrison, 1996). If this leads to better retention of learning, time costs may be a benefit; otherwise, this finding suggests that learning from an online homework system may be less efficient than learning from doing paper-based homework. The learning experience in specific online and specific paper-based practice may be similar or very different, and this probably impacts the difference in time-on-task. Both of these studies asked students to report their perception of time spent after the fact, so this data may be unreliable.

Using an accounting online homework program that allowed for multiple attempts at each problem, Hall and colleagues (2001) examined the impact of multiple attempts on self-reported engagement with the material. Students who were allowed multiple attempts reported they spent more time trying to understand the course material.

Students responded to questionnaire items about their perceptions of using online homework, and more students (43%) said they spent more time studying rather than less time (16% of respondents) with online homework than paper-based homework (Dillard-Eggers & Wooten, 2011). In the same study, just over half of students perceived better quality studying in the online environment, while 15% of the respondents thought it was of lower quality. Additionally, 55% of students believed they gained a greater understanding of the course material as a result of the online homework practice. However, when asked to compare their beliefs about online and paper-based homework, almost half of the participants reported that online was better for their learning, and 31% thought it was worse (no inferential statistics were reported). The authors suggested the students with negative views might have been resistant simply because it was a change in homework style. Now that online homework is more common in universities, many students have experience with using software programs for homework, so this finding may no longer apply. In a different study with accounting students, technical efficacy was found to explain 46% of the variance in their perceptions of usefulness (Khanlarian & Singh, 2013). This finding suggests that students' backgrounds influence their perception of usefulness.

Effect of Online Homework on Behaviour

Practitioners would be concerned if student engagement and attention were lower with online homework compared paper-based homework. Many believe that university learners are easily distracted in online settings. If scores on homework tasks are taken as a proxy for engagement, as Demirci did (2007), there is some evidence that online homework may be less engaging for students. In this example, introductory physics students had lower scores on online homework than those doing paper-based homework. The end-of-term test showed no detectable performance difference between the two groups, despite the difference in homework grades.

One of the major differences between online homework and paper-based homework is that software settings make it easier for students to attempt a question multiple times. This is logistically difficult, although not impossible, in paper-based settings. In theory, learners use information provided by software programs to make decisions about how to regulate their learning. Some students' primary homework goal is not learning. In a quasi-experiment comparing two fully online courses, economics students who were allowed two attempts (without penalty) were found to be more likely to "game" the system to inflate their grades on each of the homework assignments than those who were allowed only one attempt (Rhodes & Sarbaum, 2013). In this study, the only feedback provided after the first attempt was a student's overall score and an indication of which questions were incorrect. The performance on course tests between the two classes was the same, but the researchers suggest that guessing resulted in grade inflation. For the two-attempt class the first-attempt score for the homework assignment was lower than that of the one-attempt class, indicating that students were more likely to guess in the first "pass" of the assignment when they knew they had another attempt. The researchers concluded that students did not put forth more effort into an assignment when they were given an additional attempt, but that they were able to achieve higher homework grades. Thus, there may be some negative aspects to allowing students to attempt a question multiple times, even when performance on a particular homework assignment apparently improves. In this situation, the additional attempt did not encourage students to deeply process the material and build their understanding of the course material.

Authors of a non-experimental study claimed that students studying general chemistry who were allowed to attempt a problem set a second time addressed their misconceptions, engaged in collaborative learning, spent more time on task, and corrected their errors. These claims describe the potential cognitive benefit of attempting an incorrect question until a successful response is achieved. Unfortunately, the authors did not collect information about learning, so they could only describe the positive perceptions of what would now be considered a simple web-based system (Hall et al., 2001). Additionally, two issues limited the study: students decided if they wanted to complete the "second-chance" homework (i.e., non-randomized groups), and course grades included the score from the second-chance homework. However, the finding that 90% of the class decided to complete at least one homework assignment a second time suggests that students were motivated to work through the problems a second time.

Since most undergraduate students are taking many classes, they manage competing goals. Learners may make the decision to give up on a problem or assignment in traditional and in online learning environments. Giving up does not lead a student toward meeting a specific learning goal. The relationship between goals and giving up behaviour is controlled by one's beliefs. Those who believe their hard work contributes to their success keep trying, as their effort matters (Weiner, 1994). Simon's theory of motivation describes various degrees to which a learner may reach his or her initial goal and offers many reasons for abandoning a question or assignment: they achieve their goal, they think they are close enough to the goal (i.e. they are "satisficed"), they grow impatient, or they are discouraged by failure (Simon, 1967). Khanlarian (2010) applied Simon's theory of motivation as a framework for investigating behaviour and learning and in an online homework environment used by accounting students. She identified relationships between several constructs and student performance and these relationships changed over time. A complex structural equation model depicted relationships amongst engagement, locus of control, performance goals, mastery goals, self-efficacy, usefulness, technical efficacy, lazy user, frustration, cooperative learning, GPA, perceived ability, and performance. Feedback, humanistic learning, and student-centred control were not found to be related to performance at any time during the term. Interestingly, students' online homework grades were not predictive of test scores early in the term, but became more predictive as the course progressed. The variation in engagement and behaviour will be described in the context of achievement goal orientation theory later in this chapter.

Feedback in Online Homework

Learning from homework is a complex task. Task characteristics, such as the requirement to show one's work, nature of feedback, the likelihood of reviewing feedback, and other concepts theoretically affect the types of thinking students engage in while solving problems online (Bonham, Deardorff, & Beichner, 2003). Since the biggest difference between paper-based practice and online practice is the immediacy and specificity of feedback, this is the most-studied aspect of online homework environments. Feedback varies with respect to its intentionality, delivery, target, and content (Bangert-Drowns, Kulik, Kulik, & Morgan, 1991). In the science online homework environment, feedback usually provides information about quality and correctness of a learner's response. It is intended to assist students to arrive at the correct response and offer information that will help a learner correct errors.

Most science educators would agree that increasing the amount and quality of feedback would likely have a positive effect on learning. Feedback encompasses all "actions taken by (an) external agent(s) to provide information regarding some aspect(s) of one's task performance" (Kluger & DeNisi, 1996). As a detailed review of the feedback literature is outside the scope of this chapter, I will briefly summarize the results of a few key review articles. In 1996, Kluger and DeNisi published a review and meta-analysis about feedback intervention research published before 1992. A feedback intervention includes simple knowledge of correctness (i.e. correct or incorrect), or more informative feedback about the nature of a learner's performance (e.g., one of the carbon atoms has five bonds to it). Their criticisms of feedback intervention research performed in the prior 120 years include poor methodology, weak operationalizations, and unwarranted conclusions.

An example of the issues raised by Kluger and DeNisi (1996) is present in a review by Ammons (1956). Ammons (1956) presented several examples of decreased learning due to a feedback intervention despite the main conclusion of the review being favourable toward feedback. After performing a meta-analysis, Kluger and DeNisi concluded that on average, feedback interventions improve performance with an effect size (*d*) of 0.41. However, they found significant heterogeneity and that one third of studies involving feedback interventions produced negative effects on performance, possibly because they shift the learner's attention away from the task and onto themselves (Kluger & DeNisi, 1996).

Kluger and DeNisi (1996) propose a feedback intervention (FI) theory that builds upon behaviourism, goal setting theory, control theory, social cognition theory, and learned helplessness theory. Their theory has five arguments: "(a) Behavior is regulated by comparisons of feedback to goals or standards, (b) goals or standards are organized hierarchically, (c) attention is limited and therefore only feedback-standard gaps that receive attention actively participate in behavior regulation, (d) attention is normally directed to a moderate level of the hierarchy, and (e) FIs change the locus of attention and therefore affect behavior" (Kluger & DeNisi, 1996, p. 259). Arguments a and c relate to learner's goals and goal orientations, as well as their perception of the degree to which their response constitutes an error. They simplify the organization of goals and standards to three levels – "meta-task processes involving self" (p. 262), "task-motivational processes involving the focal task" (p. 262), and "task-learning processes involving the task details of the focal task" (p. 262). Framing effects – biased responses based on negative or positive descriptions of a situation – may affect the meta-task processes, as they connect to consequences of task performance. The literature on error framing, which will be discussed later in the chapter, may connect here, as people are encouraged to make mistakes since they can be learned from. They make an excellent case for the complexity of learning from feedback, which goes beyond the idea that learners aim to reduce the discrepancy between standards and their own performance. For example, affective processes may be impacted by a shift in attention away from the task and toward self, which occurs after a learner receives feedback.

Sometimes, increased attention to feedback does not lead to better performance, several examples of this are discussed in the review by Kluger and DeNisi (1996). When feedback interrupts the implementation of an automated script, performance may be hindered (Vallacher & Wegner, 1987). Specific types of feedback, such as outcome-only feedback, may cause a learner to "experiment with successful task strategies, resulting in poorer task performance" (Hammond, Summers, & Deane, 1973). Earley, Connolly, and Ekegren (1989) proposed that experimenting with successful task strategies means that motivated learners may be cognitively inconsistent or have "dysfunctional strategy search(es)". Even process feedback does not have a clear impact on performance. External feedback of the form of prompting (giving the user the next step) followed by feedback in the form of a correction has been shown to have a detrimental effect on transfer tasks (Carroll & Kay, 1988). Introductory organic chemistry online homework tasks require near-transfer to exam questions. Perhaps when feedback is given, learners do not make knowledge connections that are required to prepare them to succeed on transfer tasks. Understandably, if the information provided in the feedback is not new information, then there will not be an effect on performance (Kluger & DeNisi, 1996). Kluger and DeNisi (1996) suggest that learning from errors (without feedback) could be better for performance than feedback that contains too much information. Thus, the heavy inclusion of feedback in chemistry online homework systems may not lead to retention and transfer. This could explain the lack of consistent positive results when comparing learning produced through online homework practice compared to paper-based practice. Since online homework software programs provide a variety of feedback types, comparing them broadly to paper-based practice is too coarse to be meaningful.

In a small meta-analysis, Bangert-Drowns, Kulik, Kulik, and Morgan(1991) found that feedback had a small but significant effect on achievement, with an average effect size of 0.26 from 58 studies. Eighteen of these studies had results where feedback hindered learning. The researchers examined moderator and mediator variables such as pre-search (peeking at feedback), error rate on first attempt at a task, type of instruction, immediacy of feedback (delayed was better than immediate), feedback type (information to help learners arrive at the correct answer was better than correct/incorrect feedback). The study methodology was limited as aspects of instruction and feedback type were correlated in the studies used in their analysis. Online homework fits the suggestions the authors provide about giving effective feedback. It controls for pre-search in that students cannot see the feedback before attempting the problem (unlike in some other types of homework such as textbook problems with solutions manuals). They also suggest that feedback is more effective when it comes after students submit a response about which they are confident and in which they have corrected prior errors. Since students have a small grade incentive for getting each response correct, they likely submit answers to questions they are somewhat to very confident about. They describe the process a learner may go through after receiving feedback about an incorrect response, which is “mindful, metacognitively driven knowledge alteration” (p. 234).

Usually, feedback in online homework is provided after each task or item is completed. This is because of the common belief that feedback has the most positive effect on learning if it is provided immediately to the learner. However, Buzhardt and Semb (2002) did not find a difference in later performance between groups who received feedback item-by-item or on the entire practice test. They also found that most of their participants, psychology undergraduate students, preferred receiving feedback on each item. In the test environment, previous research found that item-by-item feedback increased anxiety and frustration in some students. Interestingly, students had fewer questions of instructors when they were presented with item-by-item feedback. This may indicate they were more motivated or able to resolve their own questions.

Features of feedback can negatively or positively impact learning (Hattie & Timperley, 2007) by influencing students' knowledge, motivation, or emotions. Feedback can inform a student of the correctness/incorrectness of a response, guide diagnosis of misconceptions, provide suggestions on how a student could proceed, or provide examples (Vasilyeva, De Bra, & Pechenizkiy, 2008). Researchers who aim to optimize feedback provided in online homework systems are investigating the timing, function, content, context, and personalization of feedback (Vasilyeva et al., 2008). For example, some studies have found negative effects of correctness feedback in computer-assisted instruction settings (Carroll & Kay, 1988; Lepper & Gurtner, 1989).

To design more effective feedback, researchers are taking into account correctness of response, confidence in responses, and (so-called) learning styles as "adaptive feedback" (Vasilyeva et al., 2008). When "adaptive feedback" was an option to students during a quiz in a human-computer interaction course, on average students selected to view it 75-59 percent of the time. This means that students do not always take the opportunity to examine feedback specifically designed for them. Taking into account confidence level when designing feedback has the potential to avoid situations where a learner guesses correctly and does not get helpful information about his or her ability in a task. Additionally, the researchers examined the ratio of posttest score to time spent studying feedback. They found that more confident participants have a higher score to feedback time ratio. That is, confident learners outperformed less confident learners for the same time spent studying feedback.

Despite the complexity of the impact of feedback on learning, software publishers claim that students attend to and incorporate feedback into their thinking. The literature suggests that there are three factors that explain the effect of feedback on performance: a) the content of the feedback message, b) the nature of the task, and c) the situational and personality variables (Kluger & DeNisi, 1996). Thus, investigating how students' goal and error orientation impacts their learning from feedback will move the field toward the goal of understanding how students learn from feedback in online homework settings.

Practitioners' acceptance of any type of feedback as good for learning in all types of conditions may result from this conflicting body of literature or a lack of attention to the literature at all. Since one of the main reasons online homework has been adopted is to provide feedback, practitioners and researchers should critically examine the nature and interactions between tasks, feedback, motivation, and learning.

Errors in Problem Solving

An error is “an action that is not on the (shortest) path to the intended goal” (Ohlsson, 1996). Errors may prevent someone from meeting an achievable goal, or they could occur along the way to successful attainment of a goal. Errors in problem solving result from missing or inaccurate procedural or declarative knowledge (Frese & Altmann, 1989) or faulty regulatory abilities. According to Ohlsson (1996), the specific nature of errors gives insight into the knowledge structure of a learner. When learning to solve problems, people need to activate declarative knowledge and procedures and strategically transfer what they know about general strategies to a specific strategy for a task (Phye & Sanders, 1994). Problem solving requires the learner to make decisions and self-regulate as they plan and execute a solution.

While solving online homework problems, students make errors that can be classified according to correctness, methods, approaches, and misconceptions (Roth, Ivanchenko, & Record, 2008). For example, mathematics errors were classified into five categories to investigate whether students using WeBWorK were learning from their various attempts at a problem, or if they were guessing and checking. The five categories were reworking, fixing an entry error, resubmission of the equivalent, non-strategic guessing, and non-sense. Some submissions fit into more than one category. Researchers coded 32 student submissions for three different years (1^{st}, 2^{nd}, 3^{rd}) of mathematics courses, and found that the patterns were fairly similar from year-to-year. That is, reworking and resubmission were the most common submission types and non-strategic guessing was quite low (Roth et al., 2008).

The type of cognitive processing and strategies used by students solving problems can only be inferred at from their responses. Unlike with multiple-choice problems, where specific misconceptions can be targeted, free-response questions are more complicated to score. However, we can infer problem solving characteristics such as the degree of correctness, method or approach, and misconception or error type from student-constructed responses (Lie, Taylor, & Harmon, 1996). The ability to code for these aspects of problem solving has been demonstrated by the large-scale assessment TIMSS (Trends in International Mathematics and Science Study).

Learning from Errors and Feedback

Conventional wisdom suggests that people can and do learn from their mistakes, and organic chemistry educators certainly believe that practice and making mistakes are normal parts of learning to solve problems. Those who believe in the behaviourist philosophy may reserve positive reinforcement for error-free behaviour (Skinner, 1954), but educators who have a cognitive view may encourage errors to assist learners in restructuring their knowledge (Frese et al., 1991). Errors and feedback are sources of information learners can use to improve their knowledge and skill. Making an error while solving a problem may help students improve the specificity of the conditional (if-then) relationships that make up many of the concepts in introductory organic chemistry. From a self-regulation perspective, errors can help people better estimate their current abilities and motivate them to make goals and plan for future action (D. Butler & Winne, 1995). Additionally, evidence suggests that errors may encourage students to change strategies (Ivancic & Hesketh, 2000). In organic chemistry, errors can inform the learner of if and how they should work toward altering their thinking about concepts, processes, and strategies. That is, “feedback empowers active learners with strategically useful information, thus supporting self-regulation” (Bangert-Drowns et al., 1991 p. 214).

Several researchers theorize about how people learn from knowing about their errors. In the context of feedback interventions, Kluger and DeNisi (1996) discuss both Wood and Locke’s 1990 goal setting theory and Salmoni, Schmidt, and Walter’s (1984) finding that learners think about the task after receiving negative feedback. Thinking about the task may deepen cognitive processing, which increases retention. Additionally, learning from errors requires the learner to attend to specific characteristics of the task they may have originally ignored (Kluger & DeNisi, 1996).

VanLehn (1999) used data from physics student problem solving to create a model of learning complex cognitive skills composed of rule-based and analogical reasoning methods, including the impasse-repair-reflect cycle of learning used by a computational model called Cascade. The Cascade model is able to learn physics principles from textbook information and use these principles to solve problems. In the study of physics students, VanLehn uncovered that student reasoning was “shallow” and the students were not learning as much as the material contained. Like chemistry, physics content includes conceptual, mathematical, and procedural knowledge (Vanlehn, 1999), and thus it is possible that these findings would generalize to organic chemistry learners. When solving problems, learning is not assured; and “perhaps the most impressive thing…is how little students learn” (Vanlehn, 1999). VanLehn suggested that students did not learn because they may have been “glossing over” examples instead of using self-explanation, copying examples instead of reasoning with the rules they were trying to learn/apply, and not receiving feedback. VanLehn estimates that with feedback students would be more likely to have a “learning event”, and the reward structure of the activity may have been set up for short-term gains as opposed to encouraging students to learn physics for the long term with a greater degree of mastery orientation.

A learner’s decision to attend to feedback or not depends on the correctness of their response and how confident they are that their response was correct (Webb, Stock, & McCarthy, 1994). If a learner is confident and their response is correct, they are not likely to spend much time studying the feedback since their ability has been confirmed. This was supported experimentally with general information questions given to introductory psychology students. Those with high confidence who made an error studied feedback for longer than those with high confidence who did not make an error. The relationship between confidence and feedback study time was mixed for those with lower confidence (Webb et al., 1994).

Vollmeyer and Rheinberg (2005) presented psychology students with a biology task and told half of them they would receive feedback. They found that when learners knew they would receive feedback, they implemented better strategies than students who were not going to receive feedback. This suggests that the presence of feedback changes students’ approach to learning even before they start working on a task.

Research suggests that simpler feedback may be more effective than complex feedback, because more information may allow learners to successfully complete a task without having to enact strategies on their own. When studying the structure and function of the U.S. Navy, undergraduates were grouped into four treatments with varying levels of feedback complexity. With increasing complexity of feedback, the time spent studying the feedback increased, but the conditional probability of correcting an error decreased (Kulhavy, White, Topp, & Chan, 1985).

Mathematics students who used WeBWorK were found to often resubmit a response with the same error (Roth et al., 2008). This may be specific to the nature of math response formats, but the educators decided to build in a “resubmission” alert to tell students they were submitting a response with the same error. They found that this reduced the frustration ratings, as measured by questionnaires. These researchers developed a model of student responses to learn more about how they interact with the system.

Frese and colleagues (1991) discuss four beneficial aspects of errors: errors can lead to the restructuring of mental models, errors can be used to instill error-checking mechanisms, a positive view of errors encourages exploration, and people should be able to deal with errors in “real life” as well as learning situations (Frese et al., 1991). These benefits, which span cognitive, metacognitive, and affective effects, explain how practicing problem solving in an online homework environment could improve learner’s ability to solve problems. It is vitally important that students learn to monitor their problem solving, and when they submit incorrect solutions they receive feedback that could help them restructure their knowledge. Even without feedback, the incorrect attempt may encourage them to learn. Additionally, the formative nature of online homework encourages students to attempt to apply their knowledge and risk making mistakes.

Making errors may lead to negative emotions such as frustration, anger, and hopelessness because they can confront people with the likelihood of failure (Carver & Scheier, 1990). People may also have a positive view of errors, especially if they believe that it is possible to learn from them. Although errors can be frustrating for learners (Frese et al., 1991), they can lead to significant learning. Since negative emotions lower self-efficacy and can decrease motivation (Pekrun, Goetz, Titz, & Perry, 2002), interventions that attempt to change one's orientation toward errors may improve people's emotional response to errors. When working in online homework situations, systems may grade student responses to different, more stringent standards, compared to a human grader. Evidence that this frustrates students was found when El-Labban (2003) collected survey data on student perceptions of online homework. To encourage students to trust the resources made available to them for learning, the grading criteria should be consistent.

Learning from errors requires people to process the error to find its cause (Duncan & Weiss, 1979 as cited in Zhao, 2011). Compared to a training session that was error free, driving learners using a simulator in an error-rich session demonstrated greater ability to transfer and better coping strategies (Ivancic & Hesketh, 2000). In a cooperative learning intervention of naval navigation instruction, errors allowed learners to get feedback on a "need to know" basis (Seifert & Hutchins, 1992), which may have served to reduce the cognitive load experienced by the learners. This situation differed in several ways from online homework, since team members frequently discovered each other's errors in the cooperative navigation training. However, the concept that learners typically focus on one problem at a time when doing online homework suggests that they may be successful in mastering a limited set of concepts and skills.

Instructional methods to teach computer skills became important when computers started becoming more common in the workplace. The field of error training emerged in the 1980s and 1990s as researchers aimed to design efficient computer instruction (Brodbeck, Zapf, Prümper, & Frese, 1993; Frese et al., 1991; Heimbeck, Frese, Sonnentag, & Keith, 2003; Van Der Linden, Sonnentag, Frese, & Van Dyck, 2001). There is convincing evidence from this field that errors serve a positive function in learning, as a result of experiments that "framed" learners to have a positive mindset toward making mistakes. Error framing is a method of moving an individual or team to an error-accepting psychological state through error encouragement. This was typically compared to instruction that allowed learners to avoid making errors through detailed instructions.

Frese and colleagues (1991) compared the learning gains between a group provided “error training” and a group that was discouraged from making errors. The research was done with adult, non-student participants, who were being trained to use a software system. The error-training group was provided with heuristics about learning from errors, and the error-avoidant group was not allowed to struggle with mistakes. Performance was measured with free recall and an error correction tasks, transfer, and speed. Researchers also measured trainee’s perceptions of satisfaction, frustration, mood, and strain. The error-training group outperformed the error-avoidant group in free recall and competence on the most difficult tasks. The error-training group experienced much less frustration after a task that forced them to make errors than the error-avoidance group did. The researchers attributed the difference to deeper levels of processing and higher motivation resulting from the learning-from-errors heuristics. This research suggests that people can be taught that it is desirable to make errors because they are opportunities for learning (Frese et al., 1991).

Also in workplace vocational training research, error training has been compared to other forms of instruction, such as “drill and practice”. Error training encourages learning from making mistakes, as opposed to more traditional drill and practice that engages the learner in practicing the “correct” way to do things. Constructivism and action theory have influenced vocational training, and the concepts of life-long learning, self-directed learning, and transfer of learning are highly valued in this field (Kluge, Ritzmann, Burkolter, & Sauer, 2011). Individual characteristics such as cognitive style and conscientiousness have been found to interact with general mental ability to lead to differences in training outcomes between error training and drill-and-practice treatments (Kluge et al., 2011). Kluge suggests that error-training situations require more learner effort, which explains why people with lower conscientiousness performed better in a drill-and-practice situation; they were not willing to put in the effort needed.

Mason and colleagues have conducted several related studies on learning from errors with undergraduate physics students (Cohen, A. Mason, Singh, & Yerushalmi, 2008; A. Mason & Singh, 2010b; 2010a; A. Mason, Cohen, Singh, & Yerushalmi, 2009; A. Mason, Cohen, Yerushalmi, & Singh, 2008). These studies suggest that low- and high- performing students do not always learn from the errors they make on quizzes or exams (A. Mason & Singh, 2010a). In one study, honours students scored slightly worse on final exam questions that had already been presented on a midterm exam (A. Mason & Singh, 2010a). In another study, students were asked to reflect on their own problem solutions using the correct solution, a worked example, or their overall score (the overall score treatment was described as “minimal guidance”) on the problem (Cohen et al., 2008). Self-diagnosis scores were highest for students in the treatment group that received the most guidance (correct solution), and students with knowledge below a certain threshold were not able to diagnose the problems with their solutions. This line of research suggests there is wide variability in how learners approach the error detection and planning components of learning from errors. Educators should not take it for granted that university students learn from their mistakes, even when they know they have made them.

In the case when someone fails at a task, and receives negative feedback, Wood and Locke (1990) theorize that people invoke knowledge and strategies, starting with the often-successful “universal strategy”: If someone is situationally motivated, they try harder, focus their attention, and persist in trying to complete the task well. As an example, a chemistry student who has just responded incorrectly to a question may re-read the question and search their notes for additional information before trying the question again.

While error “climate” or framing has not been studied as much in academic settings as it has in workplace training, results of research in many mathematics classrooms in Germany showed that the perceived error climate in the classroom has an effect on how students react to errors. This effect is independent of classroom goal structure and achievement motivation (Steuer, Rosentritt-Brunn, & Dresel, 2013). The researchers claim that students’ self-regulation is affected by error climate. There is some additional work published in German by Oser and colleagues (Steuer et al., 2013).

In a study comparing various types of feedback provided to people learning a software task, researchers found that feedback that signaled errors could lead to negative self-evaluation and to a unsystematic exploration (Gardner & Wood, 2009). This was contrasted with corrective feedback, following which learners engaged in systematic exploration. Unsystematic exploration led to lower levels of learning than systematic exploration. These researchers suggest that feedback needs to indicate what the error is, how the error may have come about, and what could be done to prevent a similar error. These suggestions could be translated to an online homework setting if feedback is specific about how a learner's response differs from the acceptable answer(s), plus what misconception, missing information, or slip may have led to the incorrect answer, and suggest metacognitive strategies for reducing this error in the future.

To better understand how university science students learn from making mistakes while completing online homework, we should identify the relevant characteristics of learning from errors, and examine the conditional likelihood of learning from feedback in online homework. One characteristic that may explain some of the variation in learning from errors is one's error orientation. This set of constructs has been developed into the vocational training literature and can be easily adapted for use in academic settings.

Definition and Measurement of Error Orientation

Broadly, error orientation describes one's attitudes about errors (Hetzner, Gartmeier, Heid, & Gruber, 2011). Generally, people with error-mastery orientation have positive views or approaches toward errors and those with an error-aversion orientation have negative views or approaches (Van Dyck, Van Hooft, De Gilder, & Liesveld, 2010a). Error orientation is made up of several psychological trait or state variables and has been studied almost exclusively in the work domain as the "error framing" and "error handling" literature. Rybowiak, Garst, Frese, and Batinic (1999) suggests that organizations should be interested in the error orientation of their employees because it can impact workplace culture (Rybowiak et al., 1999). Organizations can use employee error orientation information to reveal opportunities for improved efficiencies and effectiveness.

The Error Orientation Questionnaire (EOQ) was developed in Germany and used in Germany and the Netherlands (Rybowiak et al., 1999) and the Philippines (Mateo, Muring, Malayao, Daniel, & Emperio, 2013). The purpose of the original version of the EOQ was to measure worker's attitudes about errors and to gather their self-reported coping mechanisms (Rybowiak et al., 1999). Rybowiak et al.'s 1999 paper describes two studies and reports a six-scale and eight-scale version of the instrument. The researchers used data from adults in Germany and university students in the Netherlands to perform confirmatory factor analysis and examine equivalence of Dutch and English versions.

Researchers developed items by considering how errors are perceived and anticipated, as well as how people cope with errors. This information is included in a dissertation about managers' description of errors (Grefe, 1994 in Rybowiak, 1997, written in German). Using exploratory and confirmatory factor analysis, including cross-validation of their model, subscales of error competence, learning from errors, error risk taking, error strain, error anticipation, and covering up errors emerged in the measurement model.

Construct validity of the EOQ was explored by examining the relationship amongst error orientation scores and related variables: self-efficacy, self-esteem, plan-orientation, action-orientation after failure, readiness for change, control rejection, need for achievement, psychosomatic complaints, depression, negative affectivity, optimism, self-esteem, job uncertainty, and career stress (Rybowiak et al., 1999). A follow-up study suggested that Rybowiak's nomological net, which mainly investigated coping variables, is inferior to a motivational perspective on individual error orientation (Schell, Hernandez, & Rosebeary, 2008).

Rybowiak, Garst, Frese, and Batinic (1999) also developed scales to measure error competence and thinking about errors. They revised the EOQ to include enough items to create scales for these two constructs and carried out their second study with undergraduate students, asking them to respond in the context of work. They demonstrated the psychometric properties of the instrument, reporting internal consistency reliability values (Cronbach's alpha) of the subscales (dimensions) of above 0.70 except for error competence (α = 0.56). The data fit their eight-factor model, but they suggested with a larger sample size they would have also tested for a higher-order factor structure. The work by Rybowiak, Garst, Frese, and Batinic (1999) has been cited over 150 times and their instrument has frequently been used to study workplace error orientation. Table **1** includes the quoted definitions of the error orientation subscales and some points about the nature of the relationships they have with other variables.

Table 1. Error Orientation Construct Definitions and Relationships

Construct/Scale	Interpretation (Quoted from Rybowiak et al. 1999)	Relationships with Other Variables
Error competence	knowledge of how to immediately correct errors reduce their consequences (directed at short term goals)	relates to self-efficacy, to action-orientation after failure, need for achievement, and quite highly to initiative
Learning from errors	learn from errors in order to prevent them	correlates with self-efficacy, qualification, plan-orientation, need for achievement, readiness to change, and initiative
Error risk taking	having an achievement-oriented attitude that requires flexibility and taking responsibility	positive relations to need for achievement, qualification, readiness for change and initiative, as well as a negative relation to control rejection
Error strain	strained by making errors and fearing errors or reacting to errors with high emotions	positive correlations with other strain measures, such as psychosomatic complaints, depression, and negative affectivity
Error anticipation	a generalized fear of committing errors and by negative emotional reactions	correlated negatively with self-efficacy, self-esteem, and initiative and positively with control rejection, psychosomatic complaints, depression, and negative affectivity
Cover up errors	mainly the strategy of a non-self-assured person and may also be an adaptation to error-sensitive conditions at work, for example, job uncertainty	relates to low self-esteem, negative affectivity, and high control rejection, and little initiative, but also to career stress and job uncertainty
Error communication	propensity toward telling others about errors to prevent errors or improve oneself	not described in Rybowiak (1999)
Thinking about errors	reflecting on errors in order to correct them	not described in Rybowiak (1999)

Mateo, Muring, Malayao, Daniel, and Emperio (2013) investigated the appropriateness of using the EOQ with a sample of undergraduate students in the Philippines. They found a pattern of internal consistency, ranging from 0.63 to 0.89 depending on the scale, which was similar to the values calculated by Rybowiak and colleagues (.56-.89). Both studies provided evidence the scales are reliable. In another study, Schell (2008) had university students and employees in the U.S. respond to a portion of the EOQ and other measures to investigate the nomological net of polychronicity, or the "preference for working on multiple tasks at once". Cronbach's alpha values for EOQ subscales ranged from 0.61-0.74. The EOQ has been criticized for including items about both behaviours and attitudes and this may explain the low internal consistency reliability of some of the scales (Schell et al., 2008).

Tjosvold (2004) created a version of the EOQ to measure people's "blaming orientation" while working in teams to explore relationships between cooperative goals, problem solving, and learning from mistakes, but they did not include their items in their publication. Their major finding was that cooperative goals and a problem solving orientation positively related to learning from team mistakes. Blaming orientation scores were significantly positively correlated to competition and independence, and significantly negatively related to team problem solving scores.

I have adapted the original EOQ to a learning-situated EOQ to measure the learning error orientation of university organic chemistry students. While the errors in the workplace domain could be task errors or behaviour errors, those in a learning context reveal a deficiency in procedural or declarative knowledge. I will discuss the test-retest reliability of the responses in the Results, as it has not been explored for the original EOQ and other studies using the EOQ. Since Schell (2008) found a relationship between EOQ scores and achievement goal orientation scores, I also included the latter construct in my research.

Definition and Measurement of Achievement Goal Orientation

The literature on goal orientation is vast, spanning the academic, work, and sport domains. Achievement goal theory was originally described in a 1986 seminal paper by Carol Dweck (Dweck, 1986) that described adaptive and maladaptive motivational patterns in the context of achievement. Underlying, and supposedly leading to, these motivational patterns are beliefs about intelligence. Dweck's research uncovered links joining beliefs to goals to confidence to behaviour and, ultimately, performance. Dweck explained that these motivational processes affect children's ability to use their skills and knowledge, learn new skills and develop knowledge, and succeed on transfer tasks.

Achievement goals are "the purpose for which a person engages in achievement behaviour" (Elliot & Thrash, 2001) as opposed to actual goal one is trying to achieve (Anderman & Maehr, 1994). That is, achievement goal orientations represent people's reasons and methods for striving for particular goals (Anderman & Maehr, 1994). Achievement goal theory holds that these goals influence learning processes and outcomes (Huang, 2011) since they are the focus of one's competence-relevant engagement (Elliot, 1994). Achievement goals are referred to in the literature as "orientations", which suggests that these describe individual motivational traits. However, research has shown that classroom-level goal orientations can influence an individual's goal orientation and learning behaviours. Kaplan and Maehr (2007) summarize several researchers' view of goal orientation as "situated orientations for action in an achievement task" (Ames, 1992; Dweck, 1996; Nicholls, 1984).

Achievement goal theory research has uncovered several distinct goal orientations. Goal orientation theories have undergone revision concurrently with the development of tools to measure people's levels of particular orientations. Dweck (1986) suggested that maladaptive motivational processes be called performance goals and adaptive motivational processes be called learning goals. Individuals who hold performance goals want to perform well to receive positive judgments about their abilities, or to avoid negative judgments by others. Individuals who hold learning goals aim to increase competence in particular tasks. Dweck concluded that low ability learners would seek challenges and persist in tasks if they held learning goals, but would behave helplessly if they held performance goals (Dweck, 1986).

Button, Mathieu, and Zajac (1996) laid the conceptual and empirical foundation for the application of goal orientation to organizational research. Vandewalle (1997) then showed that workers displayed learning goals when they desired to learn new skills and improve their competence. In workplace achievement goal research, mastery orientation is referred to as learning goal orientation or task orientation, performance-approach goals are labeled “prove performance goal orientation” or “ego orientation” and performance-avoidance goals are labeled avoid performance orientation.

Goal orientation has also been explored in the domain of sport (Conroy, Elliot, & Hofer, 2003; Duda, 1989; Treasure & Roberts, 1994). Athletes’ task or ego orientations are predictive how they view the purpose of sport. That is, individuals with task orientation tend to believe in cooperation and sportsmanship, whereas individuals with an ego orientation tend to believe that sports enhance individual esteem and status (Duda, 1989).

In academic situations, achievement goals indicate the standards by which people evaluate their performance (Ames, 1992; Elliot et al., 2011). I will focus on personal achievement goal orientations as opposed to the perceived classroom climate orientations discussed by Ames (1992). Students working in an online homework environment in organic chemistry likely have a situational goal orientation with respect to this class, and perhaps even to specific tasks, including online homework, working through practice tests, and learning in class.

The study of achievement goal constructs over time has led to a distinction between approach and avoidance motivations (Senko, Hulleman, & Harackiewicz, 2011). This was described first for the performance goal orientation construct in Elliot’s doctoral dissertation (Elliot, 1994). In this work, Elliot confirmed that goals related to the avoidance of failure negatively impacted intrinsic motivation, whereas goals related to approach of performance did not. As with other achievement goal orientation research, Elliot suggested that mastery goals and performance-approach goals drive learners to challenges, whereas performance-avoidance goals lead to helpless behaviours.

Almost all measures of goal orientation are fixed-choice, self-report questionnaires. One exception is Ames and Archer's (1987) measure of mother's preferences for their children's goal orientation. Self-report items tend to be Likert or judgments of how much a statement is "like me". Once the achievement goal orientation literature was well established, all instruments involved multiple items to assess the degree to which individuals hold various orientations and to perform psychometrics, the importance of which was stressed by Vandewalle (1997). Many instruments are revisions or combinations of others, and goal orientation items are frequently included as a subscale on a larger measure, for example the Motivated Strategies Toward Learning Questionnaire (Pintrich, Smith, Garcia, & McKeachie, 1991). The questionnaires have been used to validate models about achievement goal theory. The "bifurcation" of performance and mastery goals into their approach and avoidance valences is described in Elliot and McGregor (2001).

In 2011, Elliot, Murayama, and Pekrun published a 3 × 2 achievement goal model and questionnaire. This model represents a leap away from the mastery- and performance definition of achievement goals. Elliot et al. argue for "the separation of task-based and self-based goals" (Elliot et al., 2011 p. 633). The "valence" of the model is positive, or approaching success, and negative, or avoiding failure. What is different from previous models is how competence is defined. Absolute competence refers to task performance, such as a correct score or understanding a concept. The achievement goal of task-approach is an aim to do well on tasks. The achievement goal of task-avoidance is an aim to avoid doing poorly. Intrapersonal competence depends on one's prior performance. The achievement goal of self-approach is to aim to do better than in the past, and self-avoidance is to avoid doing worse than in the past. Interpersonal performance involves comparison with others. The achievement goal other-approach is the aim to do better than others, and the goal of other-avoidance is to avoid doing worse than others. Thus, the 3 × 2 model has six possible orientations, self-approach, self-avoidance, task-approach, task-avoidance, other-approach, and other-avoidance. Elliot et al.'s (2011) data provides empirical support for this model. I use the 3 × 2 achievement goal questionnaire in this dissertation because of the task-focused nature of online homework learning. Also, because the questionnaire is new, I sought to determine if organic chemistry learner's responses fit the 3 × 2 structure.

The research on achievement goal orientation of university students includes experimental and observational studies. These studies examine the correlation among achievement goals, learning and other constructs, such as self-efficacy. Research has also been conducted to decide what can change a person's goal orientation and what a particular goal orientation may change in one's behaviour.

Goal orientation is likely to influence student behaviour in the online homework environment. In a review article, Cannon-Bowers and colleagues (1998) suggest that conditions of practice be conceptualized according to when they occur in relationship practicing a task. Goal orientation is a state or condition a learner is in *before* work on the task begins, and feedback is provided *after* the task is attempted (Cannon-Bowers et al., 1998). These researchers claim goal orientation is a component of the learning environment and can influence the effectiveness of training. They also suggest that directions given for a training session can influence a person's goal orientation. However, there is a lack of empirical description about how goal orientation, as a trait or state, affects how learners engage with online homework practice. Cannon-Bowers presented practical recommendations for setting up mastery goal orientation in learners before practicing psychomotor, procedural, or cognitive tasks through lecture-based instruction early in the training process.

In a five-stage model that describes how a learner receives feedback, goal orientation is considered a component of the learner's initial state, as well as prior knowledge and self-efficacy (Bangert-Drowns et al., 1991). The other four stages are search and retrieval strategies, responding to a question, evaluation of response in light of feedback, and making adjustments. These theories support the idea that goal orientation is relevant in online homework learning.

Relationship between Error Orientation and Goal Orientation

The constructs of error orientation and goal orientation have been discussed together in a few articles. Schell (2008), as previously described, showed that the error competence and error strain subscales of the EOQ were related to goal orientation in both students and employees. In this study, Schell used Vandewalle's (1997) three-factor achievement goal questionnaire, which measures people's endorsement of learning goals, prove-performance goals, and avoid-performance goals. These goals parallel the mastery-approach, performance-approach, and performance-avoidance goals suggested by research conducted by Elliot and Harackiewitz (1996) and later measured by Elliot and Sheldon (1997). Specifically, error competence scores were positively related to learning and prove-performance goal orientation. Error competence was negatively related to performance-avoidance orientation. Error strain was negatively related to learning goal orientation in students only, and positively related to prove-performance orientation and performance-avoidance orientation. Schell and colleagues suggests that people who score higher on the learning goal orientation scale may more frequently notice and correct errors as they are more persistent. Also, they explain that higher stress levels experienced by some people when they make errors may manifest itself as a performance-avoidance orientation. Individuals who have a high level of performance-avoidance goals want to avoid making mistakes to avoid looking incompetent, although this has not been tested.

The degree to which people can learn from errors was manipulated in job seekers via learning goal orientation training (Noordzij, van Hooft, van Mierlo, van Dam, & Born, 2012). Those who underwent learning goal orientation training had higher learning from failure, strategy awareness, and reemployment. This suggests that there may be a relationship between organic chemistry students' learning from errors and goal orientation.

Researchers investigated the degree to which individual differences in goal orientation and action-state orientation predicted error orientation and explored the degree to which participants focused on themselves when confronted with task failure. Learning goal orientation was significantly correlated with error mastery ($r = 0.24$, $p < 0.10$). Error aversion was significantly negatively correlated to action-state orientation after failure ($r = -0.22$, $p < .10$). The researchers developed a regression model with an adjusted R-square of 0.14 for learning goal orientation positively predicting, prove performance orientation negatively predicting, and failure-related action-state orientation negatively predicting error mastery. In a second regression model, with an adjusted R-square of 0.28, only failure-related action orientation was necessary to predict the variance in error aversion (Van Dyck, Van Hooft, De Gilder, & Liesveld, 2010a).

Some organizations aim to avoid errors, and some aim to manage errors. Arenas and colleagues (2006) note the parallel between error avoidance as an organizational culture and performance goal orientation, and between error management and mastery goal orientation.

In the domain of medicine, it is important for practitioners to report errors. Nursing educators aiming to train students not to hide errors report that their sample of students held high mastery goals, and moderate performance-approach and performance-avoidance goals (Dunn, 2014). Since those with performance-avoidance goals have a greater fear of failure, the potential shame from making an error leads them to covering up errors (Dunn, 2014). The organic chemistry online learning environment has much lower stakes than a clinical setting with potential life-and-death errors, however the link between goal orientation and error orientation is feasible. The literature suggests that both performance-approach and performance-avoidance orientations have a greater fear of failure (Elliot & Church, 1997).

In a laboratory study with undergraduate students, Van Dyck, Van Hooft, De Gilder, and Liesveld (2010a) examined error mastery or avoidance while having participants work on an error-prone task. Participants with an error mastery approach had higher levels of learning goal orientation and those with "state orientation" avoided errors. Mastery-oriented students used available cognitive resources to learn from their errors, and did not waste time and effort on feeling self-conscious. The error-prone tasks used in this study were timed-limited Tangram puzzles, which are somewhat similar to organic chemistry tasks.

Arenas and colleagues (2006) explored goal orientation and error orientation in the context of a decision-making task in the workplace. They manipulated self-efficacy and emotional state through an experimental treatment and a positive error orientation mitigated some of the negative impacts on performance. They measured error risk taking, error strain, and error communication and found several significant correlations between these scores and goal orientation scores. Error risk taking was positively correlated (r = 0.44) with learning goal orientation, error strain was negatively correlated (r = –0.34 with learning goal orientation, positively correlated with prove-performance goal orientation, and avoid performance goal orientation. Error communication was negatively correlated with prove performance goal orientation (Arenas et al., 2006). These relationships are logical and suggest that people with mastery goals are comfortable with taking risks, whereas those with performance goal orientations are strained regarding errors and would prefer not to communicate them. Through multiple regression, researchers modeled the effect of goal orientation on error orientation on performance.

Based on the organizational behaviour literature, individuals with learning goal orientations and performance goal orientations will take greater risks (Chia et al. 2003 as cited in Arenas et al., 2006). This may play out in the online learning environment. Submitting an incorrect answer is a “risk” and those with learning goal orientations will be more willing to take this risk to learn from the experience, and performance-approach individuals may be less likely to submit a wrong answer.

What is Missing from the Research?

Many of the above studies lack statistical power to explore many of the relationships between achievement goal orientations and error orientation. For example, Rybowiak, Garst, Frese, and Batinic (1999) suggested that there may be a higher-order factor structure to the EOQ, but their sample size was too small to test this. Additionally, not many have been conducted with students completing many error-prone tasks. My study will examine how error orientation and goal orientation relate to organic chemistry learner behaviours in an error-prone online homework environment. This will add to the understanding the nature of the relationship between goal and error orientation as well as provide insight into how students interact with the online homework environment.

Chapter 3.

Methods

Participants and Context

Research procedures for this study were approved as institutional research at The University of British Columbia. Participants in this study were 1212 students at the University of British Columbia. These students were enrolled in CHEM 233 (Organic Chemistry for the Biological Sciences) in term 1 of the 2011/2012 academic year. Demographic information obtained from the UBC student information system is summarized in Table **2**. Due to attrition, most of the analysis was conducted with the data from 1201 students who completed the course.

Table 2. Participant Demographics

Sex	Number of Participants
Females	757
Males	455
Citizenship	
Canada	999
Other	213
Age as of September 6, 2011	
15-17	9
18	238
19	617
20	155
21	70
22-24	68
25-27	30
Over 27	25

At the time of this research, most participants were undergraduate science majors and none were chemistry majors. Most participants were enrolled in various Bachelor of Science programs, such as general science, life science, and biology. Participant degree programs and year of study are presented in Table **3**.

Table 3. Academic Programs and Year Standing of Participants

Degree Program	Number of Students
Bachelor of Arts (BA)	41
Bachelor of Applied Science (BASC)	16
Bachelor of Commerce (BCOM)	1
Bachelor of Computer Science (BCS)	2
Bachelor of Dental Science (BDSC)	2
Bachelor of Human Kinetics (BHK)	10
Bachelor of Kinesiology (BKIN)	39
Bachelor of Science in Applied Biology (BSAB)	48
Bachelor of Science in Agriculture (BSAG)	1
Bachelor of Science in Biology (various programs) (BSC)	840
Bachelor of Science in Pharmacy (BSCP)	24
Bachelor of Science in Wood Products Processing (BSCW)	2
Bachelor of Science in Food, Nutrition, and Health (BSFN)	138
Bachelor of Science in Forest Science (BSFS)	14
Bachelor of Science in Global Resources (BSGR)	6
Exchange	1
Master of Health Administration (MHA)	1
Doctor of Philosophy (PHD)	1
Unclassified	26
Total	1213
Year Standing of Undergraduate and Unclassified Students	
1	115
2	961
3	81
4	24
5	29
Total	1210

Measures

Goal Orientation Questionnaire

To measure participants' achievement goal orientations, I used Elliot, Murayama, and Pekrun's (2011) 3 × 2 achievement goal questionnaire (AGQ), which measures task-approach and avoidance, self-approach and avoidance, and other-approach and avoidance goals. This statement was included before the list of items: "The following statements represent types of goals that you may or may not have for this class (CHEM 233: Organic Chemistry). Select the number that indicates how true each statement is of you." As in Elliot et al. (2011), the response scale used was from 1 ("not true of me") to 7 "extremely true of me". The 18-item questionnaire was administered with the UBC survey tool, Vovici by Enterprise Feedback Management. The data was stored on UBC servers to comply with the Freedom of Information and Protection of Privacy Act (FIPPA). An additional item was added in the 14th position that stated "We use this statement to discard the surveys of people who are not reading the questions. Please select 5 for this question." Example items are included in Table **4**, and the complete questionnaire is included in Appendix B.

Table 4. Sample Items from the Achievement Goal Questionnaire

Subscale	Item
Task-approach	To answer a lot of questions correctly in this class.
Task-avoidance	To avoid incorrect answers on the exams in this class.
Self-approach	To perform better on the exams in this class than I have done in the past on these types of exams.
Self-avoidance	To avoid doing worse on the exams in this class than I normally do on these types of exams.
Other-approach	To do well compared to others in this class.
Other-avoidance	To avoid doing poorly in comparison to others on the exams in this class.

Error Orientation Questionnaire

I adapted the Error Orientation Questionnaire (EOQ) used in workplace studies of coping with errors (Rybowiak et al., 1999) to improve fit to an academic setting. The original survey included eight subscales: error competence, learning from errors, error risk taking, error strain, error anticipation, covering up errors, error communication, and thinking about errors. When adapting this instrument for use in a learning setting, I substituted references to “work” with references to “problem solving”, “this class”, or “learning.” I added five items intended to measure an additional construct, error motivation. These items aimed to tap into students’ behaviour and thoughts during studying related to making mistakes. For example “When I make an error, I feel motivated to correct it” and “When deciding what to study, I choose the problems that I won’t make many errors on.” Example items are included in Table **5**, and the complete questionnaire is included in Appendix B.

Table 5. Sample Items from the Error Orientation Questionnaire and Adapted Version

Subscale	Original	Adapted
Competence	When I do something wrong at work, I correct it immediately.	When I do something wrong when solving a problem, I correct it immediately.
Learning	Mistakes assist me to improve my work.	Mistakes assist me to improve my knowledge and ability.
Risk taking	If one wants to achieve at work, one has to risk making mistakes.	If one wants to achieve in this class, one has to risk making mistakes.
Strain	I find it stressful when I err.	I find it stressful when I make mistakes during problem solving.
Anticipation		
Covering up	I do not find it useful to discuss my mistakes with other students.	I do not find it useful to discuss my mistakes.
Communication	If I cannot rectify an error by myself, I turn to my colleagues.	If I cannot manage to correct an error myself, I rely on others to help me.
Thinking	If something goes wrong at work, I think it over carefully.	After I have made an error, I think long and hard about how to correct it.
Motivation	N/A	1. When I make an error, I feel motivated to correct it. 2. When I make a lot of mistakes when solving a problem, I feel discouraged. 3. When deciding what to study, I choose the problems that I wont' make many errors on. 4. I notice when I am making the same mistakes on more than on problem. 5. If I make the same mistake twice, I think I might never learn how to do it properly.

This statement was included before the list of items: "Select the number that indicates the extent to which the statement applies to you in the context of this class (CHEM 233: Organic Chemistry)". Choices were 1 ("not at all"), 2 ("a bit"), 3 ("neither a bit, nor a lot"), 4 ("a lot"), and 5 ("totally"). As in the AGQ, I included an additional item in the 28th position that stated "We use this statement to discard the surveys of people who are not reading the questions. Please select 4 for this question." Appendix B contains the EOQ questionnaire items.

Online Homework

Most of the online homework questions used in this study were created by the system provider, Sapling Learning (Austin, Texas). Members of the teaching team created a small number of questions. All questions chosen for online homework assignments aligned to one or more course learning objectives. Questions required a variety of input formats, including multiple-choice, matching, ranking, and drawing. Most questions required students to draw one or more chemical structures and were similar to those on typical organic chemistry examinations, although the average difficulty of the questions is likely a bit lower than on examinations. Students were likely aware of this as instructors typically note that the online homework questions are best for initial practice, and the sample exams are said to be the best gauge of the actual exam. Most questions had a "hint" available that students could access without grade penalty. Most questions provide feedback when students submit an incorrect answer. Correct answer explanations were also available if the students chose to look at them. Students moved through each assignment by completing questions in any order they wished. Work was saved after each question and students could complete each assignment in more than one sitting. Assignments were untimed. For each incorrect answer submitted, students lost 5% of their grade for that question. Since each question was worth one point, each incorrect answer reduces their score on that question by 0.05 points.

Confidence and Certainty

Approximately half of the questions in the online homework included confidence and certainty prompts. Below each question, students were prompted to enter a numerical answer between 0% and 100% to answer the prompt "If this was the last time you studied this material before an exam, how sure are you that you could correct solve a similar problem? Enter a percent from 0% (definitely could not solve it) to 100% (definitely could solve it)." Beside or below this question, they were prompted "How sure are you of this prediction? Enter 1, 2, or 3: 1. Unsure 2. Somewhat sure 3. Sure". Figure 1 shows the student view of one of the online homework questions. The components on the screen include the chemical structure input, hint, check answer, next question, and the option to give up and view the solution. Below the question are the confidence and certainty judgment questions, which were numerical inputs.

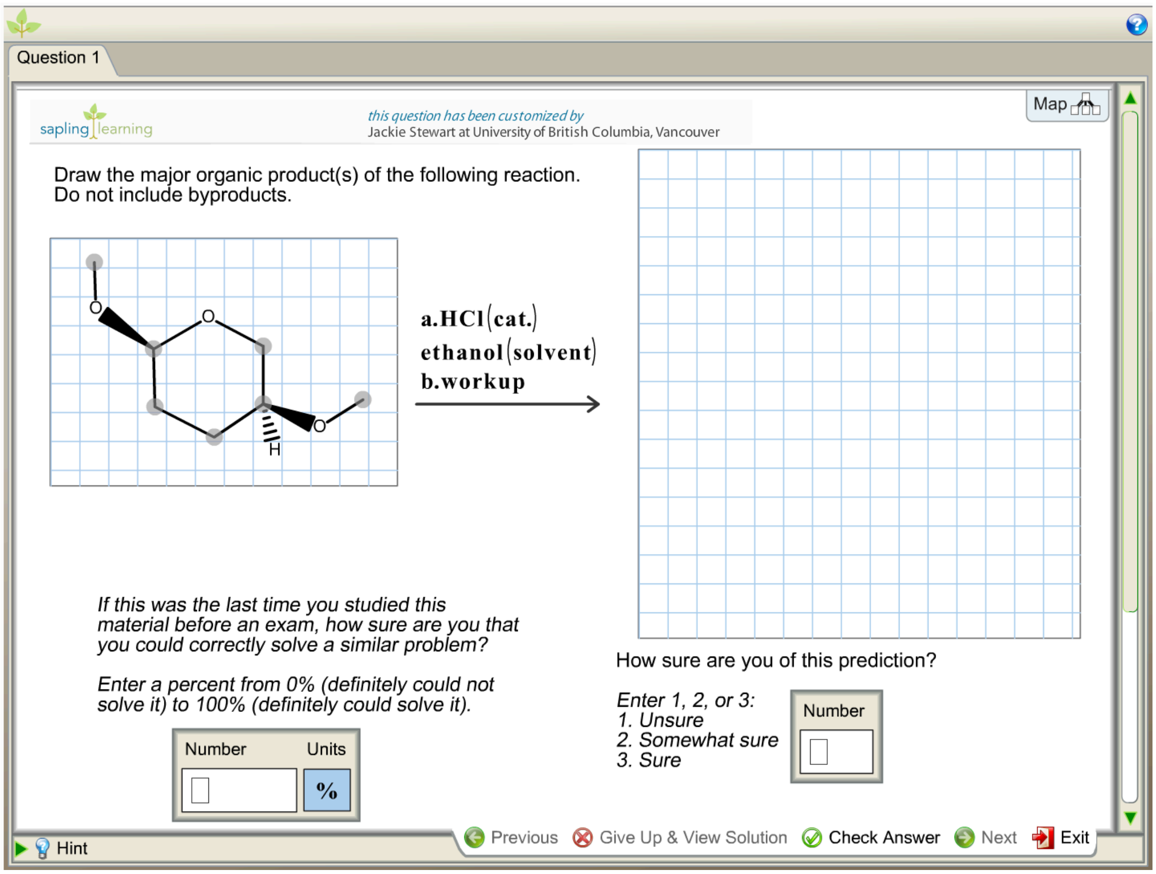

Figure 1. Screen capture image of an online homework question.

Academic Performance

Academic achievement was assessed with one-hour midterm examinations and a 2.5-hour final examination. The first midterm exam consisted of 29 multiple-choice questions and 7 multi-part constructed-response questions, with a total score of 75 points. The second midterm exam consisted of 18 multiple-choice questions and 5 multi-part constructed-response questions, with a total score of 50 points. It consisted of 51 multiple-choice questions and 7 multi-part constructed-response questions. Multiple-choice sections were graded using Scantron forms and Excel, and teams of teaching assistants whom course instructors supervised graded constructed-response questions.

Participation in class was estimated using worksheet and personal response system participation scores. Each of the instructors had their own standards and grading scheme for participation. The tasks varied between sections and included in-class activities, personal response system “clicker” questions and in-class quizzes.

Procedure

The teaching team notified students there would be upcoming surveys throughout the term and that students would receive a small bonus mark for responding. Email was the method of recruitment. Text of the message is included in Appendix C. For each AGQ and EOQ survey completed, 0.25 of a percentage point was added to students’ overall course grade. Students who submitted blank surveys or those with missing data were awarded bonus credit, to avoid coercion.

Students were randomly divided into four groups and asked to respond to the goal orientation questionnaire and error orientation questionnaire during different time periods according to the following schedule:

- Group A: Five days before each midterm.
- Group B: Five days before midterm 1 and five days after midterm 2.
- Group C: Five days after midterm 1 and five days before midterm 2.
- Group D: Five days after each midterm.

This procedure was used to study the effect of exams on goal and error orientations, and to estimate test-retest reliability. An email invitation was sent out when first round of surveys became available (Appendix C). Additionally, a pop-up announcement notified the class that the surveys were becoming available and asked them to check the course website to find out when their surveys would be available. Occasionally, students were allowed to submit a survey late and receive the bonus mark.

Students purchased access to Sapling Learning as a regular part of their course requirements. One or two assignments were due each week, and students were notified by email and learning management system announcement when a new homework assignment became available. Typically, participants had five to seven days to complete each homework assignment. When the confidence and certainty questions were included, a pilot assignment with two questions was introduced to students via an email from the course coordinator, which included the following statement “You will notice that this week's Sapling problems have "reflection" type questions added to them. I have added these because research shows that thinking about your answers while solving problems is important. You don't have to answer the reflections questions to get the points for the Sapling problems, but I recommend you do” (Appendix C).

After the pilot of two questions, the next message included the following statement: “As with last week, the problems have "reflection" type questions added to them. The reflection questions are worded slightly differently than last week so please read them carefully. I have added these because research shows that thinking about your answers while solving problems is important. You don't have to answer the reflections questions to get the points for the Sapling problems, but I recommend you do. If you do answer the reflection questions, answer both of them for each problem.” Students needed to answer *neither* of the confidence nor certainty questions or *both* of the confidence and certainty questions to avoid them interfering with the grading of the chemistry portion of the item. In assignments following the pilot, every question contained the confidence and certainty prompts. Table **6** contains a summary of the homework assignments.

Table 6. Summary of Online Homework Assignments

Due Date (month/day/year)	Assignment Topic	Number of Questions	Number of Questions with Confidence
9/18/11	Training assignment	24*	0
9/28/11	Representing organic structures	28	0
9/28/11	Bonding	27	0
9/28/11	Stereochemistry	28	0
9/29/11	Intermolecular forces	14	0
10/2/11	Fundamentals of chemical reactions	10	0
10/4/11	Redox classification	10	0
10/5/11	Organic acidity/basicity	32	0
10/16/11	Introduction to reaction mechanisms	8	0
10/30/11	Substitution and elimination reactions	51	2
11/8/11	Electrophilic addition to alkenes	11	11
11/8/11	Redox reactions	12	12
11/27/11	Nucleophilic addition to aldehydes and ketones Part 1	10	10
11/27/11	Nucleophilic addition to aldehydes and ketones Part 2	9	9
12/4/11	Carbohydrates	12	12
12/5/11	Nucleophilic acyl substitution	22	22

Note: The training assignment score was scaled to be worth 10 points, all other questions count as one point toward the student's homework grade.

Occasionally, the teaching team extended the due date for a homework assignment as part of regular classroom practice. If the extension was given before the due date of the assignment, the student did not see the correct answers in his or her account and the responses have been included in the analysis.

Chapter 4.

Analysis and Results

Goal Orientation and Achievement

How do organic chemistry students' levels of task, self, and other-approach and avoidance goals relate to achievement? To answer this question I begin with a description of course achievement variables and their relationships to each other. Then, I present results of the confirmatory factor analysis of the AGQ responses. Finally, I use correlational analysis to describe the relationship between AGQ scores and achievement.

Organic Chemistry Achievement

Students' scores for each course component were collected as part of standard classroom practice. I discuss achievement scores in four sections: course examinations (two midterm exams and one final exam), section-specific component, online homework scores, and overall course performance.

Exam Scores. Many students drop the class early in the term or withdraw after the first midterm exam. In my analysis, I included only students who completed the course by writing the final exam in December (N = 1194) or a deferred exam, per University policy, in January (N = 7). The January exam was very similar to the December exam, and the number of students who wrote the January exam is small. Students who miss the first, second, or both midterm exams due to illness account for differences in numbers of students who wrote each examination. Four students did not write the first midterm and 19 students did not write the second midterm. One student did not write either midterm exam.

To create a more representative and reliable measure of achievement, I constructed a composite exam score for each student, hereafter called weighted exam. Each of the two midterms was weighted at 18.75% and the final examination was weighted 62.5%. This corresponds to the relative weighting of these components for the course grade at 15%/15%/50%. If a student missed one midterm examination, the final exam score was substituted for the missing midterm score. This assigned a weight of 18.75% to the single midterm and a weight of 81.25% to the final examination. The weighted exam score will be used to explore relationships among achievement, goal orientation, error orientation, and online learning behaviour.

Descriptive statistics for each examination and the weighted exam score are shown in Table **7**. Histograms for all continuous variables used in this study are included in Appendix D. All variables show only minor departures from normality.

Table 7. Descriptive Statistics for Course Examinations and the Weighted Exam Score

Description	Mean Percent	Standard Deviation	Skewness	Kurtosis
Midterm 1 (*N* = 1197)	61.88	13.37	-.001	-.371
Midterm 2 (*N* = 1182)	56.08	18.79	-.195	-.587
Final Exam (*N* = 1201)	64.79	17.74	-.305	-.695
Weighted Exam (*N* = 1201)	62.57	16.02	-.226	-.661

Online Homework Performance. Students completed online homework using the Sapling Learning software. These activities were worth 15% of students' course grades if including it meant their grade was higher than without it (see Course Grades, below). In this software system, students can attempt a question multiple times and receive partial credit. During the term in which this study was conducted, the penalty was quite low, only 5% per attempt. Thus, attempting a question four times would have a student end up with a score of 80% on the question. Sapling Learning scores were created by summing the item scores. Thus, the variable Sapling percent is a percentage score of all items weighted equally.

I examined the distribution of Sapling Learning scores (as a percent) for students who created a Sapling Learning account (Figure 36 in Appendix D). Due to the way the system graded student responses, the grades are strongly negatively skewed, –2.588, and leptokurtotic, 7.968. This departure from normality suggests that linear homoscedastic relationships with other variables are not likely. Since the score itself is not particularly meaningful because different questions have different partial credit rules, transformation is a reasonable approach. After testing three transformations (square root of K-X, log10 of K-X, and inverse of K-X; where K is a constant of 1 greater than the highest score and X is the variable), I selected the logarithmic transformation to reduce skewness. The distribution of the transformed variable, transformed Sapling percent, is shown in Figure 37 in Appendix D. The distribution of the transformed Sapling percent is well within tolerances to be considered normal (skewness statistic = .233, kurtosis statistic = –.722). Lower values of the transformed Sapling percent variable correspond to better achievement on the online homework questions.

The Sapling Learning percent score was used in the calculation of the course grades for 1088 students. For the 68 students for whom the score was not used, the distribution of Sapling percent scores is not normal (Figure 35 in Appendix D). Many of these students may not have taken interactions with the system seriously and either stopped using the system at some point during the term or used the system in ways that were not oriented toward achieving a high score. Another possibility is that some of these students had unpredictably high exam scores, so much so that their Sapling Learning grade would have depressed their course grade, which explains why it was not used.

Achievement and Online Homework. What is the relationship between achievement and online homework performance? A scatterplot of Sapling Learning grades and weighted exam scores is shown in Figure 2, $r = .499$, $p < .01$, $n = 1156$. Note: all correlation coefficients are 2-tailed unless otherwise indicated. Predictably, the relationship is visually masked due to the skewed nature of the Sapling Learning grading. Additionally, due to the large skew in Sapling percent, the relationship between it and weighted exam is heteroscedastic.

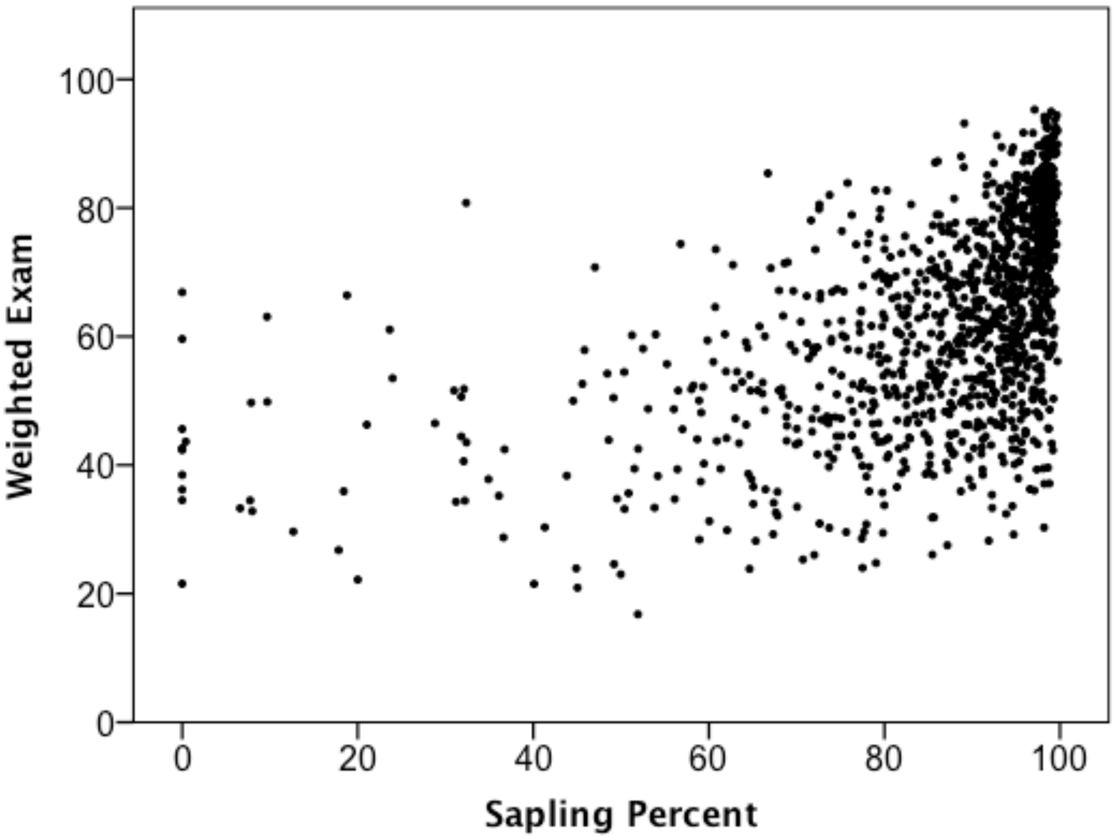

Figure 2. Scatterplot of weighted exam and Sapling Learning percent scores for all students with Sapling Learning accounts.

The relationship between weighted exam and the transformed Sapling percent score shown in Figure 3 is linear and homoscedastic, $r = -.591$, $p < .01$, $n = 1156$.

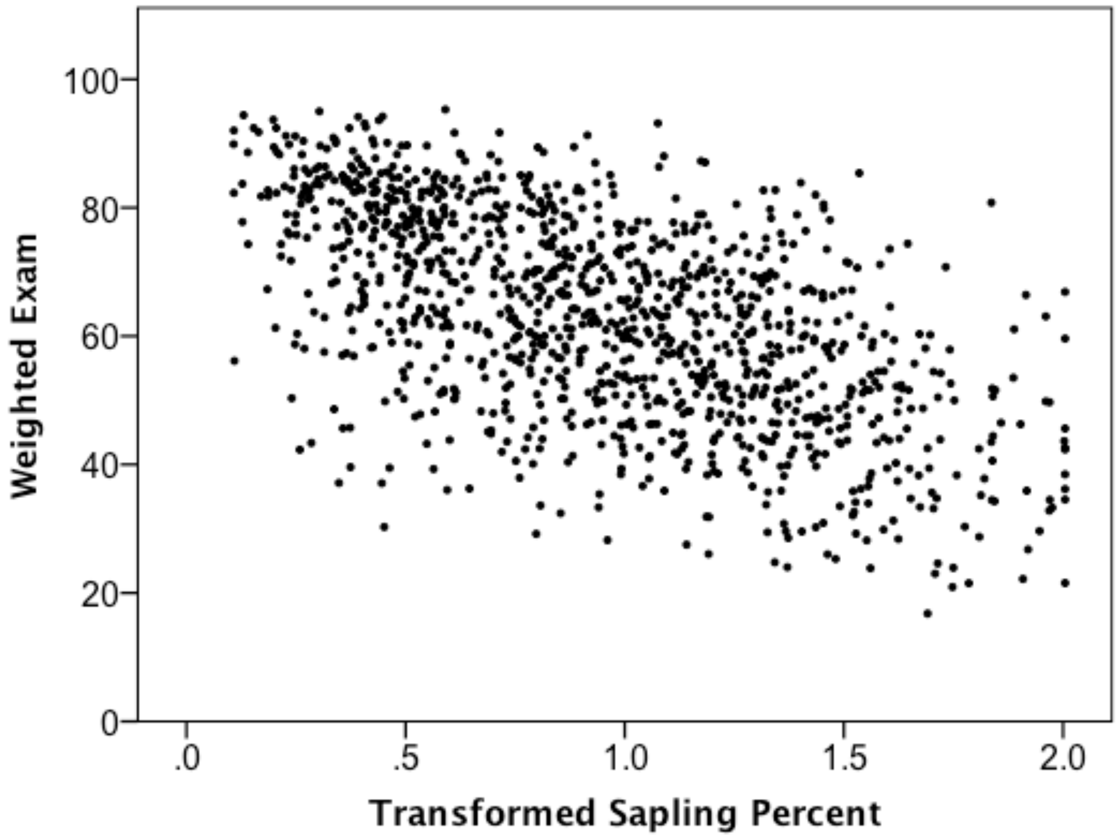

Figure 3. Scatterplot of weighted exam and the logarithmic transformation of Sapling percent for all students with Sapling Learning accounts.

Course Grades. A small incentive was added to students' grades for completing the achievement goal orientation and goal orientation questionnaires (0.25% for each questionnaire for a maximum of 1%). Students' overall grades for the course were calculated by giving them the highest score out of the two grading options shown in Table **8**.

Table 8. The Component Weights Used to Calculate Students' Course Grades

Component	Option 1	Option 1
Midterm 1	15	18
Midterm 2	15	18
Final Examination	50	59
Sapling Learning Homework	15	N/A
Section-Specific (In-Class Quizzes, Participation)	5	5

Note: The bonus for completing the course surveys was added after the highest grade was selected from the above options.

Table **9** contains four combined grades considered for use in further analyses. In this Chapter I avoid using variables in the same analysis that are constructed from the same measures to avoid multicollinearity.

Table 9. Descriptive Statistics for Course Grades

Description	Mean (N = 1201)	Standard Deviation	Skewness	Kurtosis
Option 1 (with Sapling Learning)	66.53	15.23	-.421	-.293
Option 2 (without Sapling Learning)	63.37	15.71	-.248	-.610
Course Grade (without survey incentive)	66.98	14.79	-.375	-.386
Submitted Course Grade (Largest of Options 1 and 2, with survey incentive)	67.72	14.95	-.383	-.401

Achievement Goal Endorsement

To what levels do students endorse various achievement goals? To answer this question, I describe the data and confirmatory factor analysis results.

Response Rate. The first time the Achievement Goals Questionnaire (AGQ) was offered, I obtained 1092 responses, which included several duplicates and otherwise unusable data. Five submissions did not include a student identification number and 32 did not have the correct response to the "reading" test question ("We use this statement to discard the survey of people who are not reading the questions. Please select 5 for this question."). Six submissions were exact duplicates and 19 were double submissions that had one or more differences between them. In the cases of double submission, I retained the first submission, except in one case where the submissions differed only by two blank vs. non-blank responses for which I retained the submission without blanks. I removed one submission that was missing responses to eight items. This left 1029 responses for analysis. Response rates for the four groups were comparable, ranging from 82.1%−84.7% with an overall response rate of 83.6% (Table **10**).

Table 10. Initial Achievement Goal Orientation Questionnaire Response Rates by Group

Group	Invited	Useable	Response Rate (%)
A	308	253	82.1
B	308	261	84.7
C	308	257	83.4
D	307	258	84.0
Total	1231	1029	83.6

By the time the re-test AGQ was offered, fourteen students had dropped the course. Thus, fewer students were invited to do the retest AGQ than were invited to do the initial test. The number of submissions for the retest AGQ was 949, but this also included duplicates and otherwise unusable data: twelve responses did not include student identification or contained a student ID that did not match anyone on the class list, six submissions were completely blank, 25 students submitted a survey with an incorrect reading check, five submissions were exact duplicates, and 10 submissions were non-identical duplicates were submitted. Thus, 890 responses remained, for an overall response rate of 73.1% (Table **11**).

Table 11. Re-test Achievement Goal Orientation Questionnaire Response Rates by Group

Group	Invited	Useable	Response Rate (%)
A	305	224	73.4
B	302	206	68.2
C	306	237	77.5
D	304	223	73.4
Total	1217	890	73.1

I elected not to impute missing data for entire surveys since inference to the population is a secondary goal to exploring possible relationships between constructs.

Missing Item Responses. Structural equation modeling or confirmatory factor analysis cannot be undertaken with datasets containing missing data. Between 1 and 16 responses were missing for each variable in the initial AGQ dataset. The submission with 8 missing data points and the submission with 15 missing data points were deleted and not included in the missing value analysis. This left 140 out of 1028 cases (13.6% with between one and three missing data points. Each variable had between 1 and 16 missing responses (Table **12**). There is a small increase in missing values in the second half of the survey (71 to 92), which indicates that the missing data is not only because of students abandoned the survey. However, no clear pattern emerges from Table **12**.

Table 12. Number of Missing Responses for Each Variable in the Initial AGQ

Variable	Number of Missing	Position in Survey
Task-approach Item 1	11	16
Task-approach Item 2	8	6
Task-approach Item 3	10	4
Task-avoidance Item 1	17	17
Task-avoidance Item 2	9	2
Task-avoidance Item 3	14	10
Self-approach Item 1	14	11
Self-approach Item 2	9	3
Self-approach Item 3	11	7
Self-avoidance Item 1	7	12
Self-avoidance Item 2	16	5
Self-avoidance Item 3	13	18
Other-approach Item 1	5	15
Other-approach Item 2	7	13
Other-approach Item 3	1	9
Other-avoidance Item 1	8	19
Other-avoidance Item 2	1	1
Other-avoidance Item 3	6	8

Note: Position 14 was occupied by the reading check question.

Dealing with the missing item responses is a substantial issue and Appendix E details the many decisions made on grounds of exploratory analyses. The upshot of this work was to use maximum likelihood imputation, specifically expectation-maximization (EM) to impute scores for missing item responses. This method estimates the missing data in a way that does not underestimate variances (Allison, 2003). Multivariate normality is a requirement of the EM algorithm. Allison (2003) warns that when a covariance matrix from an EM dataset is used in SEM (or CFA) for an overidentified or just identified model, confidence intervals and *p*-values should be interpreted with caution. I decided against using multiple imputation because of the exploratory nature of this research. Multiple imputation methods require creating several datasets and averaging the parameters that result from the analysis of each one.

For the descriptive statistical analysis, each missing data point was imputed by calculating the average of the other two scores that were designed to measure the same construct, which is a type of conditional mean imputation. When two values were missing from the set of three items measuring the same construct, the value from the answered item was simply replicated. This method has the risk of decreasing variance and slightly overestimating the internal consistency of each subscale. However, the relationships between AGQ subscale scores and other variables in the analysis are preliminary at this point, which justifies the use of this simpler method.

Between zero and 22 responses were missing for each variable in the re-test dataset. Since this was relatively low at 0-2.4%, no missing data analysis was performed to detect possible impact(s) on subsequent analyses due to systematic error. Each missing data point was imputed in the same way as for the initial test scores.

Descriptive Statistics. Means, standard deviations, and internal consistency-reliability (Cronbach's α) of the responses to the AGQ are included in Table **13** (N = 1029). The subscales defined by each referent and dimension of competence have been defined as in (Elliot et al., 2011) and will be referred to as variables or subscale scores. Larger values indicate a stronger endorsement of the goal.

Table 13. Descriptive Statistics and Internal Consistencies for Initial AGQ

Variable	M	SD	Observed Range	Cronbach's α
Task-approach goals	5.93	1.01	1-7	.87
Task-avoidance goals	5.80	1.03	2-7	.76
Self-approach goals	5.57	1.12	1-7	.80
Self-avoidance goals	5.57	1.17	1-7	.83
Other-approach goals	5.10	1.41	1-7	.88
Other-avoidance goals	5.29	1.35	1-7	.87

The distributions for all variables above are significantly negatively skewed and platykurtotic. This is not unexpected, since the sample is a highly motivated group of students who tend to want to perform well in organic chemistry. As hypothesized, these students have high levels of task-and other-approach orientations. The mean other-approach score is the smallest of the subscores, but it is still above 5, which is in the "true of me" side of the scale. Table 14 compares the means for students who took the survey before and after their midterm examination.

Table 14. Descriptive Statistics and Internal Consistencies by Group

	Variable	M	SD	Cronbach's α
Groups A and B (N = 514)	Task-approach goals	5.95	1.03	.87
	Task-avoidance goals	5.82	1.03	.75
	Self-approach goals	5.59	1.16	.80
	Self-avoidance goals	5.57	1.21	.82
	Other-approach goals	5.13	1.21	.86
	Other-avoidance goals	5.33	1.34	.87
Groups C and D (N = 515)	Task-approach goals	5.91	1.00	.87
	Task-avoidance goals	5.78	1.04	.76
	Self-approach goals	5.55	1.10	.79
	Self-avoidance goals	5.56	1.13	.84
	Other-approach goals	5.07	1.43	.90
	Other-avoidance goals	5.26	1.36	.87

Note: Groups A and B responded to the AGQ before the first midterm exam, and groups C and D responded after first midterm exam.

To begin exploring the factor structure of the AGQ responses, I calculated the Pearson correlation coefficient for each item pair. All items were significantly correlated with all other items, p = 0.01. The correlation table is included in Appendix F (Table 57).

Within-referent correlations of the initial AGQ. One would expect that within self-, other- and task comparisons, the approach items would be highly correlated with each other, the avoidance items would be highly correlated with each other, and approach and avoidance items would have smaller correlation coefficients. There are several instances where this prediction does not hold. One example is the correlation between task-avoidance 1 and task-approach 1 (.689), is larger than the correlation between task-approach 1 and task-approach 2 (.667) although this is not a statistically detectable difference. These items may be tapping into the same construct in this population, despite the phrasing of the statements clearly referring to approach or avoidance goals. The correlation between self-approach 1 and self-approach 2 (.524) and self-approach 2 and self-approach 3 (.462) are a little smaller than the other within-referent correlations. Self-approach 2 is more highly correlated with all two items in the avoidance factor, self-avoidance 1 and self-avoidance 3 (.480 and .506, respectively). The correlations within avoidance and approach items for the "other" referent are generally higher than those between avoidance and approach items, although the differences are not very large.

Between-referent correlations of the initial AGQ. The correlations between task-referent items and self- and other-referent items range between .3 to .4 with the exception of task-avoidance 1 and self-avoidance 3 (r = .533) and tasks avoidance 2 and other-avoidance 2 (r = .501). The correlations between self- and other-referent items all range between .2 and .3 with the exception of some of the avoidance items, which were into the low .4 range. These item correlation coefficients do not support the theory that approach and avoidance are separate factors for the task and self-referents.

Correlations among AGQ variables. To explore relationships between variable scores, I calculated Pearson correlation coefficients for each pair for the entire initial test (n = 853) (

Table **15**) and the groups who responded to the questionnaire before (n = 410) and after the midterm exam (n = 443) (Table **16**). These coefficients are generally larger than those reported by Elliot et al. (2011).

Table 15. Correlations among the AGQ Variables

Variable	1	2	3	4	5	6
1. Task-approach goals	----	.728	.611	.507	.478	.435
2. Task-avoidance goals		----	.568	.667	.403	.540
3. Self-approach goals			----	.708	.335	.361
4. Self-avoidance goals				----	.335	.526
5. Other-approach goals					----	.740
6. Other-avoidance goals						----

Note: All correlation coefficients are significant at the 0.01 level.

Table 16. Correlations among the AGQ Variables Split by Groups A/B and C/D

Variable	1	2	3	4	5	6
1. Task-approach goals	----	.736	.610	.505	.486	.447
2. Task-avoidance goals	.721	----	.587	.674	.412	.573
3. Self-approach goals	.612	.548	----	.725	.357	.364
4. Self-avoidance goals	.510	.659	.690	--------	.333	.520
5. Other-approach goals	.470	.393	.312	.338	----	.735
6. Other-avoidance goals	.422	.508	.357	.533	.744	----

Note: Values above the diagonal are for groups A and B; values below the diagonal are for groups C and D.
Note: All correlation coefficients are significant at the 0.01 level.

Correlations are largest for approach and avoidance goals of the same referent. For example, the correlation coefficient for task-approach and task-avoidance is .735, between self-approach and self-avoidance is .725, and other-approach and other-avoidance is .753. These are the largest correlation coefficients amongst the achievement goal variables. The correlation coefficients between groups A/B and C/D are quite similar, indicating the factor structure is consistent over time.

Model Validity. Do the responses fit the 3 × 2 model? To further explore the factor structure of the AGQ in this sample, I performed a confirmatory factor analysis (CFA) using LISREL. I selected LISREL because it offers the robust maximum likelihood estimation method, which is robust to departures from normality.

As a starting point, I hypothesized the model based on Elliot et al. (2011)'s findings of six factors – task, self, and other-approach/avoidance – found to fit the data from students in undergraduate psychology classes.

CFA Assumptions. The assumptions of multivariate normality and linearity were evaluated through SPSS and AMOS. The statistics for univariate and multivariate skewness and kurtosis are shown in Table **17**.

Table 17. Normality Statistics by Item

Variable	Min.	Max.	Skewness Statistic	Critical Ratio	Kurtosis Statistic	Critical Ratio
Other-avoidance 1	1	7	-0.914	-11.972	0.441	2.889
Other-avoidance 2	1	7	-0.8	-10.483	0.251	1.647
Other-avoidance 2	1	7	-0.924	-12.095	0.487	3.19
Other-approach 1	1	7	-0.532	-6.969	-0.381	-2.497
Other-approach 2	1	7	-0.92	-12.045	0.503	3.294
Other-approach 3	1	7	-0.749	-9.811	-0.001	-0.009
Self-approach 1	1	7	-0.908	-11.893	0.407	2.663
Self-approach 2	1	7	-0.903	-11.825	0.491	3.216
Self-approach 3	1	7	-0.77	-10.087	0.247	1.616
Self-avoidance 1	1	7	-0.895	-11.717	0.435	2.849
Self-avoidance 2	1	7	-1.004	-13.147	0.832	5.447
Self-avoidance 3	1	7	-1.046	-13.694	1.06	6.943
Task-avoidance 1	1	7	-0.95	-12.439	0.757	4.955
Task-avoidance 2	1	7	-1.272	-16.656	1.702	11.147
Task-avoidance 3	1	7	-1.257	-16.465	1.556	10.19
Task-approach 1	1	7	-1.036	-13.562	0.926	6.064
Task-approach 2	1	7	-0.968	-12.677	0.457	2.991
Task-approach 3	1	7	-1.051	-13.759	0.655	4.29
Multivariate					298.167	178.226

If kurtosis values above seven are used as a cut-off (West et al., 1995 as cited in Byrne, 2010), there is no evidence of departure from normality. However, the multivariate critical ratio of 178 indicates that there is a large departure from multivariate normality. Bentler (2005) in Byrne (2010) suggests that values above 5.00 indicate non-normal distributions of multivariate data. To be safe, I conducted the CFA with the robust maximum likelihood estimation method, which is theorized to produce unbiased estimates with non-normal data.

Table 18. CFA Fit Statistics for Hypothesized 3 × 2 Model for the Initial and Retest Data

Model	NPAR	Scaled Chi-Square	DF	GFI	CFI	RMSEA
Initial Test	51	8.111 (p > .9999)	120	0.906	1.000	0.0739
Retest	51	46.51 (p > .9999)	120	0.895	1.000	0.0808

The fit is reasonable, according to the fit statistic criteria in Browne & Cudeck (1993) as cited in Elliot et al. (2011), since the CFI is above 0.9 and the RMSEA is below (or very close to) 0.08. Thus, both the initial AGQ data and retest data are reasonable fits to the 3 × 2 model. Figure 4 and Figure 5 contain the model diagram with standardized estimates. Tables 19 and 20 contain the latent variable covariances.

Table 19. Latent Variable Covariances – Initial Test

Variable	1	2	3	4	5	6
1. Task-approach goals	---					
2. Task-avoidance goals	.923	---				
3. Self-approach goals	.680	.7	---			
4. Self-avoidance goals	.607	.811	.863	---		
5. Other-approach goals	.545	.52	.388	.397	---	
6. Other-avoidance goals	.525	.659	.458	.615	.852	---

Table 20. Latent Variable Covariances – Re-Test

Variable	1	2	3	4	5	6
1. Task-approach goals	---					
2. Task-avoidance goals	.957	---				
3. Self-approach goals	.79	.801	---			
4. Self-avoidance goals	.779	.864	.933	---		
5. Other-approach goals	.555	.555	.507	.518	---	
6. Other-avoidance goals	.542	.641	.546	.628	.894	---

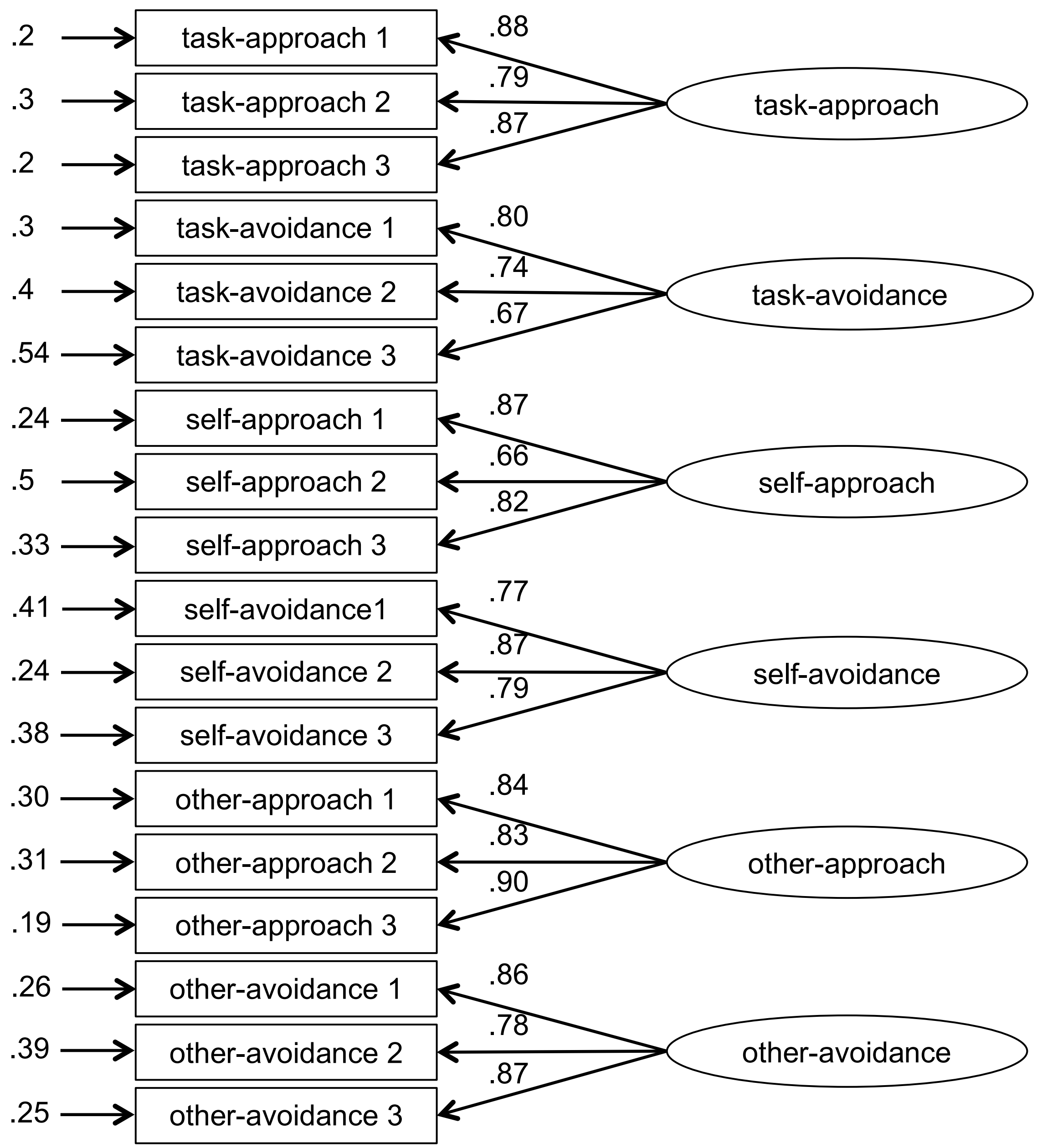

Figure 4. Confirmatory factor analysis model for AGQ test data.

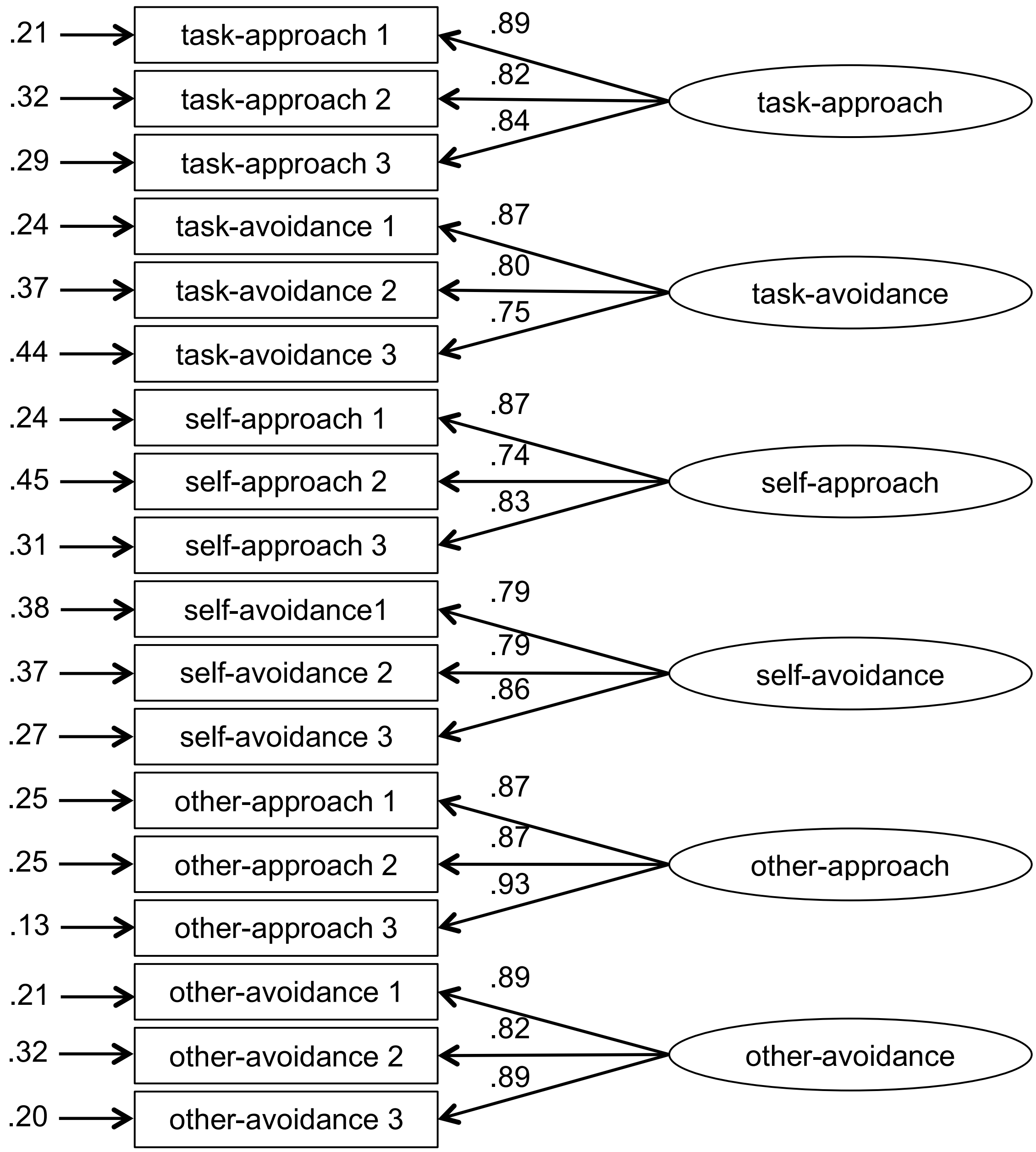

Figure 5. Confirmatory factor model for AGQ retest data.

Test-retest reliability of AGQ constructs. Are students' endorsements of achievement goals influenced by major course events? The class was divided into four groups to explore whether major course events affected students' endorsement of particular goal orientations. Variable scores (calculated as means) were correlated between the first and second administrations of the AGQ. The Pearson correlation coefficients for the test and retest occasions for the mean scores (Table **21**) and the individual items (Table 22).

Table 21. Test-retest Reliability of AGQ Variable Scores Calculated as Means

Variable	All (N = 819)	Group A (N = 204)	Group B (N = 190)	Group C (N = 219)	Group D (N = 206)
Task-approach	.571	.595	.614	.579	.514
Task-avoidance	.538	.574	.582	.543	.465
Self-approach	.513	.508	.493	.592	.469
Self-avoidance	.531	.597	.532	.542	.450
Other-approach	.678	.700	.679	.711	.628
Other-avoidance	.579	.607	.543	.568	.602

There is no clear pattern other than some subscales having higher reliability than others. For example, other-approach goals consistently have the largest correlations and self-approach the smallest. The following table includes Pearson correlation coefficients for each item by group.

Table 22. Test-retest Reliability of AGQ Item Raw Scores by Group

Item	*N*	*r*	*N*	*r*	*N*	*r*	*N*	*r*	*N*	*r*
	All		A		B		C		D	
Task-approach 1	818	.522	207	.578	190	.611	218	.473	203	.436
Task-approach 2	828	.466	208	.481	193	.501	219	.484	208	.414
Task-approach 3	831	.461	209	.473	188	.519	225	.462	209	.406
Task-avoidance 1	813	.477	206	.478	188	.520	215	.470	204	.446
Task-avoidance 2	827	.429	206	.386	191	.465	221	.436	209	.430
Task-avoidance 3	814	.447	206	.570	187	.455	216	.393	205	.376
Self-approach 1	813	.499	207	.545	185	.446	220	.525	201	.466
Self-approach 2	831	.390	203	.353	193	.429	225	.429	210	.356
Self-approach 3	820	.435	206	.490	191	.463	219	.493	204	.298
Self-avoidance 1	834	.414	210	.443	192	.369	223	.506	209	.335
Self-avoidance 2	819	.396	204	.531	189	.418	219	.294	207	.327
Self-avoidance 3	820	.446	204	.491	191	.457	217	.403	208	.425
Other-approach 1	831	.648	208	.673	191	.615	223	.683	209	.621
Other-approach 2	826	.527	208	.596	189	.528	220	.516	209	.478
Other-approach 3	836	.597	210	.619	192	.585	225	.646	209	.543
Other-avoidance 1	830	.548	208	.616	192	.509	222	.580	208	.482
Other-avoidance 2	839	.480	210	.484	194	.428	225	.464	210	.540
Other-avoidance 3	829	.491	205	.504	190	.480	223	.460	211	.518

Note: All correlations are significant to the $p < .01$ level.

Group C had the closest space between the initial and retest offerings, so one would expect the reliability coefficients to be the largest for group C. None of the item reliabilities differed very much from the others. The correlation coefficients between variables fall between .711 and .45. These values are not particularly high, indicating there may be some instability in students' endorsement of achievement goals over the course of an academic term.

Achievement Goals and Achievement. What is the relationship between achievement and achievement goals? Pearson correlations between the course examination scores and AGQ scores are shown in Table **23** (n = 1011). All achievement goal orientation variables are positively and statistically detectably related to course grade and the weighted exam grade ($p < .05$). The approach goals have slightly larger correlations than the avoidance goals.

Table 23. Pearson Correlation Between Grades and AGQ Subscale Scores

Variable	Task-Approach	Task-Avoidance	Self-Approach	Self Avoidance	Other-Approach	Other-Avoidance
Course Grade	.203**	.116**	.017	-.002	.206**	.094**
Weighted Exam	.198**	.103**	.004	-.015	.205**	.082**

Note: **Correlation significant to the .01 level.

The variable that has the strongest relationship to achievement in this sample is other-approach, and the smallest relationship is with self-avoidance. I hypothesized that self-approach would have the strongest relationship to achievement, and this is not the case. Self-avoidance does not have a negative relationship with achievement since it does not show any linear relationship.

Error Orientation

Error Orientation Endorsement

What levels of error competence, learning from errors, error risk taking, error strain, error anticipation, covering up errors, error communication, error motivation, and thinking about errors do organic chemistry students possess? As with the AGQ, students were asked to respond to the adapted error orientation questionnaire (EOQ) twice during the term. The group assignments were the same as for the AGQ. Item four on the error risk taking subscale and three error motivation items (items 2, 3, and 5) were reversed before the analysis.

Response Rate. The first time the EOQ was offered, 1032 responses were submitted from the 1231 students who were invited to participate. Of these, 14 did not have a student ID associated with the submission, 19 failed to respond correctly to the reading check question, 22 were duplicates, and 6 were blank or almost completely blank. For the re-test, 906 submissions were received, 22 had incorrect of missing student IDs, 14 did not correctly answer the reading check question, 13 were duplicates, and 6 were blank or almost completely blank. The responses rates of the EOQ are similar to those of the AGQ.

Table 24. Response Rates for Error Orientation Questionnaire

Test/Retest	Group	Invited	Useable	Response Rate (%)
Test	A	308	250	81.2
Test	B	308	253	82.1
Test	C	308	253	82.1
Test	D	307	253	82.4
Test	Total	1231	1009	82.0
Retest	A	308	223	72.4
Retest	B	308	208	67.5
Retest	C	308	235	76.3
Retest	D	307	227	73.9
Retest	Total	1231	893	72.5

Missing Data. For the initial survey, up to 1.6% of the unique submissions contained at least one blank response. For the retest survey, up to 2.0% of the unique submissions had at least one blank response. Missing item responses were handled in the same way as for the AGQ, since the mechanism of missing data should be the same.

Descriptive Statistics. Means, standard deviations, and internal consistency-reliability (Cronbach's α) of the responses to the EOQ are included in Table **25**. The subscales were defined by Rybowiak et al. (1999) study II, except for the scale I added, which I am calling error motivation.

Table 25. Descriptive Statistics and Internal Consistencies for Initial EOQ

Scale	M	SD	Observed Range	Cronbach's α
Error competence (4)	3.20	.61	1-5	.632
Learning from errors (5)	4.07	.68	1-5	.873
Error risk taking (4)	3.53	.55	1-5	.543
Error strain (4)	2.98	.90	1-5	.765
Error anticipation (5)	3.43	.64	1-5	.672
Covering up errors (5)	2.19	.79	1-5	.755
Error communication (4)	3.72	.65	1-5	.559
Thinking about errors (4)	3.73	.72	1-5	.823
Error motivation (5)	3.02	.51	1-5	.559

Note: The number of items in each subscale is shown in parentheses.

The largest mean is for learning from errors. The scale for these items asked students to indicate the extent the item applied to them, 1 was “not at all”, 2 was “a bit”, 3 was “neither a bit nor a lot”, 4 was “a lot”, and 5 was “totally”. Thus, an average above four indicates the data is centred on students thinking that they do these things “a lot” for this course. Other highly scoring subscales were error communication and thinking about errors. Covering up errors was the least endorsed subscale. Students had relatively low levels of reported error motivation, strain, and covering up errors. Table **26** shows the correlations among the scales.

Table 26. Correlations of Mean Subscale Scores for Initial EOQ

Scale	2	3	4	5	6	7	8	9
1. Error competence	.420**	.255**	.087**	.010	.007	.129**	.568**	.356**
2. Learning from errors	---	.593**	-.181**	.261**	-.335**	.275**	.521**	.486**
3. Error risk taking		---	-.301**	.177**	-.393**	.115**	.373**	.544**
4. Error strain			---	.280**	.516**	.020	-.085**	-.528**
5. Error anticipation				---	.129**	.162**	.089**	-.124**
6. Covering up errors					---	-.152**	-.171**	-.461**
7. Error communication						---	.209**	.110**
8. Thinking about errors							---	.475**
9. Error motivation								---

Error competence is highly correlated with thinking about errors (r = .586), learning from errors is highly correlated with error risk-taking and thinking about errors, and error risk taking is highly correlated with error motivation, and error strain is significantly related to covering up errors, and negatively related to error motivation.

To further explore the factor structure of the EOQ responses, I calculated the Pearson correlations for each pair of items. The correlation tables are included in Appendix F.

Model Validity. What evidence is there of validity for inferring students' view of errors as measured by the adapted Error Orientation Questionnaire? I hypothesized that the EOQ data would fit the best model published by Rybowiak (1999). Rybowiak followed the method suggested by Gerbing and Hamilton (1996), as they first performed an exploratory factor analysis then conducted confirmatory factor analysis with cross validation. The results of Rybowiak et al. (1999) show items uniquely loading on the latent variables described as competence, learning from errors, error risk taking, error strain, error anticipation. Five of nineteen items load on more than one latent variable. Following study I, the researchers tested items for two additional scales: error communication and thinking about errors. Although they did not publish the final CFA model, their correlation table suggests the two constructs did not overlap with the others. I chose to begin with a simpler model, with the suggested items loading onto the one latent variable. The fit statistics for the initial test and retest data are shown in Table **27**. Each student responded to the questionnaire twice, separated by 2-5 weeks.

Table 27. CFA Fit Statistics for Hypothesized 9-Factor Model for the Initial and Retest EOQ Data

Model	NPAR	Scaled Chi-Square	DF	GFI	CFI	RMSEA
Initial Test	116	1168.571 ($p < .0001$)	704	0.833	0.987	0.0649
Retest	116	1026.245 ($p < .0001$)	704	0.822	0.993	0.0678

According to the same criteria used for the AGQ data, the 9-factor models represented in Figures 6 to 9 reasonably fit the EOQ responses. Thus, the data from the EOQ in this study has a similar factor structure to that published version of the instrument. The adaptations of items to a learning context did not change the factor structure.

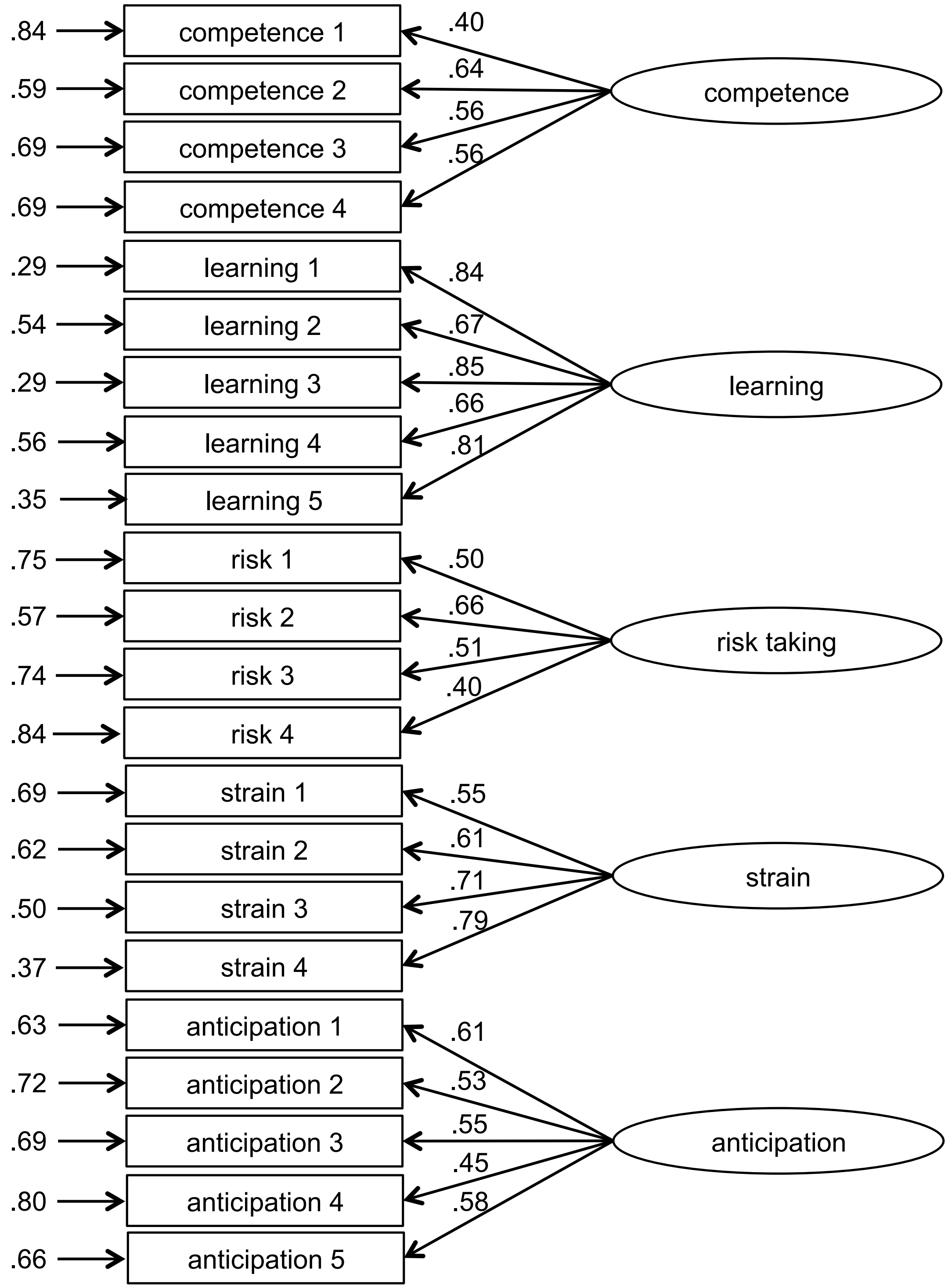

Figure 6. Part of the confirmatory factor analysis model for the initial test error orientation questionnaire data.

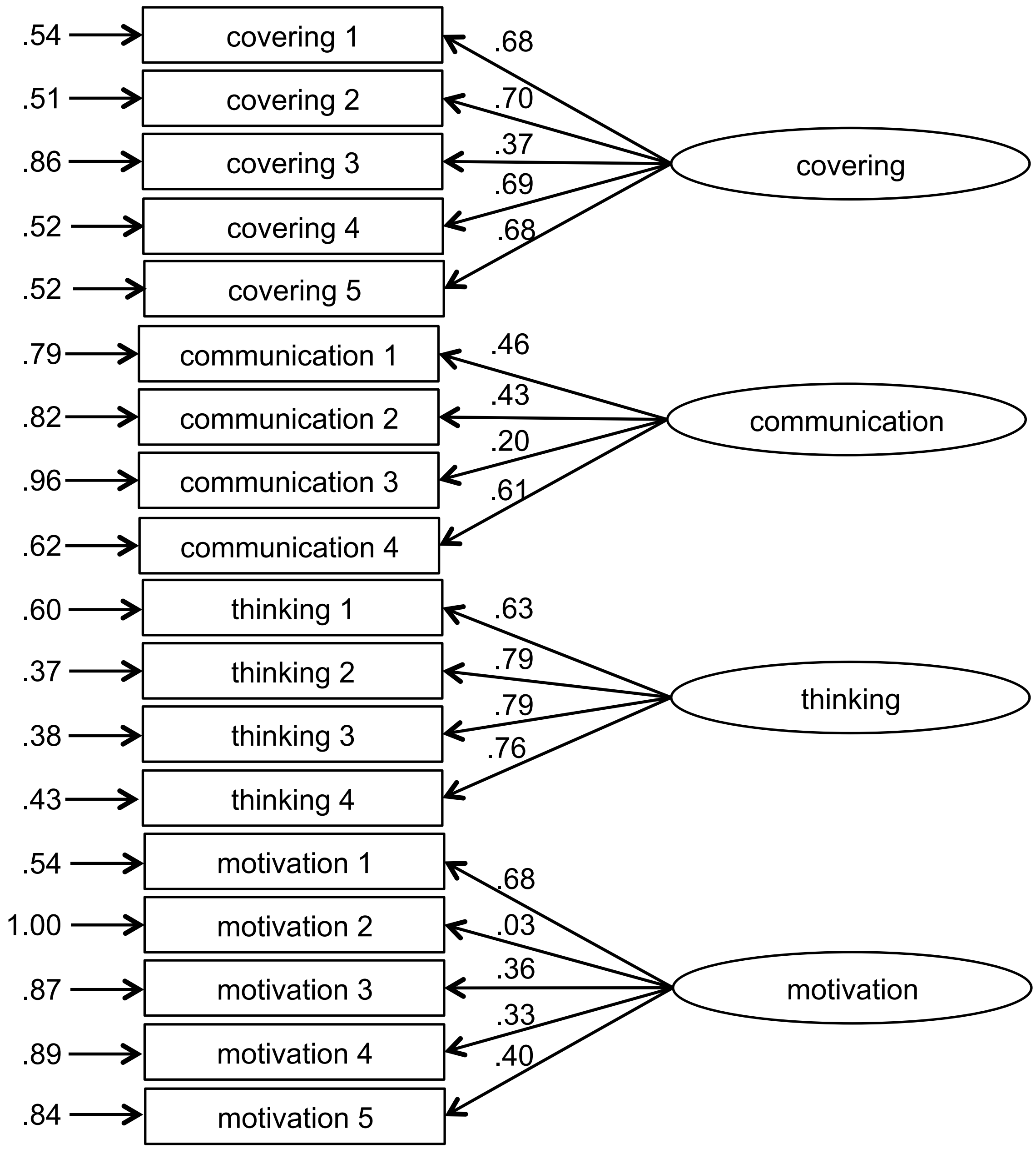

Figure 7. Part of the confirmatory factor analysis model for the initial test error orientation questionnaire data.

Table 28. Latent Variable Covariances for Initial EOQ

Variable	1	2	3	4	5	6	7	8	9
1. Error competence	---								
2. Learning from errors	.610	---							
3. Error risk taking	.549	.839	---						
4. Error strain	-.147	-.236	-.320	---					
5. Error anticipation	.028	.276	.356	.388	---				
6. Covering up errors	-.034	-.395	-.455	.702	.179	---			
7. Error communication	.704	.807	.753	-.194	.322	-.562	---		
8. Thinking about errors	.820	.602	.621	-.136	.082	-.201	.755	---	
9. Error motivation	.877	.785	.88	-.404	-.014	-.446	.861	.975	---

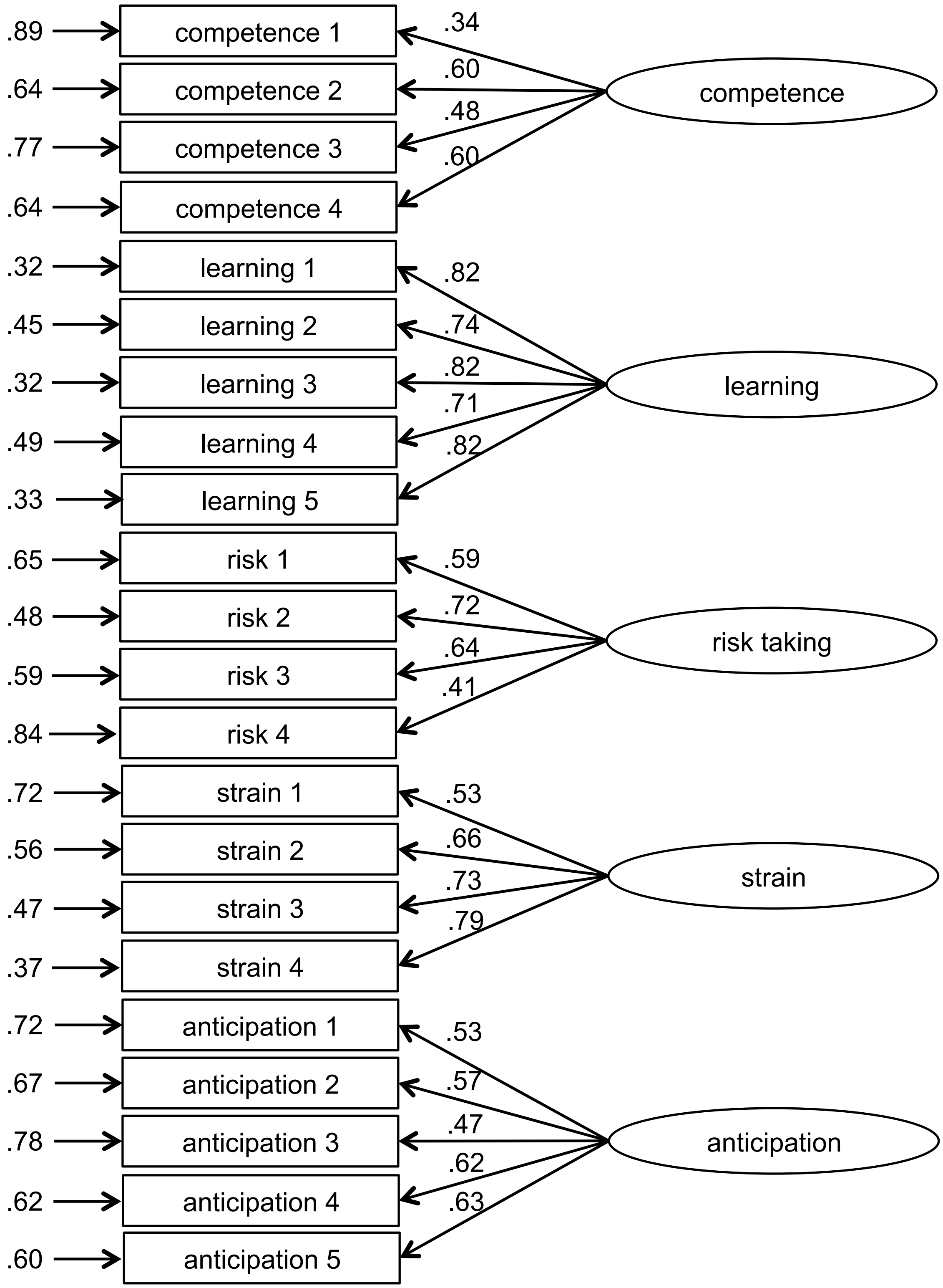

Figure 8. Part of the confirmatory factor analysis model for the retest error orientation questionnaire data.

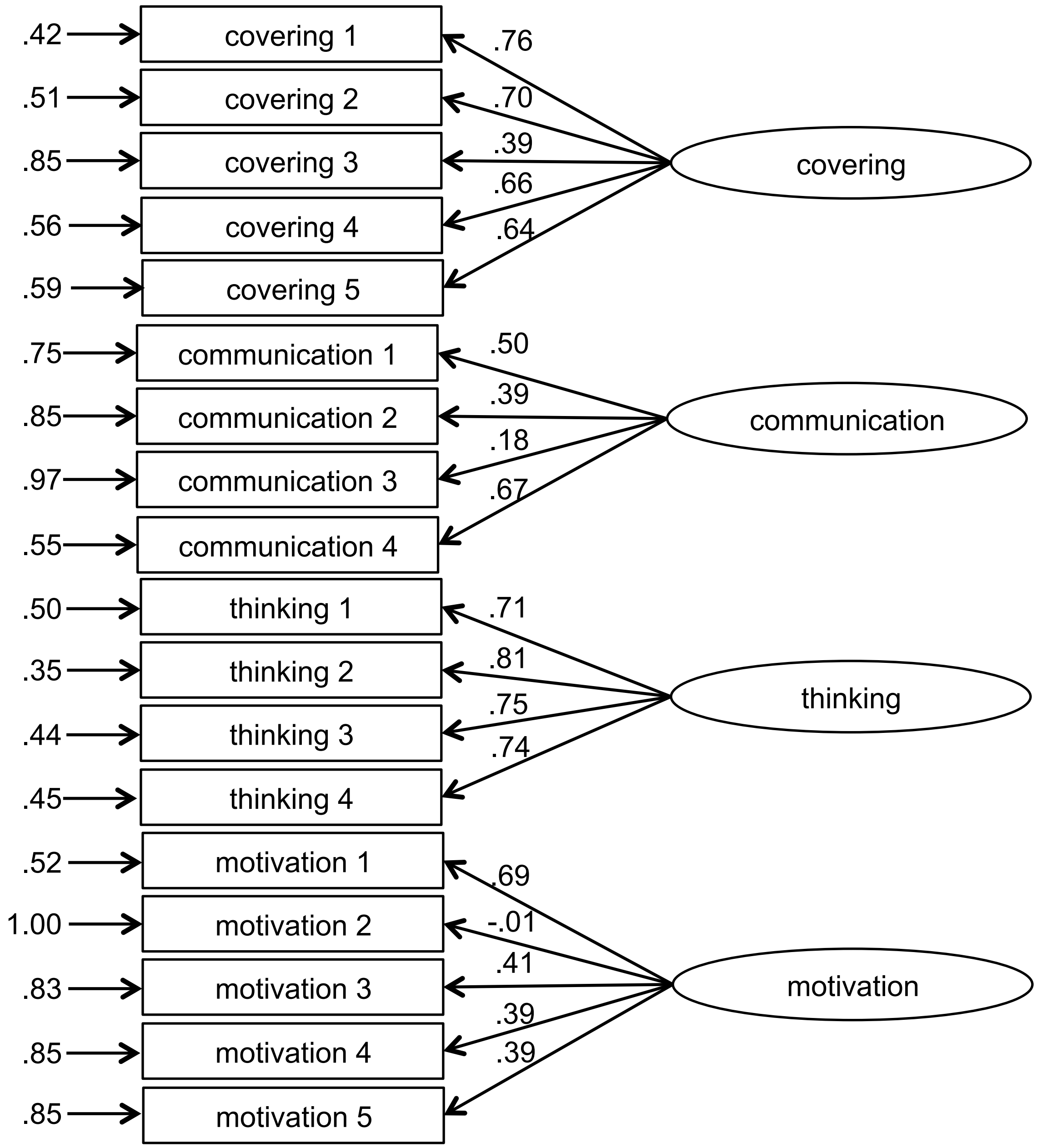

Figure 9. Part of the confirmatory factor analysis model for the retest error orientation questionnaire data.

Table 29. Latent Variable Covariances for Retest EOQ

Variable	1	2	3	4	5	6	7	8	9
1. Error competence	1								
2. Learning from errors	.663	1							
3. Error risk taking	.586	.878	1						
4. Error strain	-.181	-.258	-.267	1					
5. Error anticipation	.183	.497	.622	.134	1				
6. Covering up errors	-.073	-.412	-.444	.647	-.067	1			
7. Error communication	.738	.845	.874	-.191	.539	-.516	1		
8. Thinking about errors	.911	.702	.669	-.128	.295	-.217	.83	1	
9. Error motivation	.951	.816	.839	-.319	.326	-.443	.979	.99	1

Achievement and Error Orientation. Do higher achieving students have higher levels of learning from errors and thinking about errors? And, do lower achieving students have higher levels of error strain and covering up errors? Following the method used by Rybowiak et al. (1999), means for each subscale were calculated and used to explore relationships among error orientation subscales and other variables. The relationship between EOQ subscale scores and course exam performance is described in Table **30** (*n* = 989). The students who completed the course and the initial EOQ survey are included in this analysis.

Table 30. Relationship Between EOQ Scores and Weighted Exam Grade

Subscale	Pearson Correlation Coefficient with Weighted Exam
1. Error competence	.199**
2. Learning from errors	.125**
3. Error risk taking	.134**
4. Error strain	-.162**
5. Error anticipation	-.156**
6. Covering up errors	-.132**
7. Error communication	-.062
8. Thinking about errors	.217**
9. Error motivation	.269**

Note: **Correlation significant to the 0.01 level.

The relationship between endorsement of error orientation items and course grades are in the expected direction. Do higher achieving students have higher levels of learning from errors and thinking about errors? Yes, error competence, learning from errors, error risk taking, thinking about errors, and error motivation have positive statistically detectable correlations with the weighted exam grade. Do lower achieving students have higher levels of error strain and covering up errors? Yes, error strain, anticipation, and covering up errors have negative statistically detectable correlations with the weighted exam grade.

To explore how students' perception of learning from errors relates to achievement, students were grouped into equally sized quartiles of weighted exam scores. The quartiles are defined by weighted exam scores between zero and 50.8% (n = 300), 50.8% and 63.0% (n = 300), 63.0% and 75.8% (n = 300), and above 75.8% (n = 301). Analysis of variance indicates that there is a significant effect of weighted exam score on the levels of learning from errors at the $p < 0.05$ level, $F(3,985) = 6.678$, $p < 0.001$. Posthoc comparisons using the Tukey HSD test indicated that the mean score for the lowest weighted exam bin (M = 3.92, SD = .704) was detectably different from the two highest score bins of weighted exam (3, M = 4.09, SD = .681; 4, M = 4.18, SD = .631). Means plots help to visualize non-linearity.

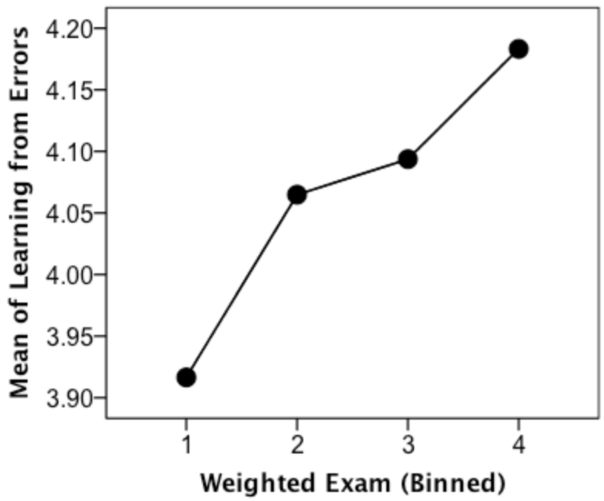

Figure 10. Means plot from the analysis of variance for the learning from errors subscale on the EOQ, by weighted exam quartile.

Achievement Goals and Error Orientation. What relationships exist between achievement goals and error orientation? The relationship between AGQ and EOQ subscale scores is shown in Table **31** (*n* = 980). The students who completed both questionnaires, and completed the course, are included in this analysis.

Table 31. Relationship Among EOQ and AGQ Subscales

Scale	Task-Approach	Task-Avoidance	Self-Approach	Self-Avoidance	Other-Approach	Other-Avoidance
Error competence	.189**	.115**	.119**	.040	.227**	.097**
Learning from errors	.265**	.191**	.170**	.111**	.083**	.046
Error risk taking	.241**	.137**	.138**	.069*	-0.01	-.038
Error strain	.004	.067*	.085**	.155**	.090**	.218**
Error anticipation	.001	.045	.094**	.119**	-.113**	.009
Covering up errors	-.154**	-.089**	-.075*	-.033	.031	.084**
Error communication	.216**	.209**	.228**	.159**	.119**	.088**
Thinkig about errors	.282**	.244**	.201**	.129**	.191**	.108**
Error motivation	.212**	.093**	.072*	-.001	.068*	-.061

None of the correlation coefficients are particularly large, although learning from errors and thinking about errors seem related to task-approach goals. Self-avoidance is weakly correlated with error strain.

Are error orientation levels affected by major course events? To answer this question, I calculated a measure of test-retest reliability for the entire sample and divided into groups (Table **32**).

Table 32. Pearson Correlations for Initial and Retest EOQ Mean Scores by Group

Subscale	All (N = 811)	A (N = 204)	B (N = 186)	C (N = 215)	D (N = 206)
1. Error competence	.536	.579	.425	.606	.534
2. Learning from errors	.579	.579	.519	.623	.589
3. Error risk taking	.574	.612	.524	.576	.578
4. Error strain	.647	.625	.709	.676	.585
5. Error anticipation	.533	.528	.531	.574	.496
6. Covering up errors	.584	.577	.658	.606	.493
7. Error communication	.525	.521	.531	.612	.440
8. Thinking about errors	.611	.595	.593	.612	.645
9. Error motivation	.628	.635	.611	.626	.637

Note: All correlations are significant to the p = .01 level.

Students in group C responded to the questionnaires the closest together in time, so one would expect stronger correlations for group C, which is not the case. Group B responded to the questionnaires the most spaced out in time, and while some of the correlations are lower, it is not true for all the subscales.

Approach to Online Homework Tasks

Out of the 1201 students who completed the course, 1156 created a Sapling Learning account. Thus, the variables described below have values for only these 1156 students. Many student-software interactions in Sapling Learning are recorded. Typically, a student opens an assignment, navigates to a question, inputs a response, and clicks “Check Answer” for correctness and qualitative feedback. Over the course of the term, students were assigned 15 sets of problems spanning a total of 283 questions.

Of those questions, the last 76 contained two optional judgments of knowledge/learning. One question requested that students rate the probability they would answer a similar question correctly on an exam, which is related to confidence. The other was about the certainty in their confidence prediction. Students voluntarily responded to the confidence and certainty judgments and their responses did not affect their grade. Technical constraints meant that they needed to respond to both or neither of the judgments of learning, otherwise the question would be marked incorrect. This section describes some of the data types that were mined from the Sapling Learning system.

Attempts. As soon as a student opens a question in the browser, they are given "zero" attempts for that question. That is, viewing a question but not submitting a solution is tracked as zero for that question. The number of attempts a student can make on each question is unlimited. The average number of attempts for each student was calculated based on the questions students submitted a response to. That is, the views with no attempts were not included. The distribution of attempts is shown in Figure 39 in Appendix D.

Homework and Item Scores. The point value awarded to a student for each item is highly related to the number of attempts, since students lose 5% of a point for each incorrect attempt. It becomes more complicated for multi-part questions, in which students may be given partial credit. Each Sapling Learning question had a maximum of one point. The distribution for the average item score for all attempted questions (one or more attempts) is shown in Figure 40 in Appendix D. The percent score, taking into account all 283 questions, is described by the variable Sapling percent and was described above (Figure 34). For some students, their average item scores and Sapling score are equivalent, but for students who did not submit many questions, these scores differ. There are a variety of factors that cause a student not to complete an assignment or question, but automatically assigning a score of zero for these questions may artificially underestimate their knowledge.

The average item score data is very negatively skewed and leptokurtotic. Much like the percent score, the scores are not particularly interpretable due to the complex nature of the partial credit. After testing three transformations (square root of K-X, log10 of K-X, and inverse of K-X; where K is a constant one larger than the highest value of the variable and X is the variable score), I selected the logarithmic transformation to deal with the substantial skewness. The distribution of the transformed variable, transformed average item score, is shown in Figure 41 in Appendix D.

Giving Up. At anytime, students may "give up" and select "View Solution" to see the correct solution to the problem. If they give up, they can no longer submit answers to the question. Not all "give up" events have zero scores because the student may have earned partial credit on the problem before giving up. In cases like this, students are awarded this fraction of a point in the item score. For this study, I use the percent of questions a student gives up on that they have attempted at least one time. Most students do not give up on questions, but approximately 200 gave up an average of 1-10 questions. Giving up on questions was not very common.

Viewing Hints. Twelve of the questions offered during the study did not have hints, leaving 272 questions with hints. Generally, hints are suggestions about how to get started solving a question. Students had the option of viewing the hint, and viewing it did not affect their score for the question. The variable Hints Viewed describes the percentage of questions attempted for which students selected to view the available hint. Unfortunately, the software does not identify when the student viewed the hint, before or after the first attempt.

Learning from Errors. I constructed a measure of learning from errors from the available metrics in Sapling Learning. If a student submits an incorrect first attempt, and then a correct second attempt, they may have learned from the error. At least, they learned enough in between the first and second attempts that they could submit a correct answer for their second attempt. The variable of correct on second attempt is a percentage based on the questions answered incorrectly for the first attempt. On average, students respond correctly on their second attempt 55% of the time. The distribution of this variable is shown in Figure 44.

Confidence and Certainty Judgments. As described in the previous chapter, students were asked to rate the probability that they would answer a similar question correctly on an exam for 76 items. Also, they were asked to rate the certainty of their prediction. These question-by-question variables are termed confidence and certainty.

The values students provided could range from a probability of 0% (could definitely not solve it) to 100% (could definitely solve it) and a prediction certainty of 1 (unsure), 2 (somewhat sure), or 3 (sure). Invalid responses were omitted from the analysis. For example, some students entered negative or very large values. For confidence, 64 values above 100 were removed, and four values below zero were removed. For certainty, 19 values of zero were removed from the set, and 100 values above three were removed. Most students used whole numbers, but some provided values with decimals. Since this part of the study was entirely voluntary, the dataset is much smaller than it is for the other behaviours, which were generated by Sapling Learning.

Are students who are more confident in their ability more certain of their performance predictions? I calculated the average confidence and certainty judgment for each student. When students answered a question incorrectly, they were prompted to try again and adjust their predictions (i.e. confidence judgments). However, students very rarely changed their judgments. Thus, I only use the confidence judgment data from the first attempt. Student participation in the confidence judgments is shown in Figure 45 (Appendix D), which shows the counts. Since they were completely optional, participation was not high. The distributions of confidence judgments for all questions are shown in Figure 46 (confidence) and Figure 48 (certainty) in Appendix D.

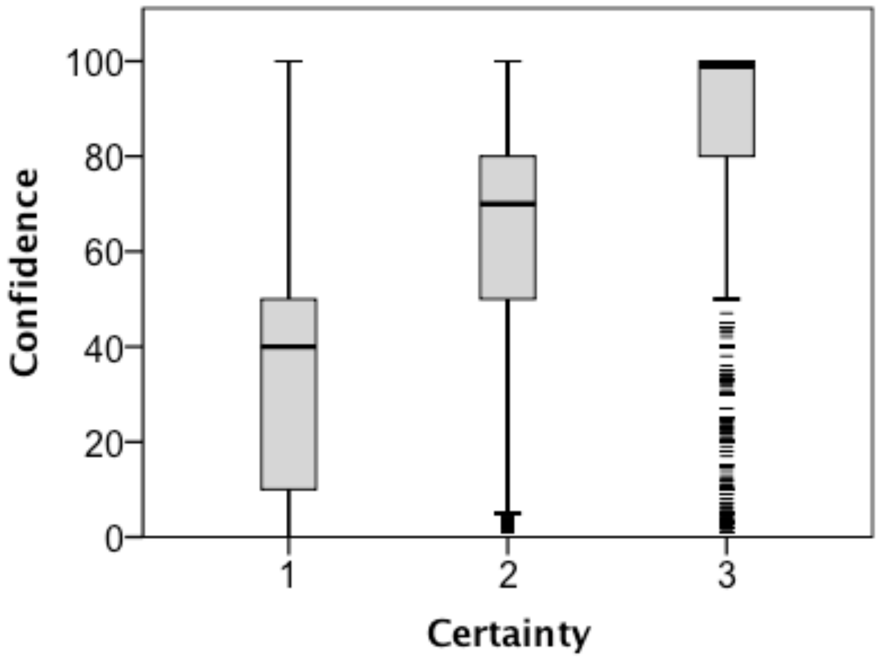

Figure 11. Boxplot of confidence judgments values for each certainty selection.

Figure 11 includes 3712 unsure certainty judgments, 12646 somewhat sure, and 10765 sure confidence judgments. Outliers are shown with small lines, and are unlabeled. This figure shows the strong relationship between students' judgments and their certainty of those judgments. Also, the distribution of confidence judgments may be narrower when students state they are sure about their prediction than when they are unsure.

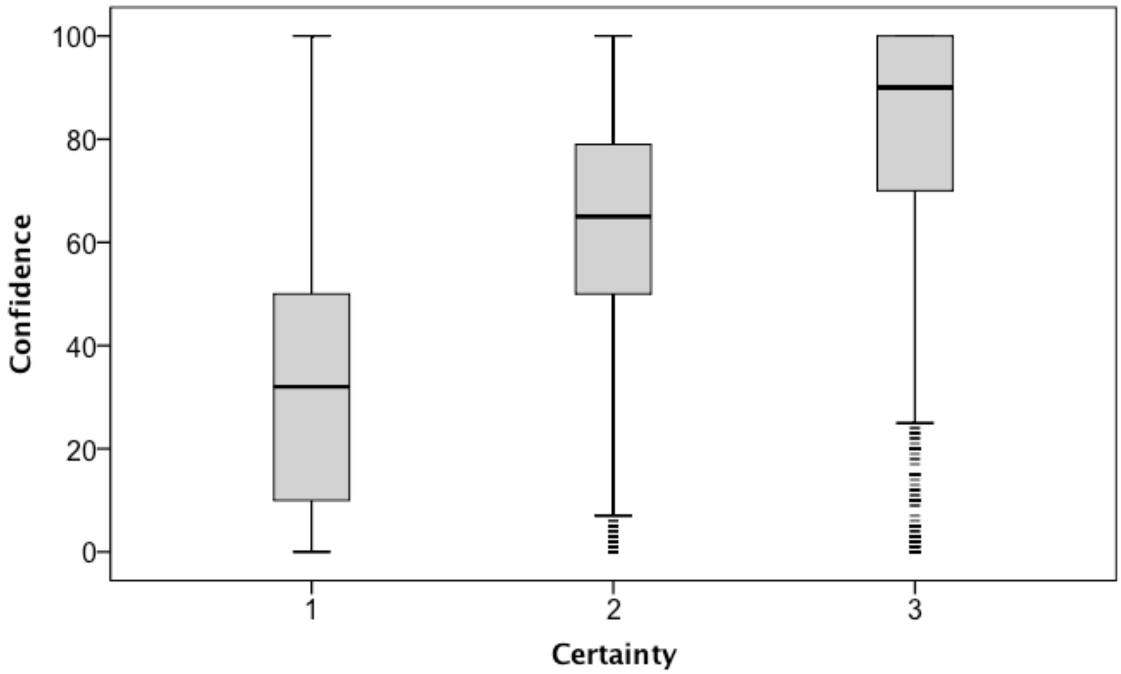

Figure 12. Split-half confidence values based on questions with lower scores.

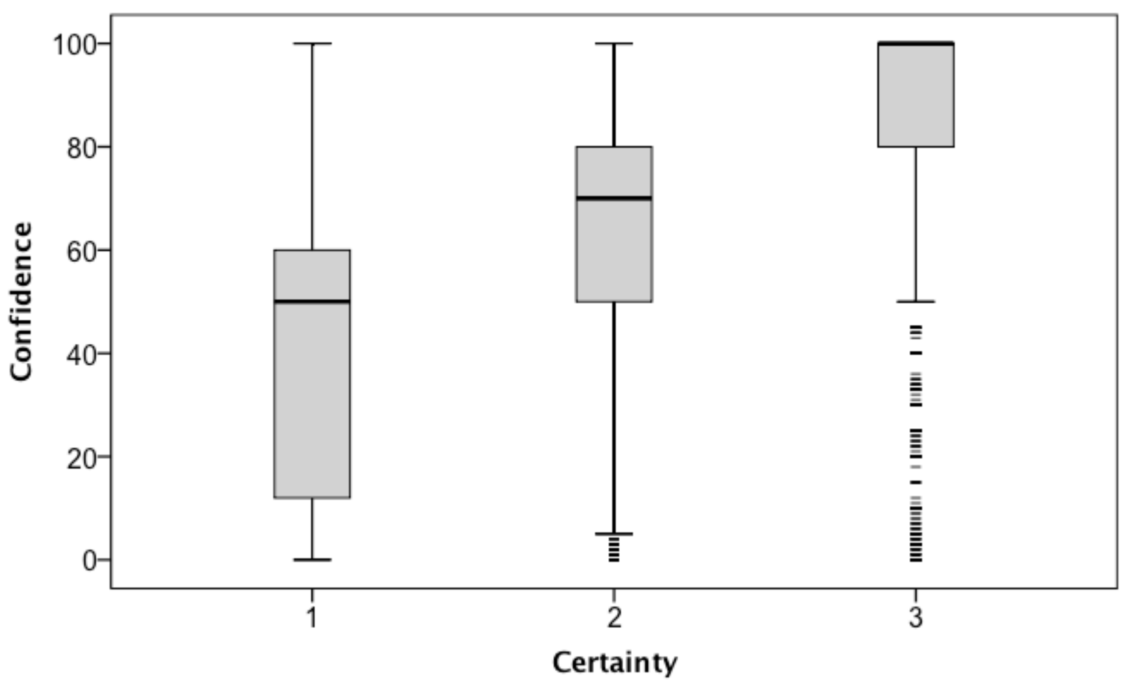

Figure 13. Split-half confidence and certainty values based on questions with higher scores.

Figure 12 and Figure 13 show that there is a wider distribution for confidence values for students who are less certain of their response for easier questions (i.e. questions with higher scores). This is the case no matter what the performance on a question ends up being. However, there is a greater variation of confidence values for easier questions, indicating some students may be under confident.

To explore how response correctness relates to confidence judgments, I created a conditional probability table that includes all available data (Table **33**). This table shows the proportion correct on first attempt for various groups of confidence and certainty levels.

Table 33. Conditional Probability of Correct Response on First Attempt

Confidence	Unsure (1)	Somewhat Sure (2)	Sure (3)
100	0.579 (55)	0.665 (320)	0.723 (3693)
75-99	0.622 (120)	0.641 (2867)	0.717 (2600)
51-74	0.426 (630)	0.533 (3315)	0.600 (554)
26-50	0.366 (186)	0.472 (339)	0.510 (125)
1-25	0.415 (407)	0.556 (380)	0.479 (203)
0	0.409 (187)	0.500 (36)	0.499 (220)

Note: The value in parenthesis is the number of responses (questions × students)

As would be expected, when students predict a confidence value of 100 and a certainty of 3, they had the highest chance of getting that particular question correct (72.3%). When students submit a confidence prediction of 0 and a certainty of 1, 40.9% of them answered that question correctly. Students who are more confident in their ability are more certain of their performance predictions.

Table **33** also shows the commonness of various judgment combinations. For example, confidence predictions between 51 and 75 are often with certainty values of 2. Confidence values above 75 are commonly paired with certainty values of 3. In all confidence levels except for 1-25, the conditional probability of a correct response increases with confidence and certainty predictions.

Continuing with describing the person-level characteristics of confidence judgments, I explored the relationship between students' average judgments, giving up, viewing hints, and exam scores. A series of scatterplots below show the pertinent relationships. For this analysis, I included data from the 310 students who responded to the confidence judgment prompts on 38 or more questions. I selected this half-point cut-off somewhat arbitrarily, but I wanted to use data that had enough confidence judgments that it was reliable. For the confidence ratings, $n = 293$, $r = .310$, $p < .01$. For the certainty ratings, $n = 293$, $r = .184$, $p < 0.01$.

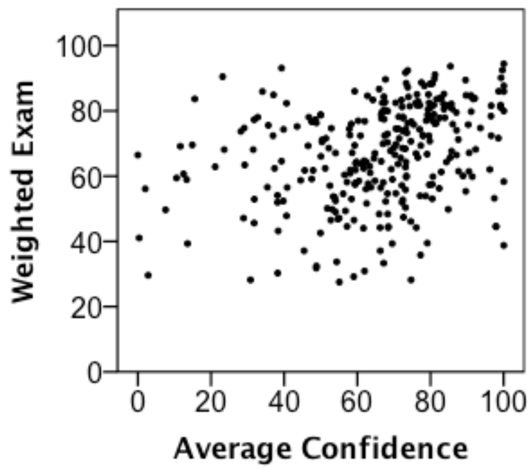

Figure 14. Scatterplot of weighted exam and average confidence judgment for those who responded to at least half (38 or more) of the confidence predictions.

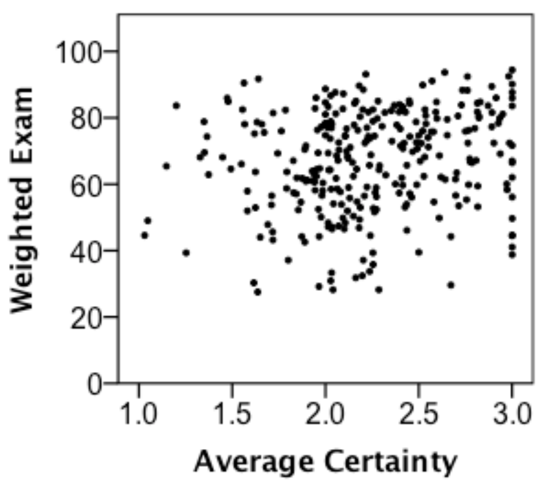

Figure 15. Scatterplot of weighted exam and average certainty judgment for those who responded to at least half of the confidence predictions.

Do students who make higher confidence judgments during online homework sessions achieve higher examination grades than those who make lower judgments? The relationship between exam scores and average confidence judgments is stronger than the relationship between exam scores and average certainty judgments, but both are positively statistically detectable correlated. Thus, students who make higher confidence judgments during online homework practice tend to score higher on course exams (n = 293, r = 0.018, p > .05).

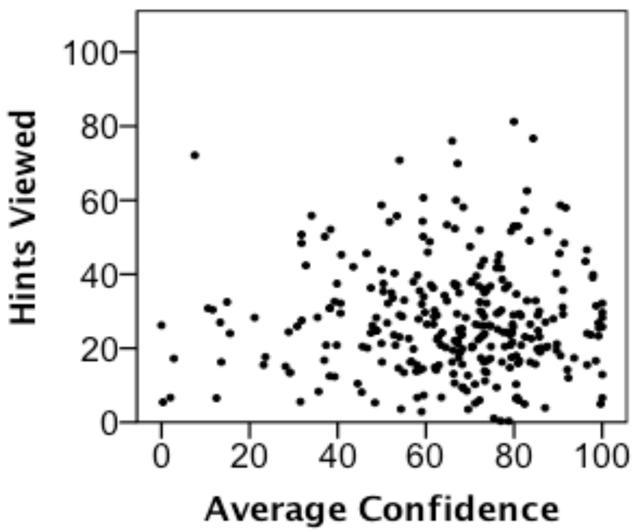

Figure 16. Scatterplot of the percentage of available hints for attempted questions and confidence judgments.

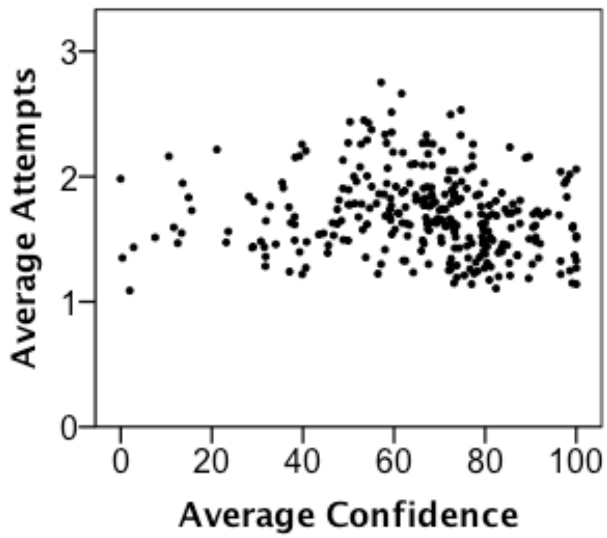

Figure 17. Scatterplot of the percentage of average attempts taken for Sapling Learning questions and confidence judgments.

Figure 17 shows that there is a weak relationship between the average number of attempts a student takes on a question and their confidence level (n = 293, r = –.137, p < 0.01).

What variation exists in how students approach their online homework? The distribution of scores, frequency of hint viewing, frequency of giving up, frequency of answering correctly, number of attempts, and nature of students' judgments of learning provides insight into how each question functioned.

The distributions of attempts and average scores for all questions are included in Table **34** and Figure 18, respectively. Most questions were solved correctly in 1, 2, or 3 attempts, but a considerable number of students attempted questions more than three times.

Table 34. Frequency of Attempts on All Sapling Learning Questions

Number of Attempts	Frequency	Percent
0	18869	6.0
1	177826	56.2
2	68025	21.5
3	25489	8.1
4	11665	3.7
5	5884	1.9
6	3365	1.1
7	1967	.6
8	1200	.4
9	731	.2
10	491	.2
11	296	.1
12	178	.1
13	141	.0
14	87	.0
15	66	.0
16	48	.0
17	37	.0
18	25	.0
19	13	.0
20	7	.0
21	13	.0
22	8	.0
23	3	.0
24	5	.0
25	7	.0
26	4	.0
27	2	.0
28, 29, 31, 32, 36, 37	1 each	.0
Total	316458	100.0

I constructed bar graphs that include the number of students who viewed hints and gave up on each question. Superimposed on this bar graph is a line graph representing the average score of the question. Figure 18 contains this graph for all questions in assignment 15. There is a large variability in how students approach specific questions. Sometimes over half the class views a hint, but other times only a couple hundred make use of the hint. Due to the variability in viewing hints and giving up, it is likely that students are making a conscious decision about whether to view a hint or give up. Otherwise, I would expect the same students to view hints on each question to be very similar. Figures for the other assignments are included in Appendix G.

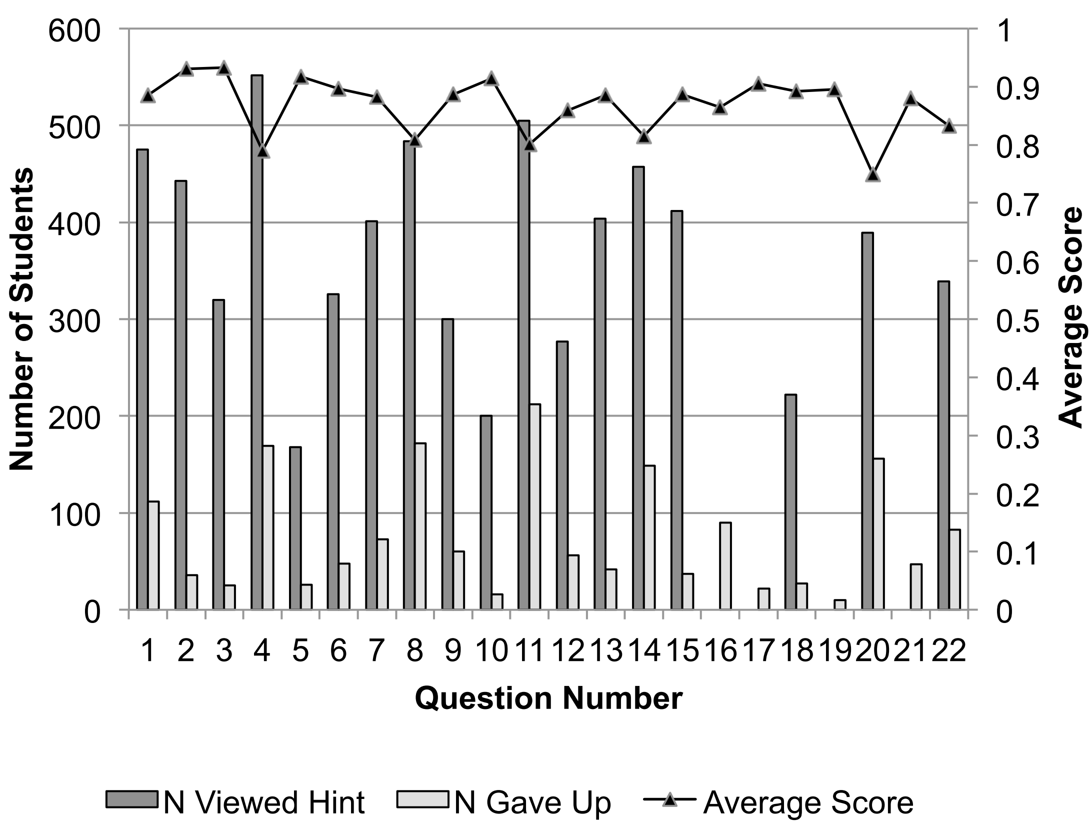

Figure 18. Number of students who viewed hint and gave up on questions from assignment 15 on the left-hand y-axis (bars).

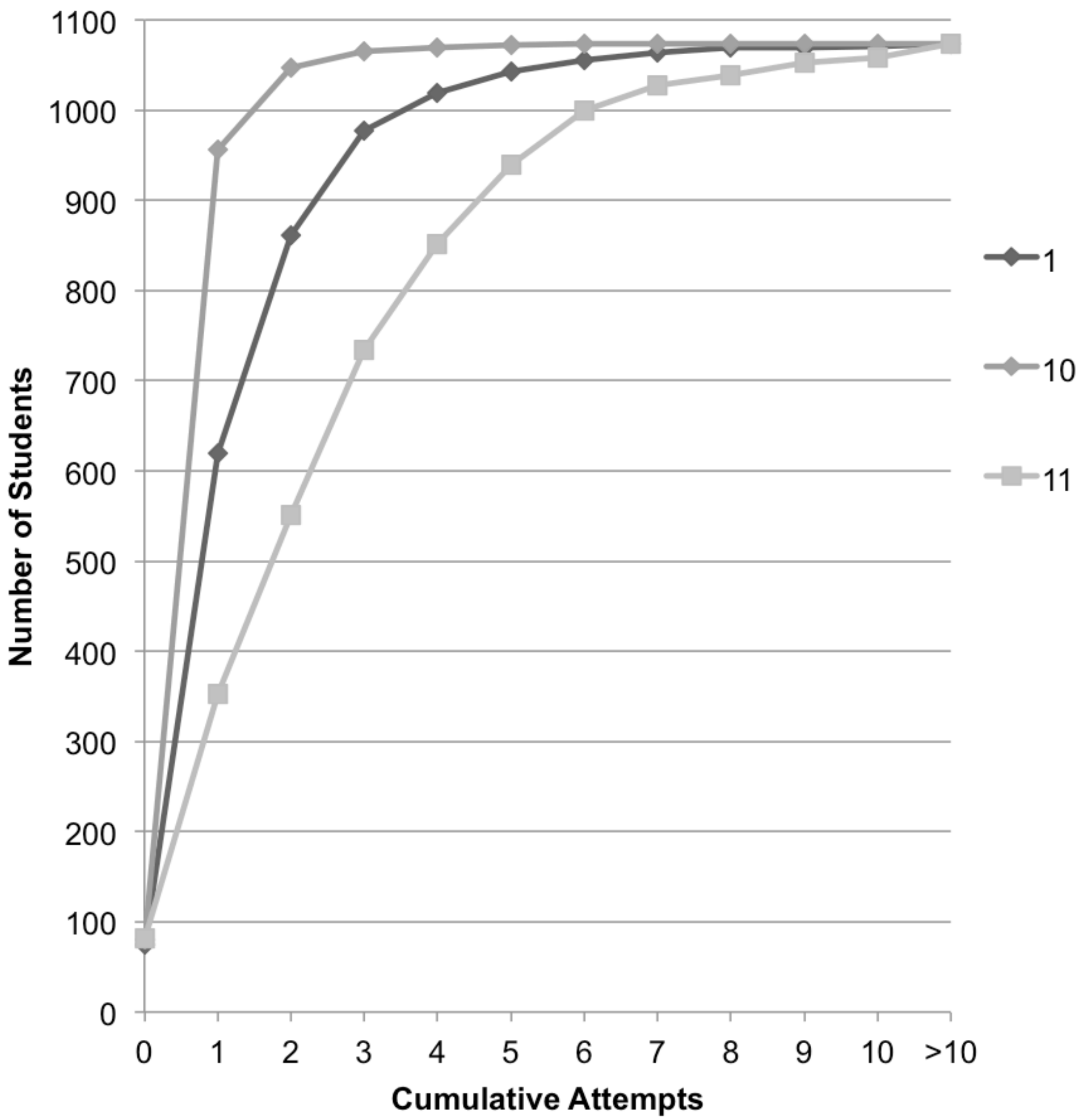

Figure 19. Line graph showing the cumulative attempts taken by students for three questions in assignment 15.

Figure 19 compares three items from the same assignment. For some items, most students obtain the correct answer after two attempts. For others, upwards of eight attempts is required. Figure 20 shows how the number of attempts taken for each question can vary widely. The average number of attempts is also quite variable (Figure 52 to Figure 70 in Appendix G).

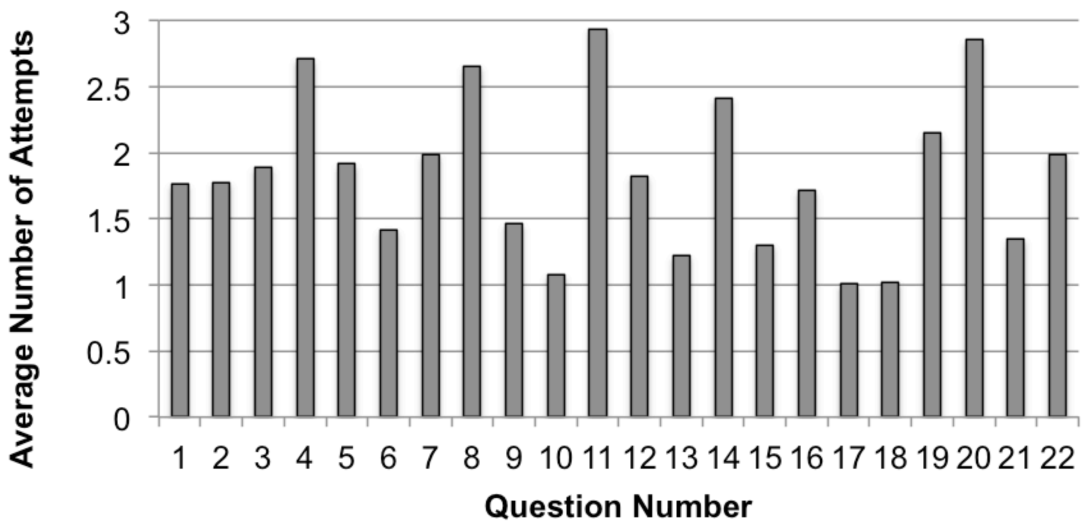

Figure 20. Average number of attempts for questions from assignment 15.

To explore the relationship between confidence judgments and scores for particular questions, I plotted the average score, average confidence, and average certainty for each question that solicited students' judgments (76 questions). Figure **21** to Figure **23** display the relationship between variable pairs.

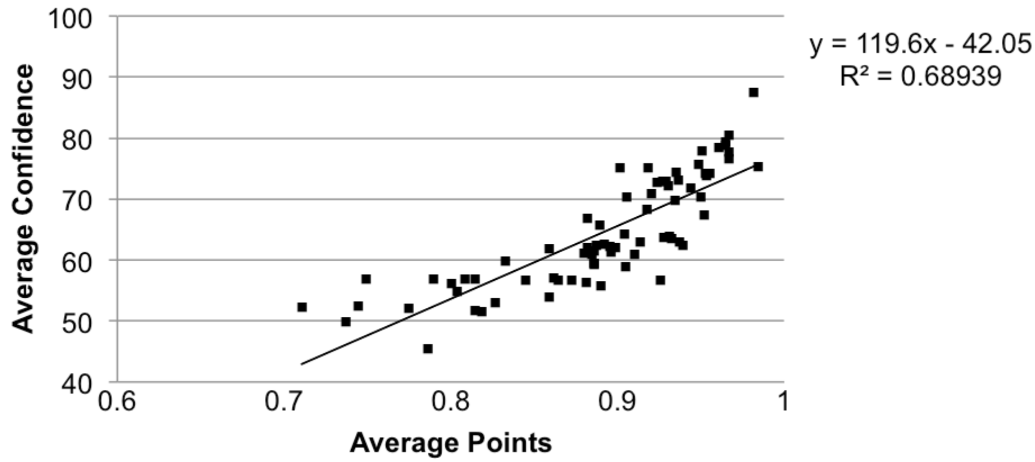

Figure 21. Scatterplot showing the average confidence prediction and average points for the 76 questions that had confidence judgments.

The relationship between the average points for a particular question and the average confidence judgment is almost linear. Figure **21** shows distortion from linearity at low values, so the relationship is somewhat curvilinear.

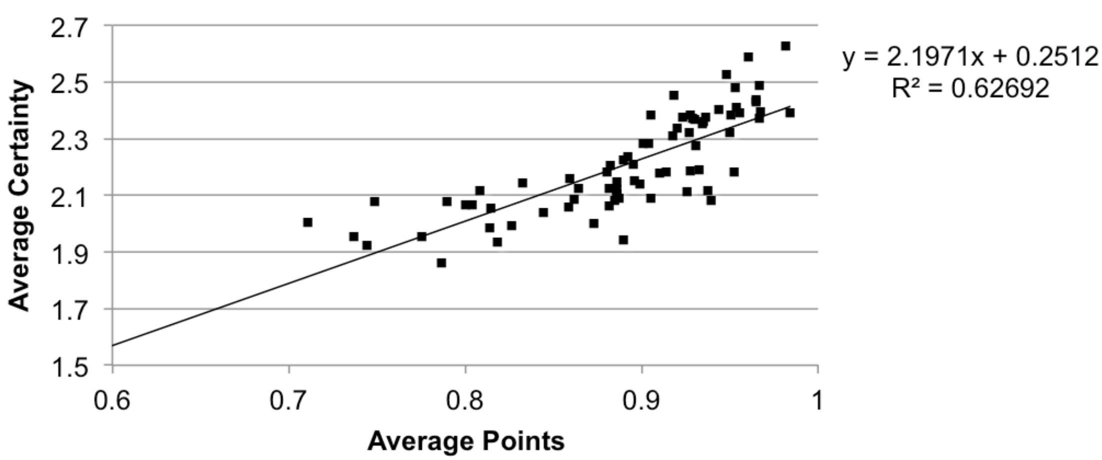

Figure 22. Scatterplot showing the average certainty and points value for the 76 questions that had confidence judgments.

Figure **22** shows a similar relationship as Figure **21**. The certainty of judgment and the average points for a question is curvilinear. The relationship may depart even more from linearity than the confidence judgment. The relationship between the confidence and certainty is linear (Figure **23**).

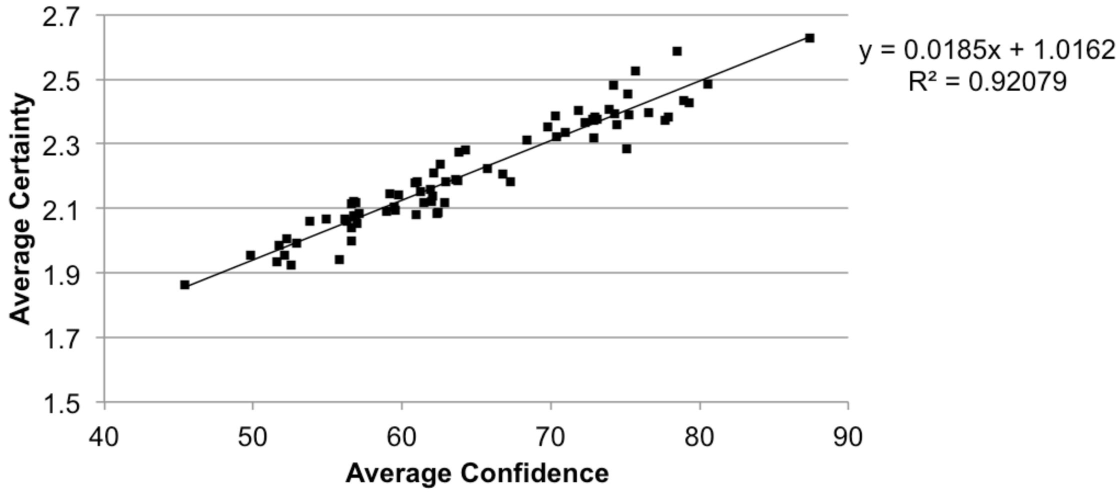

Figure 23. Scatterplot showing the linear relationship between average certainty and average confidence ratings for each question.

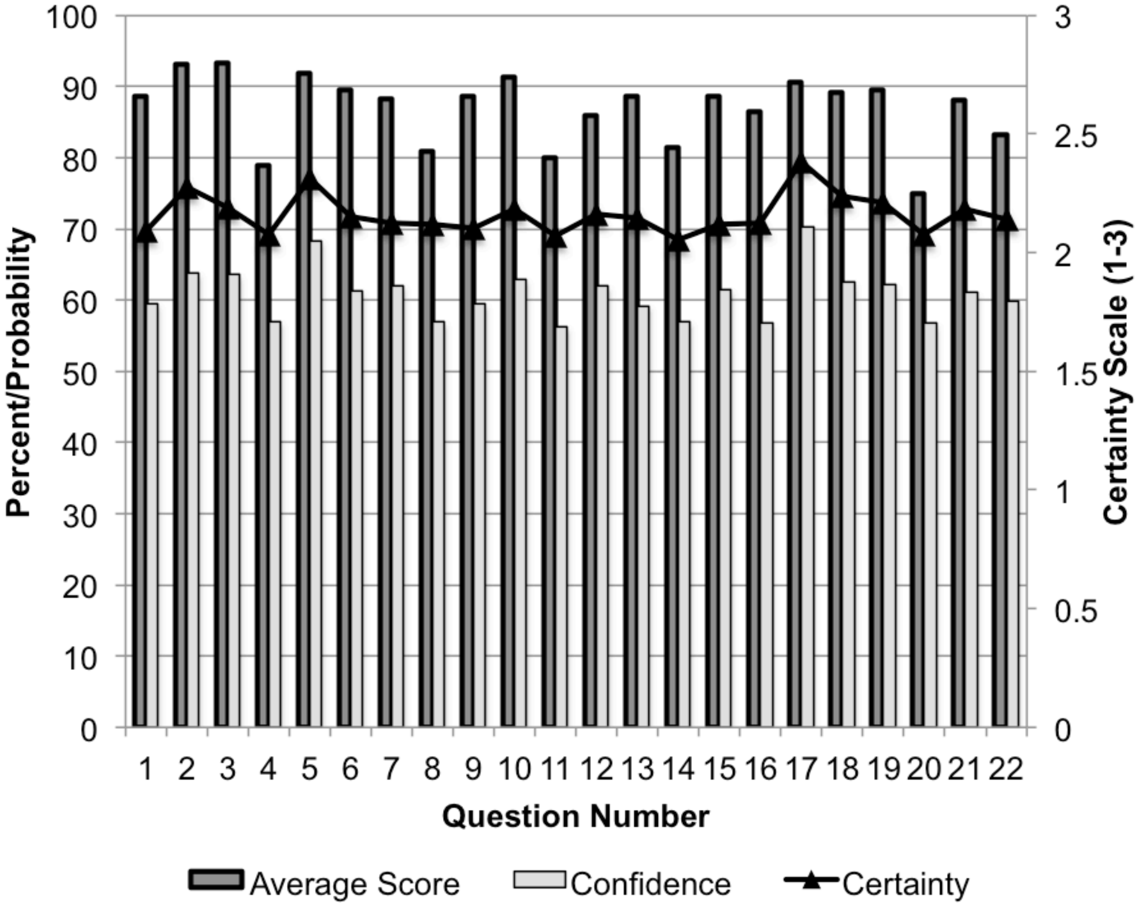

Figure 24. Confidence judgments (bars) and certainty judgments (points) for questions in assignment 15.

Figure **24** represents the relationship between average confidence, average certainty, and average score for each question in assignment 15. The confidence and certainty judgments mirror the question scores fairly closely.

Achievement, Confidence, and Learning Behaviours. To investigate if students with different levels of exam performance have different ability to make accurate predictions, I split the weighted exam score into two equally sized bins (at 63.0%) representing “high” and “low” achievement. The descriptive statistics and bivariate correlation coefficients between key variables are shown in Table **35** and Table **36**.

Table 35. Online Homework Behaviours for High- and Low-Performing Students

Variable	Mean – High Performers (N = 551)	Mean – Low Performers (N = 485	Standard Deviation – High Performers	Standard Deviation – Low Performers
Weighted exam**	75.92	49.87	7.86	9.39
Confidence**	75.83	65.43	20.03	21.85
Certainty**	2.39	2.20	0.48	0.48
Hints Used*	24.91	26.91	13.46	14.96
Gave Up**	2.35	5.48	3.51	6.60
Transformed Average Item Score**	-1.56	-1.22	0.34	0.31
Average Attempts**	1.67	1.94	0.29	0.36
Transformed Sapling Score**	.74	1.11	0.36	0.37
Learning From Errors (Correct on 2nd Attempt)**	60.13	48.91	11.84	11.05

Note: Cases were deleted list-wise for this analysis.
*Difference is significant to the 0.05 level.
**Difference is significant to the 0.01 level.

Higher-performing student are less likely to give up on questions they have attempted, earn higher item and Sapling Learning scores, and take fewer attempts on average. They are also more likely to respond correctly to questions that require a second attempt.

Table 36. Bivariate Correlation Coefficients for High- (Above Diagonal) and Low-Performing (Below Diagonal) Students.

Variable	1	2	3	4	5	6	7	8	9
1. Weighted exam	---	.116**	.084*	-.118**	-.328**	-.515**	-.457**	-.477**	.469**
2. Confidence	.111*	---	.619**	-.122**	-.071	-.042	-.072	-.056	.052
3. Certainty	.093*	.545**	---	-.156**	.019	.056	-.014	0.042	-.001
4. Hints Used	.083	.014	-.094*	---	-.025	-.009	.007	-.057	-.044
5. Gave Up	-.197**	-.150**	-.120**	.034	---	.724**	.414**	.570**	-.572**
6. Transformed Average Item Score	-.264**	-.079	-.127**	.097*	.682**	---	.786**	.778**	-.888**
7. Average Attempts	-.098*	-.027	-.113*	.153**	.237**	.615**	---	.562**	-.826**
8. Transformed Sapling score	-.332**	-.089	-.125**	-.005	.463**	.741**	.359**	---	-.656**
9. Learning from errors (Correct on 2^{nd} Attempt)	.225**	.072	.118**	-.039	-.591**	-.873**	-.702**	-.605**	---

Notes:

Cases were deleted list-wise for this analysis.

*Correlation is significant to the 0.05 level.

**Correlation is significant to the 0.01 level.

Many of the correlation coefficients between variables are similar for the high- and low-performing groups are similar but there are some striking differences (Table **36**). The relationships between weighted exam scores and the transformed average item scores and the average attempts are more positive for high performing students. For low-performing students, the statistically detectable relationships between average certainty and giving up, average item score (transformed), average attempts, and Sapling percent (transformed) are more negative. There is a stronger relationship between weighted exam and learning from errors for higher-performing students than lower-performing students. For higher-performing students, the extent to which they solve a problem correctly on the second attempt is more strongly related to their exam scores. This may suggest that students who have higher exam scores have learned from their errors more so than lower-performing students.

Confidence, Learning from Errors, and Goal Orientation. Do students learn from errors during online homework practice? To explore how students with various levels of task, self, and other-approach and avoidance goal orientations learn from the errors they make while solving online homework problems, I first examined the relationships between AGQ and EOQ survey subscales and online homework behaviours.

Table 37. Pearson Correlation Coefficients among AGQ Subscale Scores and Homework Behaviours

Behaviour	Task-Approach	Task-Avoidance	Self-Approach	Self-Avoidance	Other-Approach	Other-Avoidance
Confidence	.138**	.036	-.009	-.030	.119**	.008
Certainty	.115**	-.001	-.027	-.060	.092**	-.010
Hints Used	.037	.048	.042	.003	-.030	.002
Gave Up	-.072*	-.074*	-.022	-.007	-.105**	-.076*
Average item Score	.051	.025	.040	.005	.055	.021
Transformed Average Item Score	-.120**	-.084**	-.034	-.019	-.155**	-.092**
Attempts Average	-.078*	-.055	.029	.009	-.130**	-.080*
Sapling Percent	.128**	.116**	.072*	.051	.131**	.120**
Transformed Sapling Percent	-.143**	-.132**	-.062	-.051	-.169**	-.129**
Learning From Errors (correct 2^{nd} attempt)	.103**	.067*	.017	.003	.120**	.066*

Notes:
*Correlation is significant to the 0.05 level.
**Correlation is significant to the 0.01 level.

The strongest relationships between AGQ subscale scores and online homework behaviours are between task-approach goals and Sapling Learning scores as well as students' average confidence values. Other-approach goals are also related to confidence and Sapling Learning scores. Self-approach and other-approach subscale scores were detectably related to learning from errors, although the magnitudes of the correlation coefficients are quite small (0.120 and 0.103, $p < .05$).

Confidence, Learning from Errors, and Error Orientation. The relationships among online homework behaviours and the error orientation subscores are shown in Table **38** and Table **39**.

Table 38. Pearson Correlation Coefficients among EOQ Subscale of Competence, Learning, Risk, Strain, and Anticipation

Behaviour	Competence	Learning	Risk	Strain	Anticipation
Confidence	.232**	.178**	.114**	-.149**	-.208**
Certainty	.179**	.125**	.081*	-.117**	-.126**
Hints Used	-.076*	-.010	-.016	-.035	.003
Gave Up	-.058	-.016	-.001	.090**	.106**
Average Item Score	.108**	.049	.083*	-.086**	-.048
Transformed Average Item Score	-.167**	-.060	-.070*	.132**	.118**
Average Attempts	-.093**	-.043	-.041	.102**	.075*
Sapling Score	.095**	.027	.052	-.045	-.076*
Transformed Sapling Score	-.122**	-.046	-.052	.079	.100**
Learning From Errors (Correct on 2nd Attempt)	.180**	.071*	.081*	-.151**	-.087**

Table 39. Pearson Correlation Coefficients among EOQ Subscale of Covering Up, Communication, Thinking, and Motivation

Behaviour	Covering	Communication	Thinking	Motivation
Confidence	-.076*	.018	.193**	.253**
Certainty	-.032	.024	.174**	.178**
Hints Used	-.061	-.021	-.044	.054
Gave Up	.086**	-.136**	-.032	-.098**
Average Item Score	-.107**	.122**	.067*	.113**
Transformed Average Item Score	.134**	-.168**	-.154**	-.185**
Average Attempts	.089**	-.076*	-.112**	-.134**
Sapling Score	-.077*	.125**	.100**	.118**
Transformed Sapling Score	.111**	-.155**	-.140**	-.165**
Learning From Errors (Correct on 2nd Attempt)	-.150**	.168**	.147**	.180**

Hint viewing was not related to any of the subscales on the EOQ. The relationships between Sapling Learning scores and error competence, anticipation, communication, thinking and motivation appear to be substantial. Students' confidence and certainty judgments were related to some, but not all of the EOQ subscales. Covering up and communication of errors do not appear to relate to the confidence and certainty judgments. Giving up was slightly related to error anticipation, communication, and motivation, in that higher frequency of giving up was negatively related to students' scores on these subscales. Learning from errors was associated with higher levels of reported error competence, error communication, thinking, and motivation. Learning from errors was negatively related to error strain and covering up errors.

This analysis does not support the hypothesis that students who view errors as valuable to learning are less likely to give up and view hints than students who do not see errors as valuable for learning.

To examine whether lower-achieving students had measurable levels of error strain and covering up errors, I first examined the correlation coefficient between weighted exam and these subscale scores (r = –.162 for strain and r = –.132 for covering up errors). Students were grouped into quartiles by virtue of their weighted exam scores (at 50.8%, 63.0%, and 75.8%). There was a statistically detectable effect of achievement on the levels of error strain, $F(3,985) = 6.619$, $p < 0.001$. Posthoc comparisons using the Tukey HSD test indicated that the mean score for the highest achievement bin (4, M = 2.77, SD = .871) was significantly different from the other three groups (1, M = 3.14, SD = .882; 2, M = 3.06, SD = .925; 3, M = 2.97, SD = .886), which were not different from each other. Figure 25 contains the means plot which graphically shows the much lower mean for group 4.

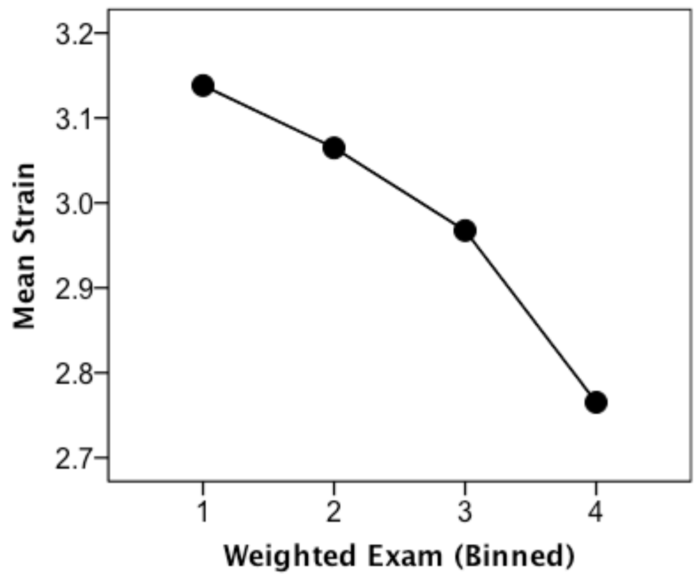

Figure 25. Means plot from the analysis of variance for the error strain subscale on the EOQ, by achievement quartile.

Achievement and Error Orientation. There was also a detectable effect of achievement on covering up errors, $F(3,985) = 4.830$, $p = 0.002$. A posthoc Tukey HSD test indicates that the levels of covering up errors for the lowest achievement group (1, $M = 2.33$, $SD = .827$) are significantly different from 3 ($M = 2.15$, $SD = .780$) and 4 ($M = 2.08$, $SD = .761$), which are detectably different from each other ($p < .05$). The evidence supports the idea that higher-achieving students have lower levels of error strain and covering up errors. Figure 26 contains the means plot which graphically shows the difference by group.

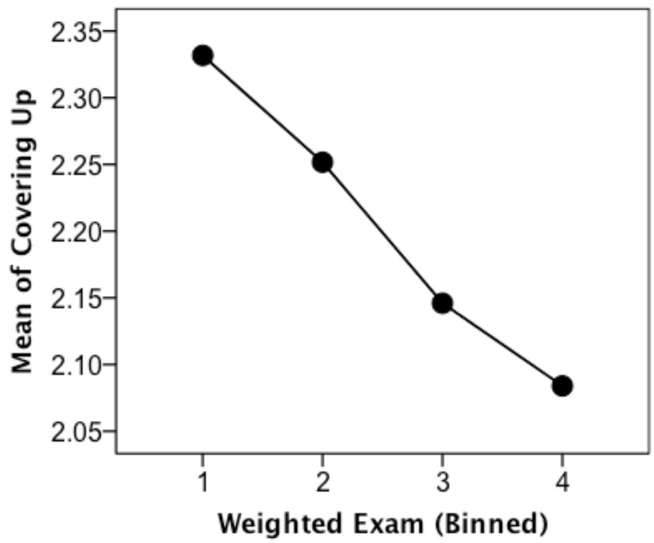

Figure 26. Means plot from the analysis of variance for the Covering Up Errors Subscale on the EOQ, by Weighted Exam quartile.

Do higher achieving students report thinking more about their errors? There was a statistically detectable effect of weighted exam on thinking about errors at the $p < 0.05$ level, $F(3{,}985) = 16.421$, $p < 0.001$. A posthoc Tukey HSD test indicates that the levels of thinking about errors for the highest achievement group (4, $M = 3.94$, $SD = .669$) are measurably different from all the other levels (1, $M = 3.53$, $SD = .704$, $p < .001$; 2, $M = 3.63$, $SD = .712$, $p < .001$; and 3, $M = 3.75$, $SD = .692$, $p = .009$). Additionally, 1 and 3 are detectably different ($p = .003$).

The evidence supports the idea that higher-achieving students report higher levels of thinking about the errors they make while solving problems. Figure 27 contains the means plot which graphically shows how the means differ by group.

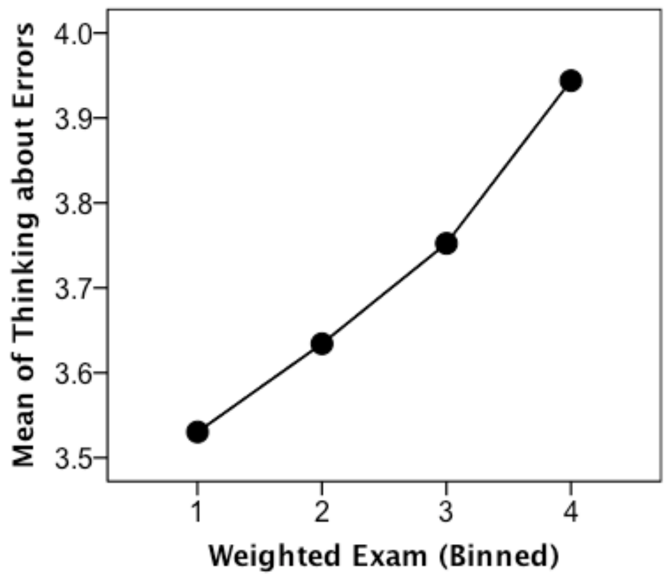

Figure 27. Means plot from the analysis of variance for the Thinking About Errors Subscale on the EOQ, by weighted exam quartile.

Achievement and Learning from Errors. Do higher-achieving students better learn from their errors? To address this question, I conducted an analysis of variance of the correct 2nd attempt variable with weighted exam groups. This analysis showed a significant difference between exam achievement on the learning from error (correct 2nd attempt) scores, $F(3,1152) = 134.020$, $p < .001$. A posthoc Tukey test showed that all groups were statistically detectably different from each other (Table **40**).

Table 40. Descriptive Statistics of Learning from Errors by Weighted Exam Group

Weighted Exam Group	Mean	Standard Deviation
1	45.79	13.95
2	50.23	11.02
3	55.34	10.49
4	64.49	11.59
All	54.11	13.70

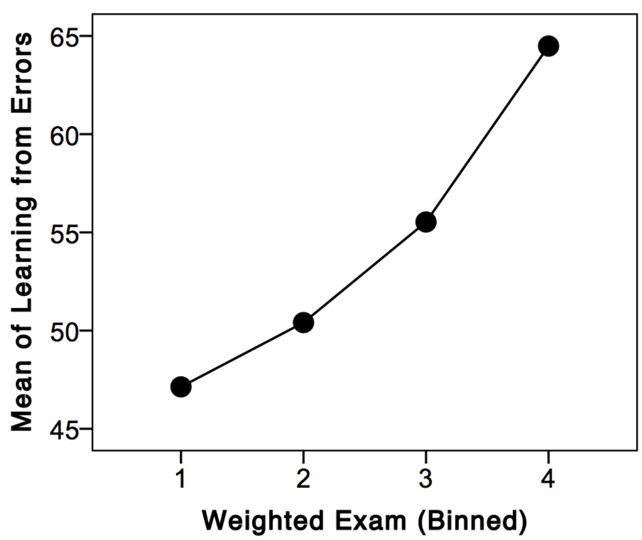

Figure 28. Means plot showing the different means of learning from errors for students with different exam scores.

Learner Profiles. Are there distinct profiles of learners based on how they interact with their online homework? I used cluster analysis to create profiles of learners according to their online learning behaviour. The first cluster analysis I performed used the variables hint viewing, giving up, and attempts. I included students who left 100 or fewer questions blank (N = 1128). Six clusters seemed appropriate after inspecting the plot from a hierarchical cluster analysis using Ward's method (squared Euclidean distance) (Burns & Burns, 2009). I then used k-means analysis to create the cluster variable for each case. Table **41** describes the means for each of the three variables after 34 iterations.

Table 41. Approaches to Online Homework Cluster Centres

Variable	1 N = 216	2 N = 60	3 N = 91	4 N = 344	5 N = 307	6 N = 110
Times gave up	2	14	15	2	2	4
Hints viewed	37	35	17	24	11	54
Average Attempts	1.79	2.09	2.01	1.82	1.69	1.81

Students in clusters 2 and 3 gave up more often than students in the other clusters. These groups also had higher average number of attempts. Most students are in clusters 1, 4, and 5, all of which do not have high levels of giving up but vary in hint use. Cluster 6 has the highest use of hints.

How do these groups of students perform on course examinations and online homework? The following tables include descriptive statistics for exam performance, Sapling performance, judgments of learning, and learning from errors grouped by cluster.

Table 42. Examination Scores by Online Homework Behaviour Cluster

	Cluster	N	Mean	Standard Deviation	Minimum	Maximum
Midterm 1 Percent	1	216	63.04	13.00	25	100
	2	60	52.11	10.29	28	83
	3	91	55.88	12.53	21	84
	4	344	62.71	13.45	32	95
	5	305	65.13	13.62	24	96
	6	110	60.99	12.11	35	91
	Total	1126	62.14	13.46	21	100
Midterm 2 Percent	1	214	58.76	17.37	10	92
	2	59	47.44	14.18	12	80
	3	88	43.30	16.73	8	80
	4	341	59.49	17.84	8	96
	5	304	59.94	19.25	10	98
	6	108	52.93	16.75	12	90
	Total	1114	56.92	18.47	8	98
Final Exam Percent	1	216	67.80	15.95	28	97
	2	60	53.00	13.92	19	83
	3	91	54.40	16.51	18	87
	4	344	68.10	17.09	21	98
	5	307	67.66	18.39	16	99
	6	110	64.30	14.87	29	93
	Total	1128	65.64	17.47	16	99
Weighted Exam	1	216	65.20	14.48	28	93
	2	60	51.70	11.80	24	75
	3	91	52.61	14.39	22	81
	4	344	65.49	15.40	24	95
	5	307	65.65	16.72	17	95
	6	110	61.56	13.56	30	90
	Total	1128	63.32	15.81	17	95

Figure 29 shows how the standardized midterm and exam scores for students in each cluster vary. Groups 1, 4, and 5 performed similarly and consistently on the course examinations. Groups 2 and 3 performed worse, and group 6 was in between. Group 6 was characterized by viewing hints often. Analysis of variance results are included in Appendix H.

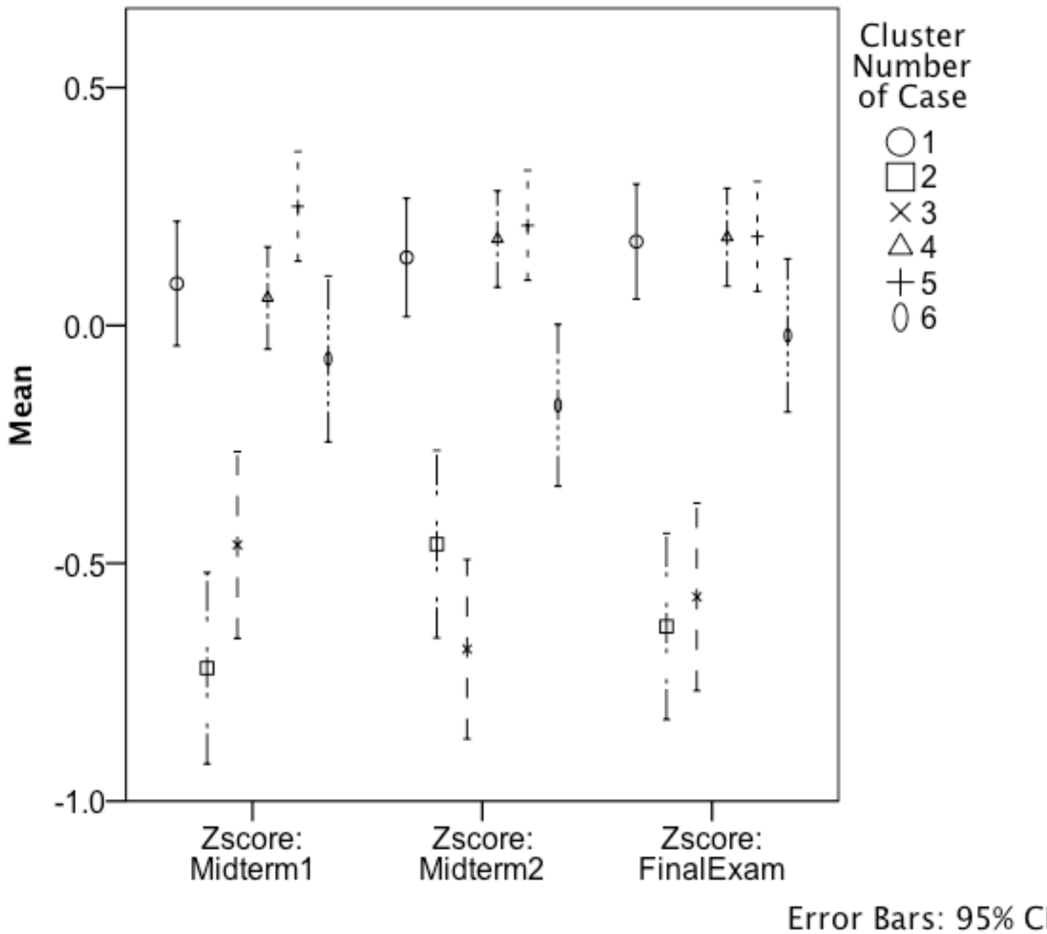

Figure 29. Standardized exam scores grouped by cluster.

How do these profiles of students perform on their online homework? Once again, students in clusters 2 and 3 perform much worse than those in 1, 4, 5, and 6 (Table 43). Since the average item score and overall percent are not equivalent, this group of students may have also missed some assignments. The score students in clusters 2 and 3 see as their online homework average grade (~75%) are much higher than their examination scores, due to the small penalty for submitting incorrect answers.

Table 43. Sapling Scores by Online Homework Behaviour Cluster

	Cluster	*N*	Mean	Standard Deviation	Minimum	Maximum
Sapling Percent	1	216	91.22	9.34	50	99
	2	60	75.14	19.23	8	95
	3	91	76.05	13.34	32	94
	4	344	89.49	11.99	7	100
	5	307	89.30	13.12	18	100
	6	110	89.21	12.27	32	99
	Total	1128	87.89	13.37	7	100
Average Item Score	1	216	0.96	0.03	1	1
	2	60	0.88	0.04	1	1
	3	91	0.87	0.05	1	1
	4	344	0.95	0.03	1	1
	5	307	0.96	0.03	1	1
	6	110	0.95	0.04	1	1
	Total	1128	0.94	0.04	1	1
Transformed Average Item Score	1	216	–1.49	0.32	–2.3	–0.8
	2	60	–0.95	0.14	–1.3	–0.64
	3	91	–0.91	0.14	–1.24	–0.45
	4	344	–1.44	0.32	–2.57	–0.6
	5	307	–1.53	0.36	–2.55	–0.62
	6	110	–1.40	0.36	–2.22	–0.8
	Total	1128	–1.40	0.37	–2.57	–0.45
Transformed Sapling Percent	1	216	0.83	0.37	0.20	1.7
	2	60	1.32	0.28	0.78	1.97
	3	91	1.34	0.22	0.87	1.84
	4	344	0.89	0.38	0.13	1.98
	5	307	0.85	0.43	0.11	1.92
	6	110	0.88	0.40	0.20	1.84
	Total	1128	0.92	0.41	0.11	1.98

How confident and certain are these learners, and to what extent do they learn from online homework errors? The six student profiles differed on the weighted exam, confidence, certainty variables, and learning from errors (correct on 2nd attempt) (Table

44). There were no differences in their goal orientation and error orientation subscale scores. Details of the analysis of variance are included in Appendix H.

Table 44. Confidence, Certainty, and Learning from Errors by Cluster

	Cluster	*N*	Mean	Standard Deviation	Minimum	Maximum
Confidence	1	202	72.06	20.17	0	100
	2	51	61.92	21.97	6	100
	3	80	63.61	22.46	0	100
	4	321	73.18	20.32	0	100
	5	276	73.13	22.63	0	100
	6	104	66.51	21.42	3	100
	Total	1034	70.98	21.55	0	100
Certainty	1	202	2.24	0.50	1	3
	2	51	2.11	0.45	1	3
	3	80	2.18	0.46	1	3
	4	320	2.37	0.45	1	3
	5	276	2.39	0.46	1	3
	6	103	2.18	0.51	1	3
	Total	1032	2.30	0.48	1	3
Learning From Errors (Correct 2nd Attempt)	1	216	57.20	11.57	27	91
	2	60	41.27	6.94	27	56
	3	91	40.11	6.90	19	53
	4	344	55.69	11.85	25	100
	5	307	58.77	12.70	19	100
	6	110	54.28	10.79	31	88
	Total	1128	54.66	12.75	19	100

What profiles of students exist with respect to achievement, confidence, and learning from errors? I followed the same clustering procedure using measures of achievement, confidence, and learning from errors. In this analysis I selected the cases that had confidence and certainty judgments on more than 37 of the questions. The goal of this was to exclude the students who were not regularly making confidence and certainty predictions. This left 295 cases that took ten iterations to cluster into eight groups (Table **45**).

Table 45. Achievement, Confidence, and Learning From Errors Cluster Centres

Variable	1 *N* = 34	2 *N* = 46	3 *N* = 10	4 *N* = 76	5 *N* = 4	6 *N* = 22	7 *N* = 59	8 *N* = 44
Weighted Exam	75	83	56	77	56	46	53	58
Confidence	41	85	15	74	7	47	60	83
Learning from errors (Correct 2nd Attempt)	61	77	44	59	74	64	44	53

Clusters 2, 4, 6, and 7 have similar weighted exam scores and confidence judgments. These well-calibrated students have various levels of learning from errors. Students in clusters 1, 3, and 5 are underconfident, and students in cluster 8 are overconfident. Interestingly, no cluster of students exists that does not learn from errors and that has high scores on the examinations.

Chapter 5.

Discussion

Goal Orientation and Achievement

How do organic chemistry students' levels of task, self, and other-approach and avoidance goals relate to achievement, online homework behaviours, and confidence? To answer this question, I performed confirmatory factor analysis on the responses to Elliot et al. (2011) achievement goal orientation questionnaire. I then determined the level each student endorsed task-approach, task-avoidance, self-approach, self-avoidance, other-approach, and other-avoidance goals at two points in the course. Finally, I compared their achievement goal endorsement with course exam grades and online homework grades through correlational analysis.

Model validity. The results support the 3 × 2 achievement goal model proposed by Elliot (2011). That is, factor analysis confirmed a model with satisfactory fit that had distinct subscales for task-approach, task-avoidance, self-approach, self-avoidance, other-approach, and other-avoidance goals. In the literature, one study used the 3 × 2 achievement goal to compare the achievement goals of traditional and non-traditional college students, but it did not include model fit statistics (Johnson & Kestler, 2013). Two others have found support for the 3 × 2 structure (C.-C. Wu, 2012; Yang & Cao, 2013). An adapted measure of the 3 × 2 AGQ, named the "Gaming Goal Orientation" scale also reports satisfactory fit statistics for the 3 × 2 structure (Quick, 2013). Currently, there are no published studies that report unsatisfactory fit statistics for the 3 × 2 model. Despite the satisfactory model fit, the approach and avoidance scores were highly correlated in this sample of organic chemistry learners (.71−.73). However, the body of goal orientation research beyond the 3 × 2 model is strongly in support of approach and avoidance being separate factor structures (Elliot, McGregor, & Gable, 1999; Marzouq, Carr, & Slade, 2012; Pintrich, 2000; Wigfield & Cambria, 2010). Further research should examine data from many learner populations to confirm that approach and avoidance goals are indeed separate constructs in the 3 × 2 perspective.

Endorsement. In this study, organic chemistry learners strongly endorsed all achievement goals, with means between 5.1 and 5.9 on a seven-point scale. Johnson and Kestler (2013) reported the mean scores of traditional and non-traditional college students enrolled in education programs that were very slightly higher than for this sample of organic chemistry learners for all subscales except for other-approach of traditional students only. However, my sample has higher mean scores than the Taiwanese elementary and junior school students in Wu's (2012) study, which ranged from 4.4 to 5.3. There are only a few groups with which to compare my sample's mean scores, so not much can be said about why the scores vary in this way. In achievement goal research using other measures, the consensus is that children become less mastery oriented and more performance oriented over time (Wigfield & Cambria, 2010). However, Wigfield and Cambria's (2010) study showed that children's performance orientation can decrease over time. There is not much data available to explain differences in means between different types of people.

Reliability and Stability. My results provide support for the stability of achievement goals over the timescale of a few weeks, as the correlations between initial and retest subscale scores ranged from .5−.7. Additionally, the responses do not seem to be affected by major course events, such as course exams. Wu (2012) calculated the internal-consistency reliability of the subscale scores as ranging from .75−.95 in junior high school and elementary school samples, which are similar to my range of .75−.90. Research has shown that children's goal orientation scores are stable over a one-year period (Wigfield & Cambria, 2010). Further research could examine shifting of students' achievement goals over time, possibly by frequent random samples of a large population to avoid test fatigue. Using the time between tests as an independent variable, one could show the rate of change over days, weeks, months, and years. The results of such research could help reduce uncertainty about the trait-like nature or situational variability of achievement goals. This is important for researchers who use achievement goal measures in their studies, since doing so requires one to make assumptions about stability when drawing conclusions.

Relation to Achievement. The relationship between task-approach and online homework performance is slightly stronger than it is for task-avoidance and online homework performance, which provides weak evidence of separation between the goal orientations. This parallels the case for other-approach and other-avoidance. Task-approach and other-approach goals were much more strongly related to course grade and exam scores than were task-avoidance and other-avoidance. This result provides support for distinguishing approach and avoidance goals of the task and other referents. The relationship between specific goals and achievement depends on the measure of achievement. Students in this class showed similarly strong endorsements of self-approach and self-avoidance goals, but the individual's scores did not relate to achievement gauged by Sapling learning scores or exam scores.

Relation to Online Homework Behaviours. Yang (2013) observed a relationship between help seeking in e-learning and achievement goals using the 3 × 2 achievement goal framework with a sample of college educational psychology students. Self-approach goals were related to help seeking but only through intrinsic motivation. Yang also showed that self-approach goals were negatively related to help seeking. My study not detect relationships between hint viewing (a form of help seeking) and goal orientation perhaps because help-seeking in the form of viewing hints was mined from the software database, as opposed to a self-report measure of help seeking. These opposing research findings could indicate that self-report and observed measures of help seeking reflect different aspects of metacognition and goals.

The strongest relationship I detected between achievement goal variables and giving up was a small relation between giving up and other-approach orientation (r = .105). This weak but statistically detectable negative relationship suggests students who want to outperform others are somewhat less likely to give up and decide to view the correct solution to a problem before solving it correctly. Since not giving up is a measure of persistence, those who are motivated by performing better than others seem to be more persistent. Future research could examine the relationship between individual attempts and giving up to see when students with different levels of achievement goal orientations abandon tasks. That is, the timing of giving up was not explored in this study but doing so would provide more detailed information about persistence in online homework learning.

Other-approach scores were also detected to relate negatively to the average number of attempts students took on online homework problems (r = –.130). This suggests the aspect of goal orientation most related to online homework behaviours is other-approach, although the effects are small.

Relation to Confidence. Only task-approach and other-approach goals were related to students' confidence judgments. Zhou (2013) used confidence judgments to construct scores of metacomprehension accuracy and showed that they were not related to goal orientation (2 × 2 achievement goal model). Zhou discusses the complex effects that multiple goals have on self-regulation, as they could "add, conteract, or interact" (p. 10).

Relation to Learning from Errors. Since task orientation is about correctly completing a task correctly or understanding a specific idea (Elliot et al., 2011), it is plausible that task-approach and avoidance scores relate to error orientation. When learners work through tasks or attempt to develop conceptual understanding, errors are common. The most important relationships between goal orientation and error orientation involve task-approach goals and learning about errors, and task-approach goals and thinking about errors. These positive, moderate relationships could indicate that they are measuring something similar, such as the drive to respond to questions correctly; or that one influences the other. Further research could examine the nomological net for these constructs to attempt to determine if they are different constructs. Rybowiak et al. (1999) examined the relationship among six error orientation constructs and other measures, the one most similar to goal orientation was need for achievement. Need for achievement is the drive toward reaching goals at work (Modick, 1978 as cited in Rybowiak et al., 1999). Error competence, learning from errors, and error risk-taking were moderately positively associated with need for achievement, while error strain, error anticipation, and covering up errors were slightly negatively related.

These results, acknowledging that the constructs are not exactly the same, appear to be consistent.

Van Dyck (2010b) related achievement goals to error mastery or aversion. Van Dyck constructed two new scales from Rybowiak et al.'s (1999) items, and called the scales mastery approach and mastery avoidance. The mean scores of the 60 students who responded to the items were 2.99 for error mastery (*SD* = 0.82) and 2.42 for error aversion (*SD* = 0.94), lower than in my sample of organic chemistry learners. Van Dyck (2010b) found error mastery was not related to any aspect of goal orientation; and, error aversion was related to avoid performance orientation (r = .24), but not mastery goal orientation or approach performance orientation. In contrast, my results showed several weak relationships among these constructs.

Error Orientation and Achievement

What levels of error competence, learning from errors, error risk taking, error strain, error anticipation, covering up errors, error communication, error motivation, and thinking about errors do organic chemistry students possess, and how do these relate to learning? To answer this question, I first performed confirmatory factor analysis on students' responses to an adapted version of Rybowiak et al.'s (1999) Error Orientation Questionnaire at two points in the course. I then compared students' subscale scores to achievement, online homework behaviours, learning from errors, and confidence through correlational analysis.

Model Validity. The error orientation data fit the factor structure of the published workplace error orientation questionnaire despite the items being modified from the original. Not all subscales had good internal-consistency reliabilities; error risk taking, error communication, and motivation had low Cronbach alpha values in the .5 range. Rybowiak et al.'s (1999) reliability scores ranged from .56 to .89 and Mateo et al.'s (2013) ranged from .63 to .89. Risk item 4, which needed to be reverse-coded, showed weaker correlations with the other error risk-taking items and could be reworded or omitted in future versions of this instrument.

I included items judged to measure the extent to which students are or are not motivated by errors. The items are quite varied (see Appendix G) and correlate highly with the responses to other items. For example, motivation item 2, “When I make a lot of mistakes, I feel discouraged” with competence item 2, “When I do something wrong when solving a problem, I correct it immediately“ correlates r = .437. It also has a moderate negative correlation with strain 1, “I find it stressful when I make errors during problem solving” (r = .555) and error strain 2, “I am often afraid of making mistakes” (r = –.482). The error motivation construct is new, and should be further explored by comparing it to other motivation constructs. Good candidates for comparison would be motivation as measured but the Motivated Strategies toward Learning Questionnaire (Pintrich et al., 1991) and the need for cognition scale (Cacioppo, Petty, & Feinstein, 1996).

Endorsement. Organic chemistry learners in this study endorsed the various error orientation subscales to different degrees. Learning from errors had the greatest mean (4.07) and covering up errors had the smallest mean (2.19). Covering up errors was also the only subscale to have a mean below 2.5, which would be the centre of a 5-point scale. This indicates students in my sample generally do not believe that they cover up errors. Compared to the Dutch students in Rybowiak et al.’s (1999) study 2, organic chemistry learners in my sample had higher scores on learning from errors, error strain, error anticipation, error communication, and thinking about errors. The mean scores were the same for both groups for covering up errors. These results indicate that error orientation in learning settings may be fundamentally different than learning from errors in a workplace setting.

Error competence relates to one’s knowledge about how to recover from errors and reduce negative consequences errors may have. In a learning setting, a negative consequence is that the learner is not able to currently complete a task. During learning, negative outcomes are minor, especially in low-stakes practice settings. Negative outcomes of errors are mainly for tests, which are summative assessments. In workplace settings, there may be more immediate and important negative consequences of errors, and thus people may develop greater competence in dealing with errors as they gain experience in the workplace. Current instructional practices do not typically focus on the consequences of errors, so it is not surprising students scored lower on this subscale.

Learning from errors is one main goal of practice in educational settings, and students in organic chemistry endorse these statements strongly because they generally have high goals for their performance on course examinations. In the workplace, learning from errors may have a more abstract meeting, perhaps to improve one's workplace performance over time in a general sense. Error risk taking scores were only slightly different between groups, but one could imagine types of risks differ in the workplace and learning settings. In low-stakes practice, making errors relates to taking responsibility and having flexibility. Error strain may be higher in the sample of organic chemistry learners because they are more highly driven to succeed in organic chemistry than the Dutch students are to succeed in their part-time jobs.

Organic chemistry learners anticipated errors in their academic work more than the Dutch students do in their work. As with the other constructs, whether this is due to differences between samples or the perspective being measured (learning or workplace) is not possible to determine with the data available.

Care should be taken not to generalize these comparisons beyond this sample. Since classroom goal climate can influence student goal orientation, different results may be obtained with other academic settings. Similarly, workplace error culture is highly variable, and the Dutch sample was from students with a wide variety of part-time jobs, all with potentially different error climates.

Stability over Time. This study does not support the idea that major course events, such as course exams, impact error orientation. The correlation between initial and retest scores for groups separated by when they responded to the questionnaire do not show a clear trend or weakened correlations. Like with goal orientation, it appears error orientation may be stable over a time period of a few weeks. Future research could examine this more closely to discover the limits of this inference.

Relation to Achievement. Some components of error orientation relate to achievement in this group of students. This is the first time achievement has been compared to students' responses on an error orientation questionnaire in a academic setting. Error communication scores did not relate to achievement. These items asked students about the usefulness of discussing their errors with classmates and teachers. Science educators are using pedagogies such as peer instruction to emphasize collaborative, cultural aspects of knowledge construction. Higher-performing students in this study did not report greater perceived usefulness of discussing their errors with other students or teachers. This could be because students sometimes do not perceive the value of peer discussion (Crouch & Mazur, 2001), or it could be because discussing with peers and teachers has a similar level of helpfulness to all students, regardless of their level of knowledge. The finding also could be a combination of these two factors.

Since error competence, learning from errors, error risk-taking, thinking about errors, and error motivation showed positive correlations with achievement, and error strain, anticipation, and covering up errors had negative correlations with achievement, researchers could experiment with treatments that manipulate the extent to which classrooms support making errors, as in the error training literature Frese et al., 1999; Heimbeck, 2003). Additional experiments could explore direct instruction about error orientation and error handling skills, similar to the error framing literature. For example, an instructional designer could develop activities that teach learners errors are beneficial to learning. These types of studies would help establish causality between error orientation and achievement.

Relation to Online Homework Behaviours. The several weak relationships between error orientation and online homework behaviours indicate that error orientation relates to behaviours in an error-prone environment. Hint viewing does not relate to error orientation, but the tendency to give up does. For example, a weak (r = .11) relationship was observed between giving up and error anticipation. Rybowiak et al. (1999) describes the pessimistic nature of those with high error anticipation, and perhaps students are sufficiently pessimistic about the question they tend to give up more than those with lower levels of error anticipation. The relationship between error communication and giving up also was weak (r = –.14). Those who are more likely to give up have a slight tendency to report they see less value in discussing their errors with others.

The average number of attempts a student makes on a question has several causes, including knowledge and persistence. Error strain has a weak positive relationship with the average number of attempts and thinking about errors and error motivation have weak negative relationships. In the case of strain, the causality could go in either direction or be reciprocal over time; the more errors a student makes, the more strained they become. Or, more highly stressed individuals make more errors because of the stress. It would be useful to examine students' progression in both behaviours and error orientation over the course of a term to gain some evidence about the mechanism behind these relationships.

Relation to Confidence. The only type of error orientation that did not show a relationship to confidence is error communication, although covering up errors was very small (r = –.08). The other error orientations showed weak to moderate relationships to students' average confidence, with directionality as expected. Error competence, learning, risk taking, thinking about errors, and motivation all had weak positive relationships to average confidence. Error strain and anticipation had weak negative relationships to average confidence. Rybowiak et al. (1999) did not examine confidence or judgment of learning accuracy, but he did examine the relationship between error orientation and self-efficacy for six subscales. In his study, error competence, learning from errors, and risk taking showed moderate positive relationships with self-efficacy and error strain, anticipation; covering up errors showed weak to moderate negative relationships. Since the relationships have a similar directionality, this lends support for the relationship between confidence and self-efficacy and error orientation. Since certainty was very similar to confidence, the relationships between certainty and error orientation are similar.

Relation to Learning from Errors. The relationship between the constructed measure of learning from errors and the self-reported scores about learning from errors is very weak (r = .07). This measure of learning from errors does not distinguish between different types of errors. For example, students have made a small "slip" or guessed incorrectly. Or, students could submit an incorrect answer because of a misconception or deficiency in knowledge. There is likely a lot of noise associated with this measure. Further research should mine online homework data and use triangulation to determine how to better measure learning from errors in the online homework environment. The exploratory work in this study informs this next stage of research.

Other subscales on the EOQ relate more strongly to learning from errors, that is error competence, communication, thinking, and motivation showed weak positive relationships and strain, anticipation, and covering up showed weak negative relationships. Despite the error associated with learning from errors, it does have detectable relationships in the sensible direction with error orientation subscales.

Online Homework Question Characteristics

How do students approach their online homework, and what is the impact of this on learning? To answer this question, I explored the extent to which students used hints, how frequently they gave up, and the average number of attempts taken on questions that varied by average score, average confidence, and average certainty. This purpose of this broad question is to gain a better understanding of the function of the online homework questions. I then examined the relationship between behaviours, confidence, certainty, goal orientation, and error orientation. Finally, I performed two cluster analyses to determine the profiles of learners with respect to online homework behaviours and achievement, confidence, and learning from errors.

The number of students viewing hints was different on each homework question, and hints sought by each student were not related to their confidence, tendency to give up, achievement on exams or online homework. Hint use was negatively minimally related to average certainty. This suggests that less certain students may be using hints to confirm their responses rather than to inform them. More research should be done to explore the sequentially unfolding interactions with the online homework program to clarify this issue. Generally, more hints are viewed on lower-scoring (more difficult) questions and more students give up on lower scoring questions.

Online Homework Behaviours and Achievement

The intention of offering online homework is to improve student learning and achievement on course examinations. While many studies have shown a positive correlation between online homework scores and exam grades, none have examined specific interactions with the system in relation to achievement. I showed that higher achieving organic chemistry learners not only receive higher scores on their online homework, they also use fewer hints and give up less often. They require far fewer attempts to correctly answer a question. This could suggest that students who require multiple attempts at a problem do not learn enough to get to the same level as the higher achieving students, a contradiction to the notion that errors might benefit learning. On the other hand, lower scoring students may have scored even less well if they were not offered opportunities to identify and correct errors. Further research could examine pre- and post-homework session achievement to narrow the effect of multiple attempts on learning.

Future research could explore the possibility of providing students with feedback that their behaviours are maladaptive or providing greater incentives to take their work with the program more seriously. This study provides evidence of something practitioners have likely suspected, that is, not all students engage with practice problems in a deep way that leads to increased knowledge and problem solving ability.

Confidence

The relationships between confidence, certainty, goal orientation and error orientation were discussed earlier in this chapter. Students who made higher predictions of future success were more certain in their predictions. This adds another meta-dimension to metacognitive accuracy, as I asked students to rate the accuracy of their metacognitive judgments. Since the probability of a successful homework response rose with confidence and with certainty, it does not seem as though higher achieving students in this sample suffer from cognitive bias.

There was no link between average confidence scores and learning from errors, although there was a weak relationship detected between certainty and learning from errors (r = .12). This finding suggests that when students predict a greater certainty of their responses, they have a higher likelihood of answering correctly on the second attempt if their first attempt is incorrect. To explain this effect, one could explore the nature of errors students made. Perhaps students with higher learning from errors scores make less conceptual, more accidental errors.

Learner Profiles

What profiles of students exist with respect to achievement, confidence, and learning from errors? Six profiles of learners resulted from the cluster analysis that included giving up, hints viewed, and average attempts. The profiles with the highest frequency of giving up scored the lowest scores on examinations. They also had lower confidence and were less likely to learn from errors. These relationships could be due to general characteristics of these clusters, but they were not related to learners' goal and error orientations. Perhaps a general academic motivational variable accounts for both the less engaged interactions with the online homework system and the lower course performance.

Eight profiles of learners with respect to exam scores, confidence, and learning from errors indicate that there is no group of learners with low learning from errors and high exam scores. This means that learners that do well on examinations typically respond to an online homework question correctly after a first unsuccessful attempt.

These findings could be used to better customize online learning environments for learners with different profiles. Examples of such customization would be to provide process feedback to learners who give up on many questions. Students may not be aware that this behaviour has a negative association with exam performance. Also, it is possible that some students learn best when the wrong-answer penalty is higher, as it provides motivation to more carefully consider each question attempt. However, some students' error strain may increase with the size of the grade penalty. These customizations are possible with today's technology, but they need to be informed by research. Such customizations could improve learning, which, in turn, could enhance students' experience in organic chemistry and improve attitude toward science in general.

Limitations

As with any study that uses self-report measures, the validity of interpretations rely on the level of student effort in responding to the questionnaire items. I have not optimized the measure of learning from errors (correct second attempt). It is possible that the correct second attempt measure is not the best way to estimate learning from errors in online homework. Another measure, perhaps using think-aloud methodology, could be used to validate and triangulate these data. Also, students are still learning from errors even on their third and higher attempts, and the learning from errors variable I constructed ignores this. Additionally, students may respond incorrectly on their first attempt of question even if they had the knowledge required to solve the problem. For example, they could have selected the wrong atom label by mistake.

In this study I investigated the learning process of students who completed the course. Students who withdraw from the course are likely different from those who complete it, for example they may respond differently to online homework and have different achievement goals and error orientations. Future work should explore how these students approach their online homework.

There are several other settings in which organic chemistry learners must learn from errors. They use ungraded problem sets, textbook questions, and practice exams in their studying. I do not know the portion of learning that occurs in each setting. However, students spend a considerable amount of time responding to the hundreds of online homework questions and it is reasonable to think that a portion of their learning comes from this experience. Also, the types of errors students make in their online homework attempts were not analyzed in this study. This data is available but manually judging the nature of student errors was beyond resources available at this time.

Some are concerned that performance approach goals in general are over-emphasized in research, and that other important goals students are neglected by researchers (Costa & Remedios, 2014). In this study, the emphasis of the AGQ referent of “others” refers to classmates, but students have other people in their lives that can influence their goals, such as their family members and teachers. Thus, the AGQ only measures some aspect of the goal orientation of learners.

Additionally, the only measure of achievement used is course examination scores. Expert chemists and instructors created these examinations, but I do not know the extent to which the exam measures a mixture of organic chemistry knowledge and skill with factors such as speed of processing, reading comprehension, or other unrelated skills. The examination and the online homework questions have a large degree of overlap, but the topic emphasis and level of difficulty of the items is likely not the same.

I have taken care not to assume causality with the correlations among variables analyzed in this study. Controlled experiments or quasi-experiments could be used to answer many of the remaining questions in this area.

Conclusions

In this dissertation I relate goal orientation, error orientation, confidence, online homework behaviours, and learning in an organic chemistry context. Organic chemistry learners strongly endorsed all achievement goals, and these goals were stable over time. This study provides evidence in support of the validity of Elliot et al.'s (2011) 3 × 2 achievement goal questionnaire. When working in an online homework environment, learners selectively viewed hints, but hint-viewing behaviour was not found to relate to any other variables. Statistically significant, but not educationally significant relationships were found between giving up and other-approach orientation scores.

The adaptation of the Error Orientation Questionnaire showed a similar stable factor structure to the published instrument, and the correlations with achievement support the theory that higher achieving students have more positive views of the purpose of errors in learning. Students most strongly endorsed the items in the latent variable called learning from errors. Learning from errors during online homework practice did not relate to students' self-reported scores on the learning from errors subscale of the Error Orientation Questionnaire. Distinct profiles of learners based on online homework behaviours were shown to have different achievement levels on course examinations. Future research could take an experimental approach to error framing in an online homework context and explore the impact on online homework behaviours and learning.

References

Allain, R., & Williams, T. (2006). The effectiveness of online homework in an introductory science class. *Journal of College Science Teaching*, *35*(6), 28–30.

Allison, P. D. (2003). Missing data techniques for structural equation modeling. *Journal of Abnormal Psychology*, *112*(4), 545–557. doi:10.1037/0021-843X.112.4.545

American Chemical Society, "College to Careers" *Organic Chemistry*. Retrieved August 7th, 2014 from http://www.acs.org/content/acs/en/careers/college-to-career/areas-of-chemistry/organic-chemistry.html.

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, *84*(3), 261–271.

Ames, C., & Archer, J. (1987). Mothers' beliefs about the role of ability and effort in school learning. *Journal of Educational Psychology*, *79*(4), 409–414. doi:10.1037/0022-0663.79.4.409

Ammons, R. B. (1956). Effects of knowledge of performance: A survey and tentative theoretical formulation. *The Journal of General Psychology*, *55*, 279–299.

Anderman, E. M., & Maehr, M. L. (1994). Motivation and schooling in the middle grades. *Review of Educational Research*, *64*(2), 287–309. doi:10.3102/00346543064002287

Anderson, J. R., Corbett, A. T., & Koedinger, K. R. (1995). Cognitive tutors: Lessons learned. *The Journal of the Learning Sciences*, *4*(2), 167–207. doi:10.1207/s15327809jls0402_2

Arenas, A., Tabernero, C., & Briones, E. (2006). Effects of goal orientation, error orientation and self-efficacy on performance in an uncertain situation. *Social Behavior and Personality: an International Journal*, *34*(5), 569–586.

Arzi, H. J., Ben-Zvi, R., & Ganiel, U. (1986). Forgetting versus savings: The many facets of long-term retention. *Science Education*, *70*(2), 171–188. doi:10.1002/sce.3730700211

Bandura, A. (1997). *Self-efficacy: The exercise of control* (Vol. d). New York, NY: W.H. Freeman.

Bangert-Drowns, R. L., Kulik, C. L. C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, *61*(2), 213–238. doi:10.3102/00346543061002213

Barr, D. A., Matsui, J., Wanat, S. F., & Gonzalez, M. E. (2010). Chemistry courses as the turning point for premedical students. *Advances in Health Science Education*, *15*, 45–54. doi:10.1007/s10459-009-9165-3

Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*.

Bodner, G. M. (1991). I have found you an argument: The conceptual knowledge of beginning chemistry graduate students. *Journal of Chemical Education*, *68*(5), 385–388. doi:10.1021/ed068p385

Bonham, S. W., Deardorff, D. L., & Beichner, R. J. (2003). Comparison of student performance using web and paper-based homework in college-level physics. *Journal of Research in Science Teaching*, *40*(10), 1050–1071. doi:10.1002/tea.10120

Bonham, S., Beichner, R., & Deardorff, D. (2001). Online homework: Does it make a difference? *The Physics Teacher*, *39*(5), 293–296. doi:10.1119/1.1375468

Bradley, A. Z., Ulrich, S. M., & Jones, M., Jr. (2002). Teaching the sophomore organic course without a lecture. Are you crazy? *Journal of Chemical Education*, *79*(4), 514–519. doi:10.1021/ed079p514

Brodbeck, F. C., Zapf, D., Prümper, J., & Frese, M. (1993). Error handling in office work with computers: A field study. *Journal of Occupational and Organizational Psychology*, *66*, 303–317.

Burch, K. J., & Kuo, Y.-J. (2010). Traditional vs. online homework in college algebra. *Mathematics and Computer Education*, *44*(1), 53–63.

Burns, R., & Burns, R. (2009). Chapter 23-Cluster Analysis. In *Business Research Methods and Statistics using SPSS* (p. 560). Business Research Methods and Statistics using SPSS. Retrieved from http://www.uk.sagepub.com/burns/chapters.htm

Butler, D., & Winne, P. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, *65*(3), 245–281.

Butler, M. B., & Zerr, R. J. (2005). The use of online homework systems to enhance out-of-class student engagement. *International Journal for Technology in Mathematics Education*, *12*(2), 51–58.

Button, S. B., Mathieu, J. E., & Zajac, D. M. (1996). Goal orientation in organizational research: A conceptual and empirical foundation. *Organizational Behavior and Human Decision Processes*, *67*(1), 26–48. doi:10.1006/obhd.1996.0063

Buzhardt, J., & Semb, G. B. (2002). Item-by-item versus end-of-test feedback in a computer-based PSI course. *Journal of Behavioral Education*, *11*(2), 89–104. doi:10.1023/A:1015479225777

Byrne, B. M. (2013). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Routledge.

Cacioppo, J. T., Petty, R. E., & Feinstein, J. A. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, *119*(2), 197–253.

Cannon-Bowers, J. A., Rhodenizer, L., Salas, E., & Bowers, C. A. (1998). A framework for understanding pre-practice conditions and their impact on learning. *Personnel Psychology*, *51*(2), 291–320.

Carroll, J. M., & Kay, D. S. (1988). Prompting, feedback and error correction in the design of a scenario machine. *International Journal of Man-Machine Studies*, *28*, 11–27.

Carver, C., & Scheier, M. (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological Review*, *97*(1), 19–35.

Chamala, R. R., Ciochina, R., & Grossman, R. B. (2006). EPOCH: An organic chemistry homework program that offers response-specific feedback to students. *Journal of Chemical Education*, *83*(1), 164–169. doi:10.1021/ed083p164

Cheng, K., Thacker, B., Cardenas, R., & Crouch, C. (2004). Using an online homework system enhances students' learning of physics concepts in an introductory physics course. *American Journal of Physics*, *72*(11), 1447–1453.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*(2), 121–152. doi:10.1207/s15516709cog0502_2

Clark, R. E. (1994). Media will never influence learning. *Educational Technology Research and Development*, *42*(2), 21–29. doi:10.1007/BF02299088

Cohen, E., Mason, A., Singh, C., & Yerushalmi, E. (2008). Identifying differences in diagnostic skills between physics students: Students' self-diagnostic performance given alternative scaffolding. *AIP Conference Proceedings*, *1064*, 99–102.

Cole, R., & Todd, J. (2003). Effects of web-based multimedia homework with immediate rich feedback on student learning in general chemistry. *Journal of Chemical Education*, *80*(11), 1338–1343.

Conroy, D. E., Elliot, A. J., & Hofer, S. M. (2003). A 2 × 2 Achievement Goals Questionnaire for Sport: Evidence for Factorial Invariance, Temporal Stability, and External Validity. *Journal of Sport & Exercise Psychology*, *25*, 456–476.

Cooper, H., Robinson, J. C., & Patall, E. A. (2006). Does homework improve academic achievement? A synthesis of research, 1987-2003. *Review of Educational Research*, *76*(1), 1–62. doi:10.3102/00346543076001001

Copolo, C. E., & Hounshell, P. B. (1995). Using three-dimensional models to teach molecular structures in high school chemistry. *Journal of Science Education and Technology*, *4*(4), 295–305. doi:10.1007/BF02211261

Costa, L. D., & Remedios, R. (2014). Different methods, different results: Examining the implications of methodological divergence and implicit processes for achievement goal research. *Journal of Mixed Methods Research*, *8*(2), 162–179. doi:10.1177/1558689813495977

Crippen, K., Schraw, G., & Brooks, D. (2005). Using an interactive, compensatory model of learning to improve chemistry teaching. *Journal of Chemical Education*, *82*(4), 637–640. doi:10.1021/ed082p637

Crouch, C. H., & Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American Journal of Physics*, *69*(9), 970–977. doi:10.1119/1.1374249

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, *125*(6), 627–668. doi:10.1037//0033-2909.125.6.627

Demirci, N. (2007). University students' perceptions of web-based vs. paper-based homework in a general physics course. *Eurasia Journal of Mathematics*, *3*(1), 29–34.

Demirci, N. (2010). The effect of web-based homework on university students' physics achievements. *The Turkish Online Journal of Educational Technology*, *9*(4), 156–161.

Different Types of Chemistry. (2009). Different Types of Chemistry. *University of Wisconsin La Crosse Department of Chemistry and Biochemistry*. Retrieved May 20, 2014, from http://www.uwlax.edu/chemistry/html/types.htm

Dillard-Eggers, J., & Wooten, T. (2011). Evidence on the effectiveness of on-line homework. *College Teaching Methods Styles Journal*, *4*(5), 9–16.

Duda, J. L. (1989). Relationship between task and ego orientation and the perceived purpose of sport among high school athletes. *Journal of Sport & Exercise Psychology*, *11*, 318–335.

Duis, J. M. (2011). Organic chemistry educators' perspectives on fundamental concepts and misconceptions: An exploratory study. *Journal of Chemical Education*, *88*(3), 346–350. doi:10.1021/ed1007266

Dunlosky, J., & Lipko, A. (2007). Metacomprehension. *Current Directions in Psychological Science*, *16*(4), 228–232. doi:10.1111/j.1467-8721.2007.00509.x

Dunn, K. E. (2014). Insight into error hiding: Exploration of nursing students' achievement goal orientations. *The Journal of Nursing Education*, *53*(2), 93–96. doi:10.3928/01484834-20140122-02

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*. Retrieved June 2, 2013, from http://www.nisdtx.org/cms/lib/TX21000351/Centricity/Domain/21/j%20carlisle/Motivational%20Processes.pdf

Earley, P. C., Connolly, T., & Ekegren, G. (1989). Goals, strategy development, and task performance: Some limits on the efficacy of goal setting. *Journal of Applied Psychology*, *74*(1), 24–33.

Eichler, J. F., & Peeples, J. (2013). Online homework put to the test: A report on the impact of two online learning systems on student performance in general chemistry. *Journal of Chemical Education*, *90*(9), 1137–1143. doi:10.1021/ed3006264

El-Labban, W. (2003). *Assessment of the effect of online homework on achievement in chemistry*. Retrieved from ProQuest Dissertation & Theses Database.

Elliot, A. J. (1994). *Approach and avoidance achievement goals*. Madison. Retrieved from ProQuest Dissertation & Theses Database.

Elliot, A. J., & Church, M. A. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology*, *72*(1), 218–232.

Elliot, A. J., & Harackiewicz, J. (1996). Approach and avoidance achievement goals and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology*, *70*(3), 461.

Elliot, A. J., & McGregor, H. A. (2001). A 2 × 2 achievement goal framework. *Journal of Personality and Social Psychology*, *80*(3), 501–519. doi:10.1037/0022-3514.80.3.501

Elliot, A. J., & Sheldon, K. M. (1997). Avoidance achievement motivation: a personal goals analysis. *Journal of Personality and Social Psychology*, *73*(1), 171–185.

Elliot, A. J., & Thrash, T. M. (2001). Achievement goals and the hierarchical model of achievement motivation. *Educational Psychology Review*, *13*(2), 139–156.

Elliot, A. J., McGregor, H. A., & Gable, S. (1999). Achievement goals, study strategies, and exam performance: A mediational analysis. *Journal of Educational Psychology*, *91*(3), 549–563.

Elliot, A. J., Murayama, K., & Pekrun, R. (2011). A 3 × 2 achievement goal model. *Journal of Educational Psychology*, *103*(3), 632–648. doi:10.1037/a0023952

Ericsson, K. (2006). The influence of experience and deliberate practice on the development of superior expert performance. *The Cambridge Handbook of Expertise and Expert Performance*, 683–703.

Frese, M., & Altmann, A. (1989). The Treatment of Errors in Learning and Training. In *Developing skills with information technology* (pp. 65–86). Chichester: John Wiley & Sons.

Frese, M., Brodbeck, F., Heinbokel, T., Mooser, C., Schleiffenbaum, E., & Thiemann, P. (1991). Errors in training computer skills: On the positive function of errors. *Human-Computer Interaction*, *6*(1), 77–93.

Fynewever, H. (2008). A comparison of the effectiveness of web-based and paper-based homework for general chemistry. *Chemical Educator*, *13*, 264–269. doi:10.1333/s00897082142a

Gabel, D. L. (1999). Improving teaching and learning through chemistry education research: A look to the future. *Journal of Chemical Education*, *76*(4), 548–544. doi:10.1021/ed076p548

Gabel, D. L., Samuel, K. V., & Hunn, D. (1987). Understanding the particulate nature of matter. *Journal of Chemical Education*, *64*(8), 695–697. doi:10.1021/ed064p695

Garcia, T. (1993). Women and minorities in science: Motivational and cognitive correlates of achievement (pp. 1–15). Presented at the Annual Meeting of the American Education Research Association, Atlanta.

Gardner, D., & Wood, R. (2009). Errors, feedback, learning and performance. *Australian and New Zealand Journal of Organisational Psychology*, *2*(1), 30–43.

Gerbing, D. W., & Hamilton, J. G. (1996). Viability of Exploratory Factor Analysis as a Precursor to Confirmatory Factor Analysis. *Structural Equation Modeling-a Multidisciplinary Journal*, *3*(1), 62–72. doi:10.1080/10705519609540030

Hacker, D.J., Bol, L., & Bahbahani, K. (2008). Explaining calibration accuracy in classroom contexts: the effects of incentives, reflection, and explanatory style. Metacognition and Learning, *3*, 101-121. doi: 10.1007/s11409-008-9021-5

Hall, R. W., Butler, L. G., & Kestner, N. R. (1999). Combining feedback and assessment via web-based homework. *Campus-Wide Information Systems*, *16*(1), 24–26.

Hall, R. W., Butler, L. G., McGuire, S. Y., McGlynn, S. P., Lyon, G. L., Reese, R. L., & Limbach, P. A. (2001). Automated, web-based, second-chance homework. *Journal of Chemical Education*, *78*(12), 1704–1708.

Hammond, K. R., Summers, D. A., & Deane, D. H. (1973). Negative effects of outcome-feedback in multiple-cue probability learning. *Organizational Behavior and Human Performance*, *9*, 30–34.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. doi:10.3102/003465430298487

Hauk, S., & Segalla, A. (2005). Student perceptions of the web-based homework program WeBWorK in moderate enrollment college algebra classes. *Journal of Computers in Mathematics and Science Teaching*, *24*(3), 229–253.

Heimbeck, D., Frese, M., Sonnentag, S., & Keith, N. (2003). Integrating errors into the training process: The function of error management instructions and the role of goal orientation. *Personnel Psychology*, *56*(2), 333–361.

Henderleiter, J., Smart, R., & Anderson, J. (2001). How do organic chemistry students understand and apply hydrogen bonding? *Journal of Chemical Education*, *78*(8), 1126–1130. doi:10.1021/ed078p1126

Hetzner, S., Gartmeier, M., Heid, H., & Gruber, H. (2011). Error orientation and reflection at work. *Vocations and Learning*, *4*(1), 25–39. doi:10.1007/s12186-010-9047-0

Horowitz, G. (2009). It's not always just about the grade: Exploring the achievement goal orientations of pre-med students. *The Journal of Experimental Education*, *78*(2), 215–245. doi:10.1080/00220970903352746

Huang, C. (2011). Discriminant and criterion-related validity of achievement goals in predicting academic achievement: A meta-analysis. *Journal of Educational Psychology*, 1–27. doi:10.1037/a0026223

Huesgen, B. (2012). *Impact of an online homework program on the achievement of college students enrolled in first semester general chemistry*. Retrieved from ProQuest Dissertation & Theses Database.

Ivancic, K., & Hesketh, B. (2000). Learning from errors in a driving simulation: Effects on driving skill and self-confidence, *43*(12), 1966–1984.

Johnson, M. L., & Kestler, J. L. (2013). Achievement goals of traditional and nontraditional aged college students: Using the 3 × 2 achievement goal framework. *International Journal of Educational Research*, 1–12. doi:10.1016/j.ijer.2013.03.010

Johnstone, A. H. (1991). Why is science difficult to learn? Things are seldom what they seem. *Journal of Computer Assisted Learning*, *7*(2), 75–83. doi:10.1111/j.1365-2729.1991.tb00230.x

Johnstone, A. H. (1993). The development of chemistry teaching: A changing response to changing demand. *Journal of Chemical Education*, *70*(9), 701–705. doi:10.1021/ed070p701

Jungic, V., Kent, D., & Menz, P. (2012). On online assignments in a calculus class. *Journal of University Teaching and Learning Practice*, *9*(1), 1–15.

Kaplan, A., & Maehr, M. L. (2007). The contributions and prospects of goal orientation theory. *Educational Psychology Review*, *19*(2), 141–184. doi:10.1007/s10648-006-9012-5

Khanlarian, C. J. (2010, December 15). *A longitudinal study of web-based homework*. Greensboro. Retrieved from ProQuest Dissertation & Theses Database.

Khanlarian, C. J., & Singh, R. (2013). An exploratory study of the online learning environment. *Issues in Accounting Education*, *29*(1), 117–147. doi:10.2308/iace-50614

Kind, V. (2004). *Beyond appearances: Students' misconceptions about basic chemical ideas* (2nd ed., pp. 1–84). Durham: Royal Society of Chemistry.

Kingsbury, C. A. (1986). Report of the symposium on teaching in organic chemistry. *Journal of Chemical Education*, *63*(4), 315–317. doi:10.1021/ed063p315

Kluge, A., Ritzmann, S., Burkolter, D., & Sauer, J. (2011). The interaction of drill and practice and error training with individual differences. *Cognition, Technology & Work*, *13*(2), 103–120. doi:10.1007/s10111-010-0158-z

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*(2), 254–284.

Krylova, I. (1997). *Investigation of causes of differences in student performance on the topics of stereochemistry and reaction mechanisms in an undergraduate organic chemistry course*. Dissertation Abstracts International.

Kulhavy, R. W., White, M. T., Topp, B. W., & Chan, A. L. (1985). Feedback complexity and corrective efficiency. *Contemporary Educational Psychology*, *10*, 285–291.

Lafarge, D. L., Morge, L. M., & Méheut, M. M. (2014). A new higher education curriculum in organic chemistry: What questions should be asked? *Journal of Chemical Education*, *91*(2), 173–178. doi:10.1021/ed300746e

Lawson, A. (1983). Predicting science achievement: The role of developmental level, disembedding ability, mental capacity, prior knowledge, and beliefs. *Journal of Research in Science Teaching*, *20*(2), 117–129. doi:10.1002/tea.3660200204

Lepper, M. R., & Gurtner, J. L. (1989). Children and computers: Approaching the twenty-first century. *American Psychologist*, *44*(2), 170–178.

Lepper, M. R., & Woolverton, M. (2002). The wisdom of practice: Lessons learned from the study of highly effective tutors. *Improving Academic Achievement: Impact of Psychological Factors on Education*, 135–158.

Lie, S., Taylor, A., & Harmon, M. (1996). *Scoring techniques and criteria. Third International Mathematics and Science Society* (Vol. 1, pp. 1–232). Chestnut Hill, MA.

Litherland, K., Carmichael, P., & Martínez-García, A. (2013). Ontology-based e-assessment for accounting: Outcomes of a pilot study and future prospects. *Journal of Accounting Education*, *31*(2), 162–176. doi:10.1016/j.jaccedu.2013.03.003

Lynch, D. J., & Trujillo, H. (2011). Motivational beliefs and learning strategies in organic chemistry. *International Journal of Science and Mathematics Education*, *9*(6), 1351–1365. doi:10.1007/s10763-010-9264-x

Mahal, K. (2013). *UBC AMS Academic Experience Report 2013* (pp. 1–198).Retrieved from http://www.ams.ubc.ca/wp-content/uploads/2013/07/AMS-Academic-Experience-Survey-2013-FINAL-Electronic.pdf

Marzouq, S., Carr, A., & Slade, L. (2012). A “personal opposites” approach to understanding achievement goal questionnaires. *Psychology of Education Review*, *36*(2), 17–25.

Mason, A., & Singh, C. (2010a). Do advanced physics students learn from their mistakes without explicit intervention? *American Journal of Physics*, *78*(7), 760–767. doi:10.1119/1.3318805

Mason, A., & Singh, C. (2010b). Helping students learn effective problem solving strategies by reflecting with peers. *American Journal of Physics*, *78*(7), 748–754.

Mason, A., Cohen, E., Singh, C., & Yerushalmi, E. (2009). Self-Diagnosis, Scaffolding and Transfer: A Tale of Two Problems. *AIP Conference Proceedings*, *1179*, 27–30.

Mason, A., Cohen, E., Yerushalmi, E., & Singh, C. (2008). Identifying Differences In Diagnostic Skills Between Physics Students: Developing A Rubric. *AIP Conference Proceedings*, *1064*, 147–150.

Mason, D. (Ed.). (2009). *Comparison of homework systems (four web-based) used in first-semester general chemistry*. Denton. Retrieved from ProQuest Dissertation and Thesis Database.

Mateo, C. A., Jr, Muring, M., Malayao, S. O., Jr, Daniel, D. L., & Emperio, M. (2013). Exploring the applicability of Error Orientation Questionnaire (EOQ) in Maritime students: validity and reliability. *Medina College Research Journal*, *1*(1), 38–43.

Morrissey, D. J., Kashy, E., & Tsai, I. (1995). Using computer-assisted personalized assignments for freshman chemistry. *Journal of Chemical Education*, *72*(2), 141–146.

Nicholls, J. G. (1984). Achievement-Motivation – Conceptions of Ability, Subjective Experience, Task Choice, and Performance. *Psychological Review*, *91*(3), 328–346.

Noordzij, G., van Hooft, E. A. J., van Mierlo, H., van Dam, A., & Born, M. P. (2012). The effects of a learning-goal orientation training on self-regulation: A field experiment among unemployed job seekers. *Personnel Psychology*, *66*(3), 723–755. doi:10.1111/peps.12011

Ohlsson, S. (1996). Learning from error and the design of task environments. *International Journal of Educational Research*, *25*(5), 419–448.

Parker, L. L., & Loudon, G. M. (2013). Case study using online homework in undergraduate organic chemistry: Results and student attitudes. *Journal of Chemical Education*, *90*(1), 37–44. doi:10.1021/ed300270t

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*, *37*(2), 91–105. doi:10.1207/S15326985EP3702_4

Peng, J. C. (2009). Using an online homework system to submit accounting homework: Role of cognitive need, computer efficacy, and perception. *Journal of Education for Business*, *84*(5), 263–268. doi:10.3200/JOEB.84.5.263-268

Penn, J. H., Nedeff, V. M., & Gozdzik, G. (2000). Organic chemistry and the Internet: A web-based approach to homework and testing using the WE_LEARN system. *Journal of Chemical Education*, *77*(1), 227–231. doi:DOI:10.1021/ed077p227

Phye, G. D., & Sanders, C. E. (1994). Advice and feedback: Elements of practice for problem solving. *Contemporary Educational Psychology*, *19*, 286–301.

Pintrich, P. R. (2000). An achievement goal theory perspective on issues in motivation terminology, theory, and research. *Contemporary Educational Psychology*, *25*, 92–104. doi:10.1006/ceps.1999.1017

Pintrich, P., Smith, D., Garcia, T., & McKeachie, W. (1991). A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ). *National Center for Research to Improve Postsecondary Teaching and Learning*.

Pribyl, J. R., & Bodner, G. M. (1987). Spatial ability and its role in organic chemistry: A study of four organic courses. *Journal of Research in Science Teaching*, *24*(3), 229–240. doi:10.1002/tea.3660240304

Process Oriented Guided Inquiry. (n.d.). Process Oriented Guided Inquiry. Retrieved January 22, 2014, from https://pogil.org/

Quick, J. N. (2013). Modeling gameplay enjoyment through feature preferences, goal orientations, usage, and gender. Retrieved from ProQuest Dissertation and Thesis Database.

Rhodes, M. T., & Sarbaum, J. K. (2013). Online homework management systems: Should we allow multiple attempts? *Department of Economics Working Paper Series*. Retrieved February 10, 2014, from http://bae.uncg.edu/assets/research/econwp/2013/13-14.pdf

Richards-Babb, M., Drelick, J., Henry, Z., & Robertson-Honecker, J. (2011). Online homework, help or hindrance? What students think and how they perform. *Research and Teaching*, *40*(4), 81–93.

Rixse, J. S., & Pickering, M. (1985). Freshman chemistry as a predictor of future academic success. *Journal of Chemical Education*, *62*(4), 313–315. doi:10.1021/ed062p313

Romero, C., & Ventura, S. (2010). Educational data mining: A review of the state of the art. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, *40*(6), 601–618. doi:10.1109/TSMCC.2010.2053532

Roth, V., Ivanchenko, V., & Record, N. (2008). Evaluating student response to WeBWorK, a web-based homework delivery and grading system. *Computers & Education*, *50*(4), 1462–1482. doi:10.1016/j.compedu.2007.01.005

Rybowiak, V., Garst, H., Frese, M., & Batinic, B. (1999). Error orientation questionnaire (EOQ): reliability, validity, and different language equivalence. *Journal of Organizational Behavior*, *20*(4), 527–547. doi:10.1002/(SICI)1099-1379(199907)20:4<527::AID-JOB886>3.0.CO;2-G

Salmoni, A. W., Schmidt, R. A., & Walter, C. B. (1984). Knowledge of results and motor learning: A review and critical reappraisal. *Psychological Bulletin*, *95*(3), 355–386.

Schell, K. L., & Conte, J. (2008). Associations among polychronicity, goal orientation, and error orientation. *Personality and Individual Differences*, *44*(1), 288–298.

Schell, K. L., Hernandez, J., & Rosebeary, M. (2008). The error orientation quesionnaire: A motivational perspective on validity. *Society for Industrial and Organizational Psychology 3rd Annual Conference*, 1–20.

Seery, M. K. (2009). The role of prior knowledge and student aptitude in undergraduate performance in chemistry: a correlation-prediction study. *Chemistry Education Research and Practice*, *10*(3), 227–232. doi:10.1039/B914502H

Seifert, C., & Hutchins, E. (1992). Error as opportunity: Learning in a cooperative task. *Human-Computer Interaction*, *7*(4), 409–435.

Senko, C., Hulleman, C. S., & Harackiewicz, J. M. (2011). Achievement goal theory at the crossroads: Old controversies, current challenges, and new directions. *Educational Psychologist*, *46*(1), 26–47. doi:10.1080/00461520.2011.538646

Seymour, E., & Hewitt, N. M. (1997). *Talking about leaving: Why undergraduates leave the sciences* (pp. x–429). Boulder: Westview Press.

Simon, H. A. (1967). Motivational and emotional controls of cognition. *Psychological Review*, *74*(1), 29–39.

Skinner, B. F. (1954). The science of learning and the art of teaching. *Harvard Educational Review*. doi:10.1037/11324-010

Small, M. Y., & Morton, M. E. (1983). Spatial visualization training improves performance in organic chemistry. *Research in College Science Teaching*, *13*(1), 41–43.

Sommet, N., Pulfrey, C., & Butera, F. (2013). Did my M.D. really go to university to learn? Detrimental effects of numerus clausus on self-efficacy, mastery goals and learning, *8*(12), e84178. doi:10.1371/journal.pone.0084178

Stankov, L., Lee, J., Luo, W., & Hogan, D. J. (2012). Confidence: A better predictor of academic achievement than self-efficacy, self-concept and anxiety? *Learning and Individual Differences*, *22*(6), 747–758. doi:10.1016/j.lindif.2012.05.013

Steuer, G., Rosentritt-Brunn, G., & Dresel, M. (2013). Dealing with errors in mathematics classrooms: Structure and relevance of perceived error climate. *Contemporary Educational Educational Psychology*, *38*(3), 196–210. doi:10.1016/j.cedpsych.2013.03.002

Szu, E., Nandagopal, K., Shavelson, R. J., Lopez, E. J., Penn, J. H., Scharberg, M., & Hill, G. W. (2011). Understanding academic performance in organic chemistry. *Journal of Chemical Education*, *88*, 1238–1242. doi:10.1021/ed900067m

Tabachnick, B. G., & Fidell, L. S. (2007). *Using Multivariate Statistics* (5 ed.). Pearson Higher Ed.

Taraban, R., Anderson, E. E., Hayes, M. W., & Sharma, M. P. (2005). Developing on-line homework for introductory thermodynamics. *Journal of Engineering Education*, *94*(3), 339–342.

Thoennessen, M., & Harrison, M. J. (1996). Computer-assisted assignments in a large physics class. *Computers & Education*, *27*(2), 141–147.

Tien, L. T., Roth, V., & Kampmeier, J. A. (2002). Implementation of a peer-led team learning instructional approach in an undergraduate organic chemistry course. *Journal of Research in Science Teaching*, *39*(7), 606–632. doi:10.1002/tea.10038

Tjosvold, D., Yu, Z., & Hui, C. (2004). Team learning from mistakes: The contribution of cooperative goals and problem-solving. *Journal of Management Studies*, *41*(7), 1223–1245.

Treasure, D. C., & Roberts, G. C. (1994). Perception of Success Questionnaire: Preliminary validation in an adolescent population. *Perceptual and Motor Skills*, *79*, 607–610.

Turner, R. C., & Lindsay, H. A. (2003). Gender differences in cognitive and noncognitive factors related to achievement in organic chemistry. *Journal of Chemical Education*, *80*(5), 563–568. doi:10.1021/ed080p563

Vallacher, R. R., & Wegner, D. M. (1987). What do people think theyre doing – action identification and human-behavior. *Psychological Review*, *94*(1), 3–15.

Van Der Linden, D., Sonnentag, S., Frese, M., & Van Dyck, C. (2001). Exploration strategies, performance, and error consequences when learning a complex computer task. *Behaviour & Information Technology*, *20*(3), 189–198.

Van Dyck, C., Van Hooft, E., De Gilder, D., & Liesveld, L. (2010a). Proximal antecedents and correlates of adopted error approach: A self-regulatory perspective. *The Journal of Social Psychology*, *150*(5), 428–451.

Van Dyck, C., Van Hooft, E., De Gilder, D., & Liesveld, L. (2010b). Proximal antecedents and correlates of adopted error approach: A self-regulatory perspective. *The Journal of Social Psychology*, *150*(5), 428–451. doi:10.1080/00224540903366743

Vandewalle, D. (1997). Development and validation of a work domain goal orientation instrument. *Educational and Psychological Measurement*, *57*(6), 995–1015.

VanLehn, K. (2011). The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems. *Educational Psychologist*, *46*(4), 197–221. doi:10.1080/00461520.2011.611369

Vanlehn, K. (1999). Rule-learning events in the acquisition of a complex skill: An evaluation of Cascade. *Journal of the Learning Sciences*, *8*(1), 71–125.

Vasilyeva, E., De Bra, P., & Pechenizkiy, M. (2008). Immediate Elaborated Feedback Personalization in Online Assessment. In *Times of Convergence. Technologies Across Learning Contexts* (Vol. 5192, pp. 449–460). Berlin, Heidelberg: Times of Convergence. Technologies Across Learning Contexts. doi:10.1007/978-3-540-87605-2_50

Vollmeyer, R., & Rheinberg, F. (2005). A surprising effect of feedback on learning. *Learning and Instruction*, *15*(6), 589–602. doi:10.1016/j.learninstruc.2005.08.001

Webb, J. M., Stock, W. A., & McCarthy, M. T. (1994). The effects of feedback timing on learning facts: The role of response confidence. *Contemporary Educational Psychology*, *19*, 251–265.

Weiner, B. (1994). Integrating social and personal theories of achievement striving. *Review of Educational Research*, *64*(4), 557–573.

Wigfield, A., & Cambria, J. (2010). Students' achievement values, goal orientations, and interest: Definitions, development, and relations to achievement outcomes. *Developmental Review*, *30*(1), 1–35. doi:10.1016/j.dr.2009.12.001

Wu, C.-C. (2012). The Cross-Cultural Examination of 3. *Procedia – Social and Behavioral Sciences*, *69*, 422–427. doi:10.1016/j.sbspro.2012.11.429

Wu, H. K., Krajcik, J. S., & Soloway, E. (2001). Promoting understanding of chemical representations: Students' use of a visualization tool in the classroom. *Journal of Research in Science Teaching*, *38*(7), 821–842. doi:10.1002/tea.1033

Yang, Y., & Cao, L. (2013). Differential influences of achievement approach goals and intrinsic/extrinsic motivation on help-seeking in e-learning. *Knowledge Management & E-Learning*, *5*(2), 153-169.

Zhao, B. (2011). Learning from errors: The role of context, emotion, and personality. *Journal of Organizational Behavior*, *32*(3), 435–463.

Zhou, M. (2013). University student's goal profiles and metacomprehension accuracy. *Educational Psychology*, *33*(1), 1–13.

Appendix A.

Simon Fraser University Ethics Approval

Minimal Risk Expedited Approval

SFU ENGAGING THE WORLD

OFFICE OF RESEARCH ETHICS

ORE Office of Research Ethics Balancing Research with Ethical Considerations

8888 University Drive
Discovery 2
1 778 782 3447 Manager
1 778 782 6593 Director
Email: dore@sfu.ca
http://www.sfu.ca/vpresearch/ethics

Date	File	Approval	Principal Investigator
19 July 2012	2012s0376]	Approved	Stewart, Jaclyn

Title	Start Date	End Date
The influence of goal and error orientations on learning from errors in organic chemistry online homework.	19 July 2012	19 July 2015

SFU Position	Department / School	Supervisor
Graduate Student	Education	Winne, Phil

Hello Jaclyn,

Your application has been categorized as 'Minimal Risk' and approved by the Director, Office of Research Ethics on behalf of the Research Ethics Board, in accordance with University Policy r20.01 (http://www.sfu.ca/policies/research/r20.01.htm)

The Research Ethics Board reviews and may amend decisions made independently by the Director, Chair or Deputy Chair at the regular monthly meeting of the Board.

Please acknowledge receipt of this Notification of Status by email to dore@sfu.ca and Include the file number as shown above as the first item in the Subject Line.

You should get a letter shortly. Note: All letters are sent to the PI addressed to the Department, School or Faculty for Faculty and Graduate Students. Letters to Undergraduate Students are sent to their Faculty Supervisor.

Good luck with the project,

Hal Weinberg, Director

Appendix B.

Questionnaire Items

Table 46. Elliot et al.'s (2011) Achievement Goal Questionnaire Items

Subscale	Item
Task-approach	1. To get a lot of questions correctly in this class. 2. To know the right answers to the questions on the exams in this class. 3. To answer a lot of questions correctly on exams in this class.
Task-avoidance	1. To avoid incorrect answers on the exams in this class. 2. To avoid getting a lot of questions wrong on the exams in this class. 3. To avoid missing a lot of questions on the exams in this class.
Self-approach	1. To perform better on the exams in this class than I have done in the past on these types of exams. 2. To do well on the exams in this class relative to how well I have done in the past on such exams. 3. To do better on the exams in his class than I typically do in this type of situation.
Self-avoidance	1. To avoid doing worse on the exams in this class than I normally do on these types of exams. 2. To avoid performing poorly on the exams in this class compared to my typical level of performance. 3. To avoid doing worse on the exams in this class than I have done on prior exams of this type.
Other-approach	1. To outperform other students on the exams in this class. 2. To do well compared to others in the class on the exams. 3. To do better than my classmates on the exams in this class.
Other-avoidance	1. To avoid doing worse than other students on the exams in this class. 2. To avoid doing poorly in comparison to others on the exams in this class. 3. To avoid performing poorly relative to my fellow students on the exams in this class.

Table 47. Rybowiak et al.'s (1999) and Adapted Error Orientation Questionnaire Items for Four Subscales

Subscale	Original	Adapted
Competence	When I have made a mistake, I know immediately how to correct it	1. When I have made a mistake, I know immediately how to correct it
	When I do something wrong at work, I correct it immediately	2. When I do something wrong when solving problem, I correct it immediately
	If it is at all possible to correct a mistake, then I usually know how to go about it	3. I usually know how to go about correct the mistakes I make
	I don't let go of the goal, although I may make mistakes	4. I don't let go of the goal, although I may make mistakes
Learning	Mistakes assist me to improve my work	1. Mistakes assist me to improve my knowledge and ability
	Mistakes provide useful information for me to carry out my work	2. Mistakes provide useful information about what I know and don't know
	My mistakes help me to improve my work	3. My mistakes help me to improve
	My mistakes have helped me to improve my work	4. My mistakes have helped me to improve my knowledge and ability
		5. Errors help me improve my knowledge and ability
Risk taking	If one wants to achieve at work, one has to risk making mistakes	1. If one wants to achieve in this class, one has to risk making mistakes
	It is better to take the risk of making mistakes than to 'sit on one's behind'	2. It is better to take the risk of making mistakes than not to attempt solving problems
	I'd prefer to err, than to do nothing at all	3. I'd prefer to make mistakes, than do nothing at all
		4. I typically choose problems to practice that I know I won't make a mistake on
	To get on with my work, I gladly put up with things that can go wrong	
Strain	I find it stressful when I err	1. I find it stressful when I make errors during problem solving
	I am often afraid of making mistakes	2. I am often afraid of making mistakes
	I feel embarrassed when I make an error	3. I feel embarrassed when making an error if my instructor will find out
	If I make a mistake at work, I 'lose my cool' and become angry	4. I feel embarrassed when making an error if my classmates will find out
	While working I am concerned that I could do something wrong	

Table 48. Rybowiak et al.'s (1999) and Adapted Error Orientation Questionnaire Items for Three Subscales

Subscale	Original	Adapted
Anticipation	In carrying out my task, the likelihood of errors is high	1. In solving problems, the likelihood of making errors is high
	Whenever I start some piece of work, I am aware that mistakes occur	2. Whenever I start solving a problem, I am aware that mistakes occur
	Most of the time I am not astonished about my mistakes because I expected them	3. Most of the time I am not astonished about my mistakes because I expect them
	I anticipate mistakes happening in my work	4. I anticipate mistakes happening when I am learning
	I expect that something will go from time to time	5. I expect that I will make errors from time to time
Covering up	Why mention a mistake when it's obvious?	1. I don't think there is a point in sharing the mistakes I make with others
	It is disadvantageous to make one's mistakes public	2. It puts me at a disadvantage to tell others about my errors
	I do not find it useful to discuss my mistakes	3. I do not find it useful to discuss my errors with my instructor
		4. I do not find it useful to discuss my errors with my classmates
	I would rather keep my mistakes to myself	5. I would rather keep my mistakes to myself
Communication	When I make a mistake at work, I tell others about it in order that they do not make the same mistake	1. When I make a mistake in solving a problem, telling others about it might prevent them from making the same mistake
	If I cannot rectify an error by myself, I turn to my colleagues	2. If I cannot correct an error by myself, I ask a classmate or peer for help
	If I cannot manage to correct a mistake, I can rely on others	3. If I cannot manage to correct an error by myself, I rely on others to help me
	When I have done something wrong, I ask others, how I should do it better	4. When I have made an error when solving a problem, I find out how I should do it correctly

Table 49. Rybowiak et al.'s (1999) and Adapted Error Orientation Questionnaire Items for Two Subscales

Subscale	Original	Adapted
Thinking	After I have made a mistake, I think about how it came about	1. After I have made a mistake, I think about how it came about
	If something goes wrong at work, I think it over carefully	2. If I make an error while solving a problem, I think it over carefully.
	After a mistake has happened, I think long and hard about how to correct it	3. After I have made an error, I think long and hard about how to correct it.
	When a mistake occurs, I analyze it thoroughly	4. When I make an error, I analyze it thoroughly
	I often think "How could I have prevented this?"	
Motivation	N/A	1. When I make an error, I feel motivated to correct it
		2. When I make a lot of mistakes when solving a problem, I feel discouraged
		3. When deciding what to study, I choose the problems that I wont' make many errors on
		4. I notice when I am making the same mistakes on more than on problem
		5. If I make the same mistake twice, I think I might never learn how to do it properly

Appendix C.

Participant Contact

Included here is the text from the email I sent to students announcing the release of a new Sapling online homework assignment on Oct. 22nd, 2011. Wording for this email was chosen carefully to avoid coercion as well as avoiding promises of improved performance if students respond to the confidence and certainty questions. This information was also posted on the learning management system to increase the likelihood students were notified of why the judgment questions were appearing on their homework questions.

“Hi everyone,

The next Sapling assignment is now available. It covers SN1/SN2/E1/E2. The due date is Oct. 30th at 11:55 PM. Complete your assignment early to avoid dealing with a potentially slow system close to the deadline.

You will notice that two of this week's Sapling problems have "reflection" type questions added to them. I have added these because research shows that thinking about your answers while solving problems is important. You don't have to answer the reflections questions to get the points for the Sapling problems, but I recommend you do.

Have a great weekend!

Jackie”

For the course surveys, here is the text for the email announcement of these surveys that was sent out on Oct. 1st, 2011. Students were also told at the start of term that “bonus surveys” will be available and that they would receive more information about this via email.

“Dear CHEM 233 Student,

In order to learn more about the group of students who take CHEM 233, I would like to request your participation in some online questionnaires. It is important to hear from all of you to make the results more meaningful, so we are giving a small bonus for completing the surveys during the available dates (1% for the four short surveys). We will use the results to make improvements in the course from year to year. Please answer the questions with respect to how you feel about your learning in this course (CHEM 233).

Your instructors will NEVER see your individual responses. The results will only be reported anonymously as group averages; your identity will never be revealed.

Be sure to enter your student ID correctly otherwise you won't receive the bonus mark.

Links to the surveys are on the course Vista site. The surveys are staggered, so only half of you will have immediate access to the surveys. Each of you has a personalized folder on Vista with the survey submission dates, as you will see when you click on the "Surveys & Quizzes" Vista folder.

If you have any questions, concerns, or comments please contact me at (email address).

Thank you very much for your participation!

Jackie Stewart”

Appendix D.

Variable Distributions.

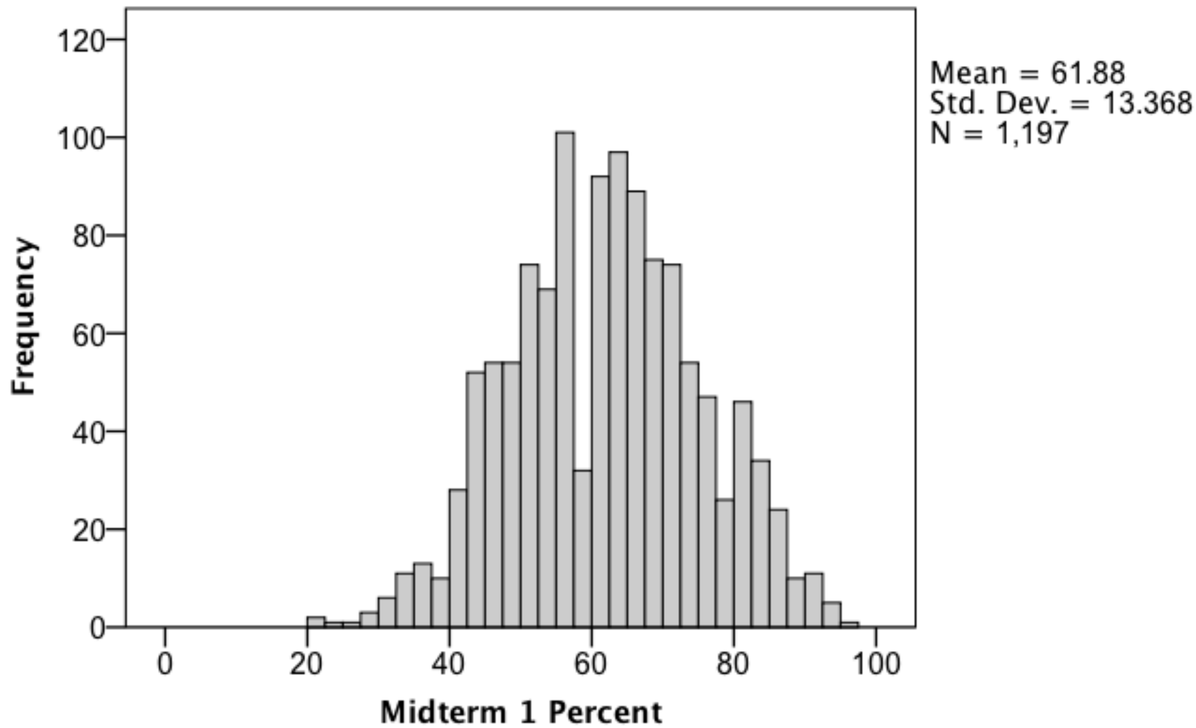

Figure 30. Distribution of students' midterm 1 grades as a percentage.

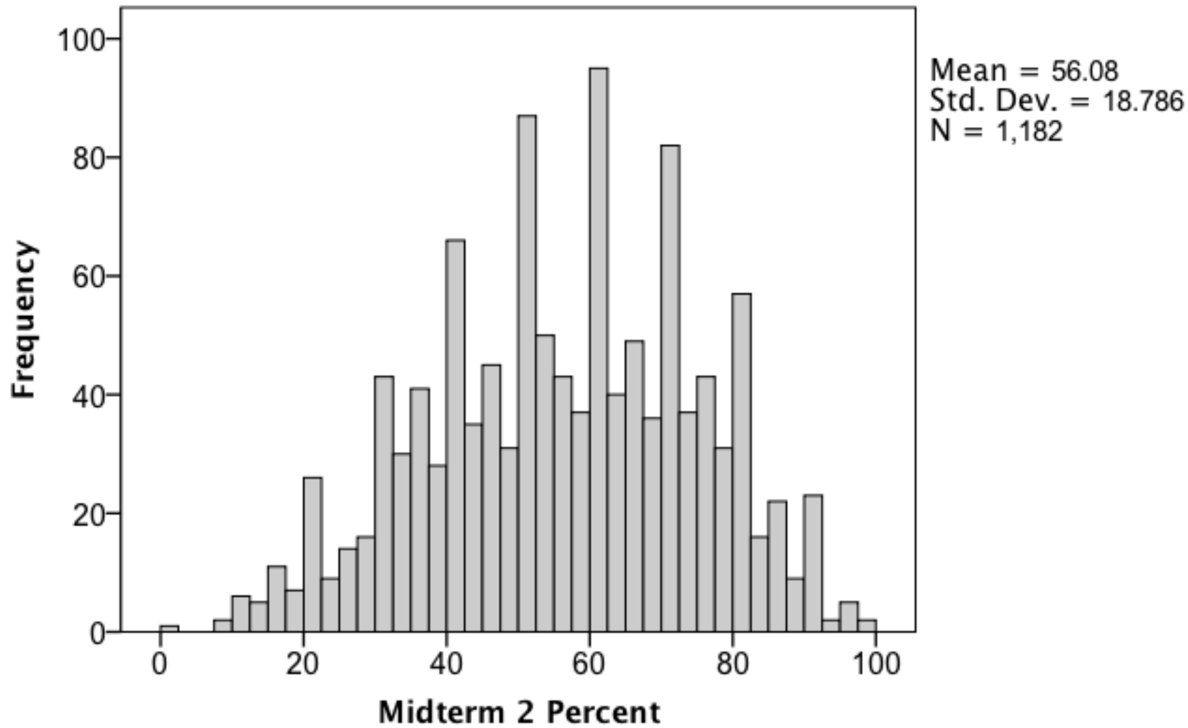

Figure 31. Distribution of students' midterm 2 grades as a percentage.

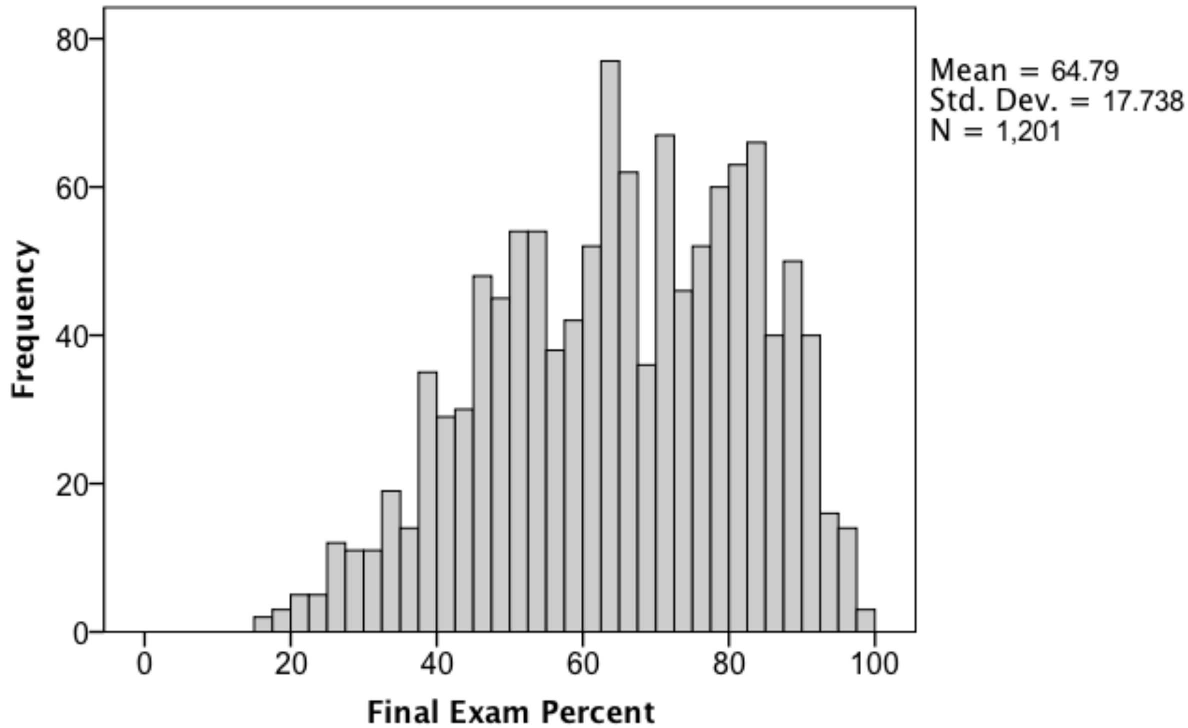

Figure 32. Distribution of students' final exam grades as a percentage.

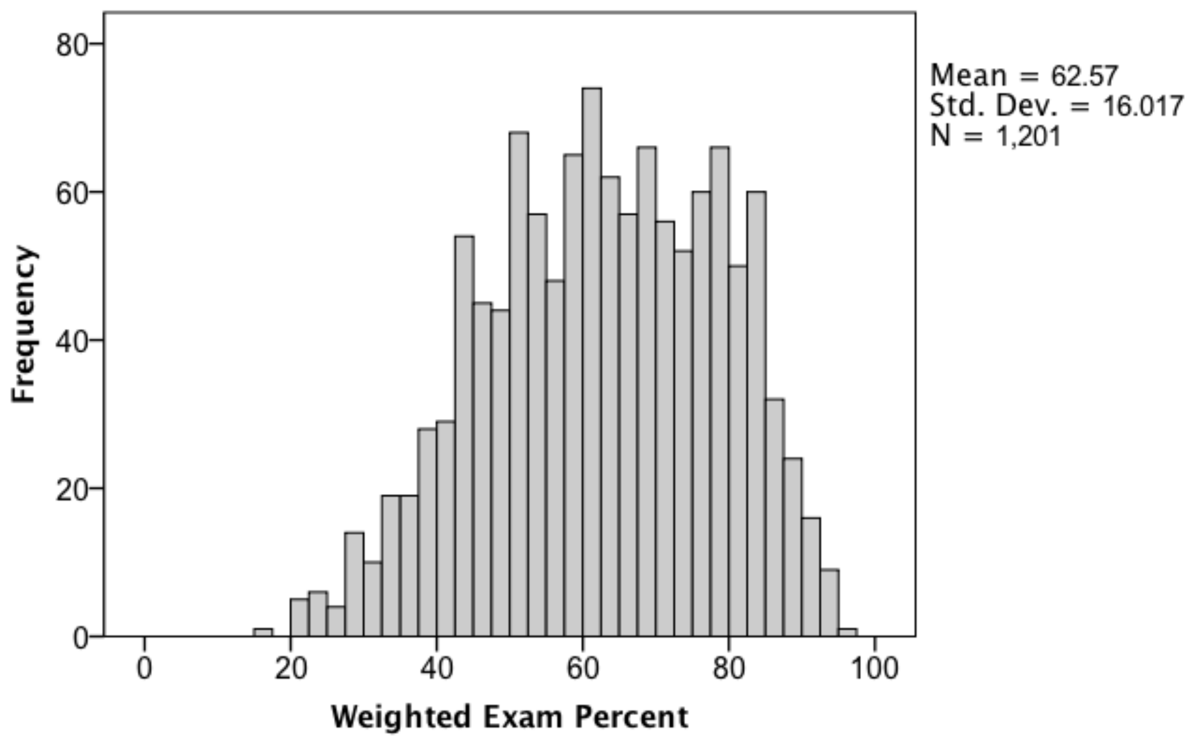

Figure 33. Distribution of the weighted exam scores (composite midterm and final exam).

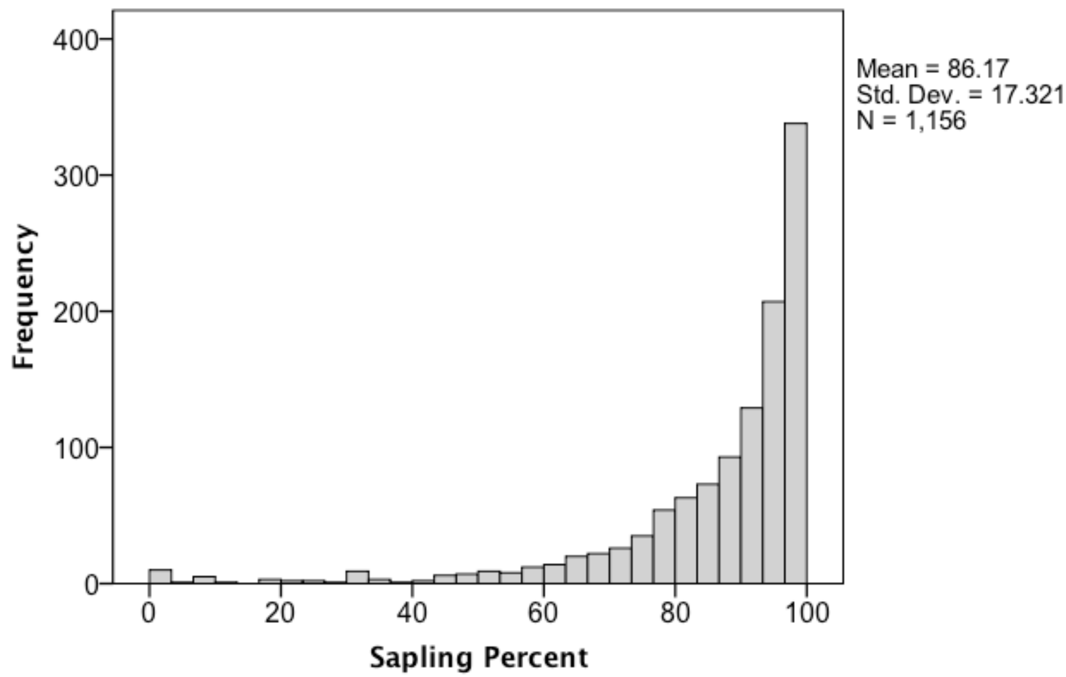

Figure 34. Distribution of students' Sapling Learning scores (SapPercent) for all students who created an account on the Sapling Learning site.

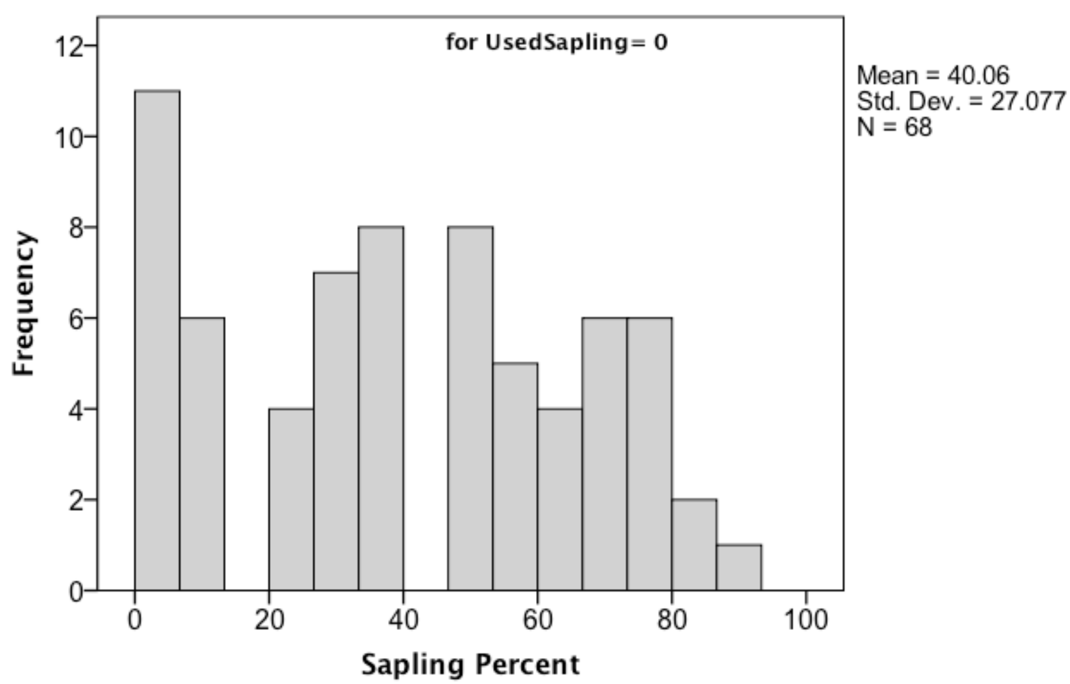

Figure 35. Distribution of Sapling Learning scores for the 68 students whose scores were not used in calculating their overall grade.

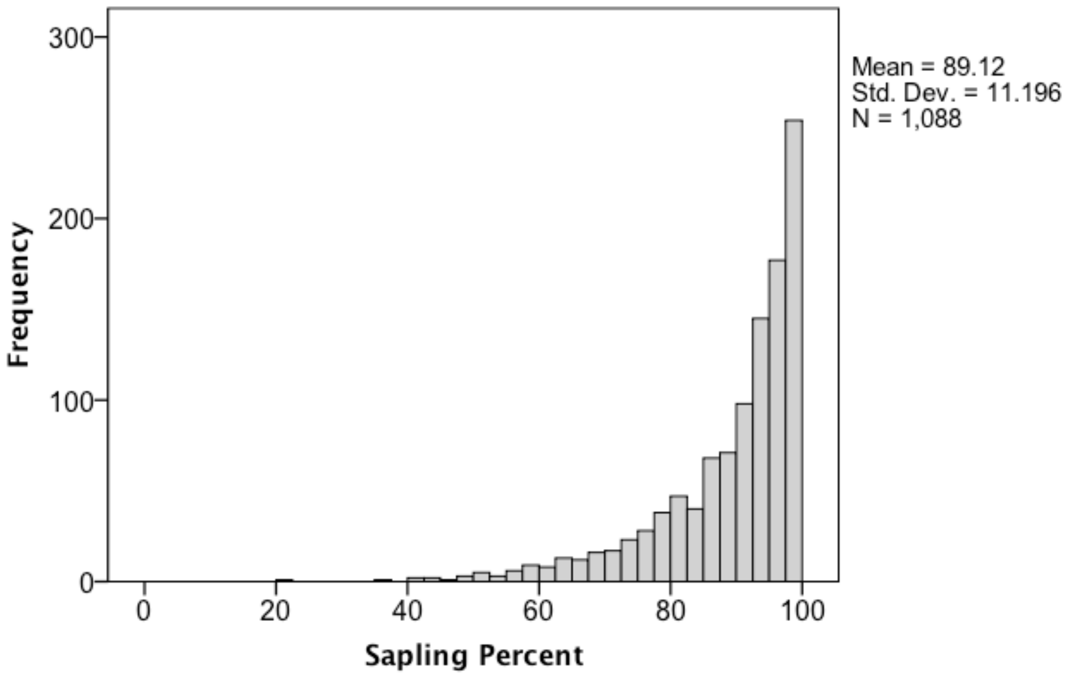

Figure 36. Distribution of Sapling Learning scores for the 1088 students whose grade was used in the calculation of their overall grade.

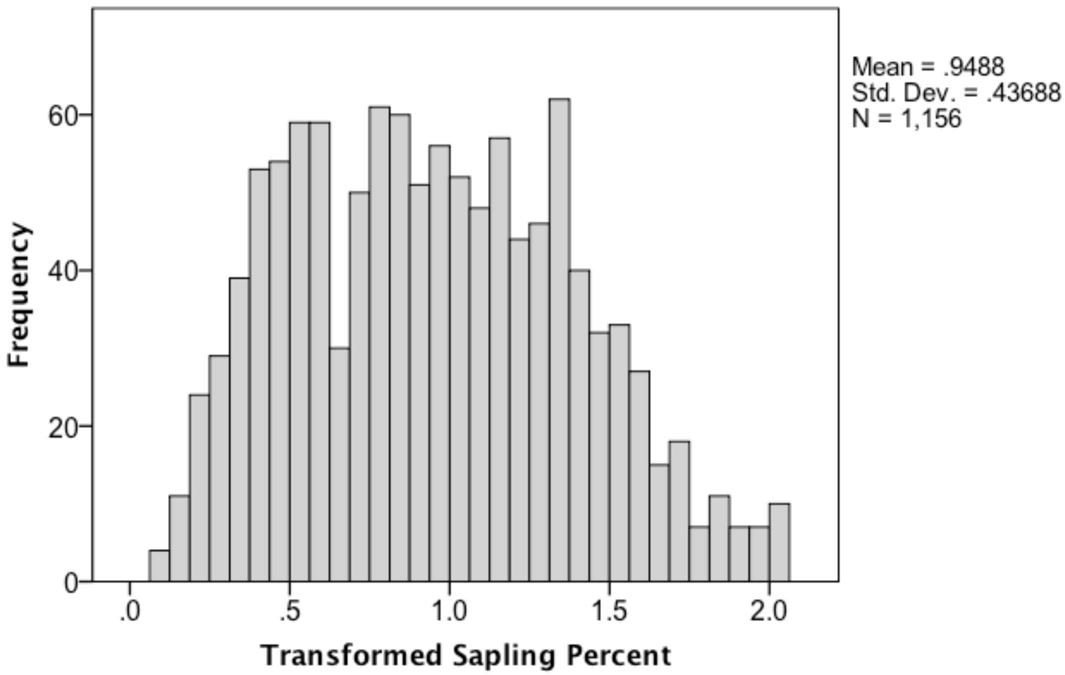

Figure 37. Distribution of transformed Sapling percentage grade.

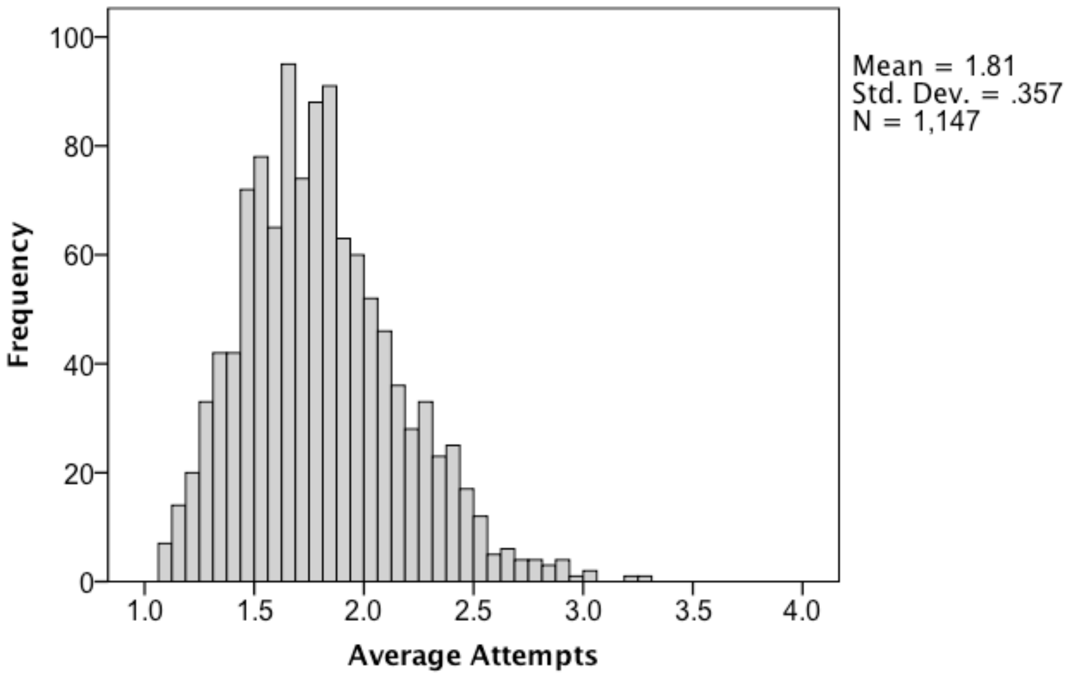

Figure 38. Distribution of the average number of attempts for Sapling Learning questions attempted one or more times.

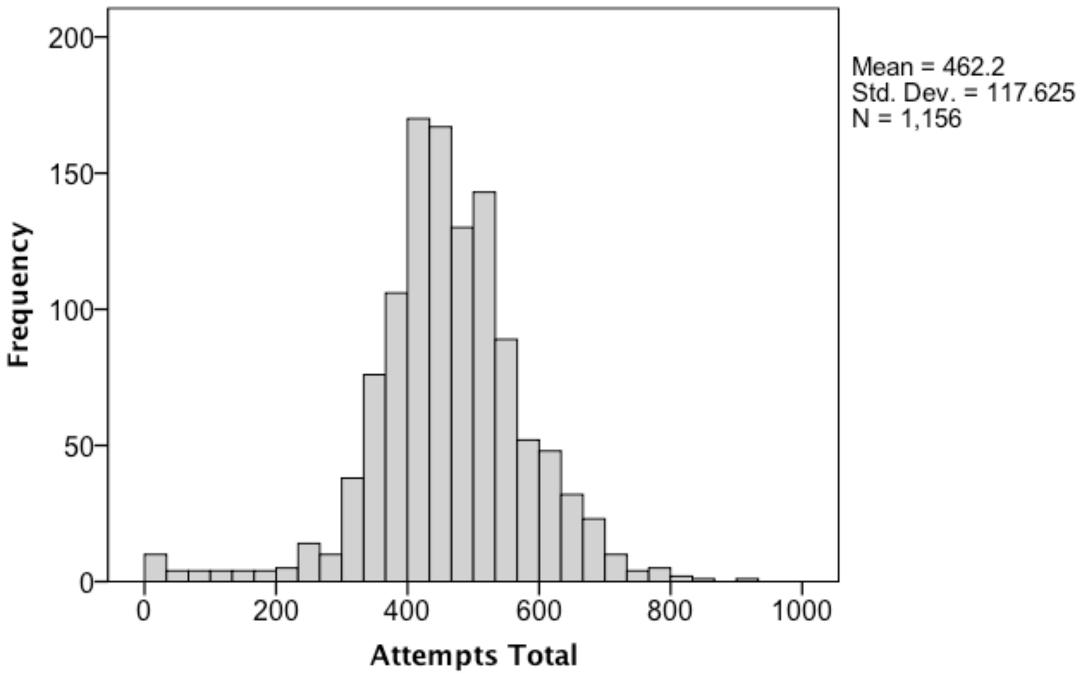

Figure 39. Distribution of the total number of attempts students took on questions over the entire term.

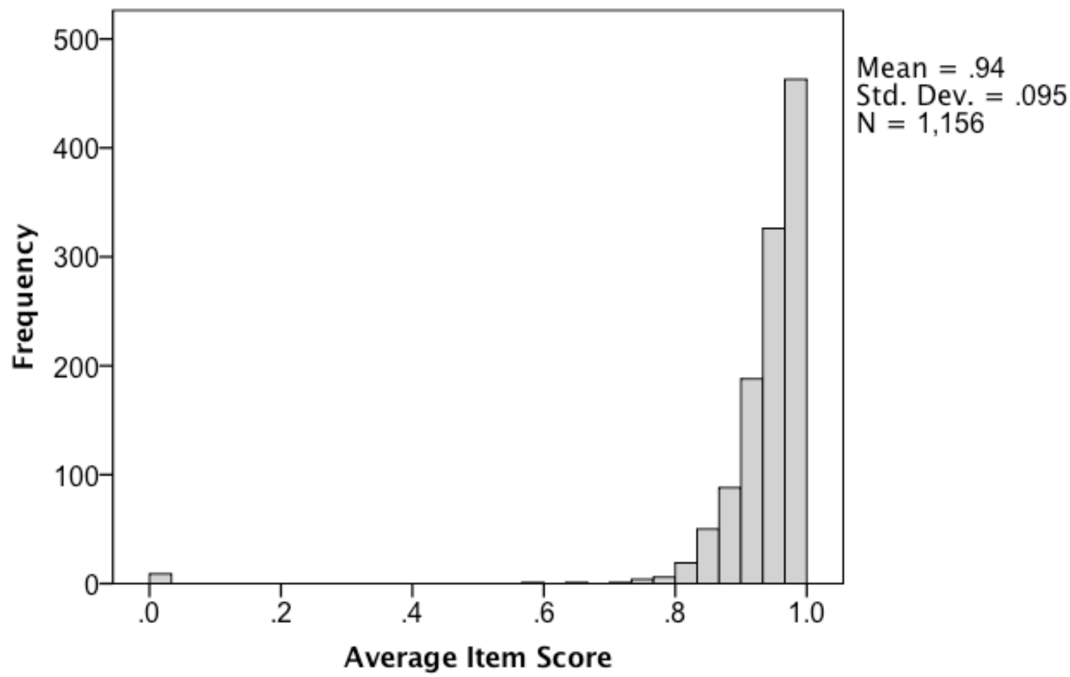

Figure 40. Distribution of average item score.

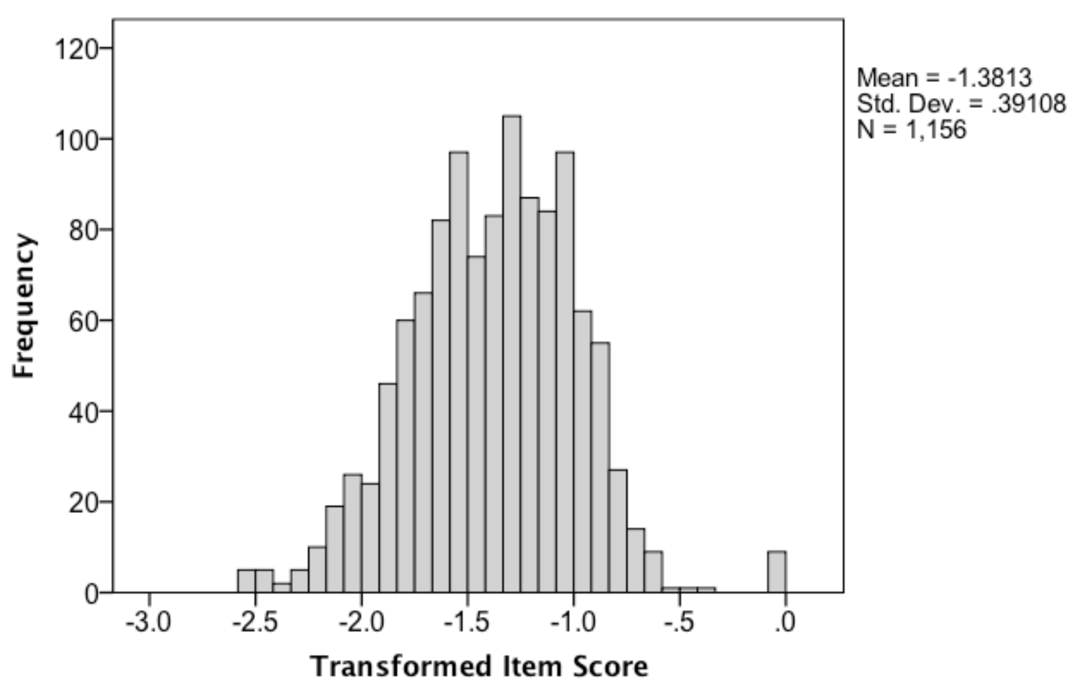

Figure 41. Histogram of transformed average item score.

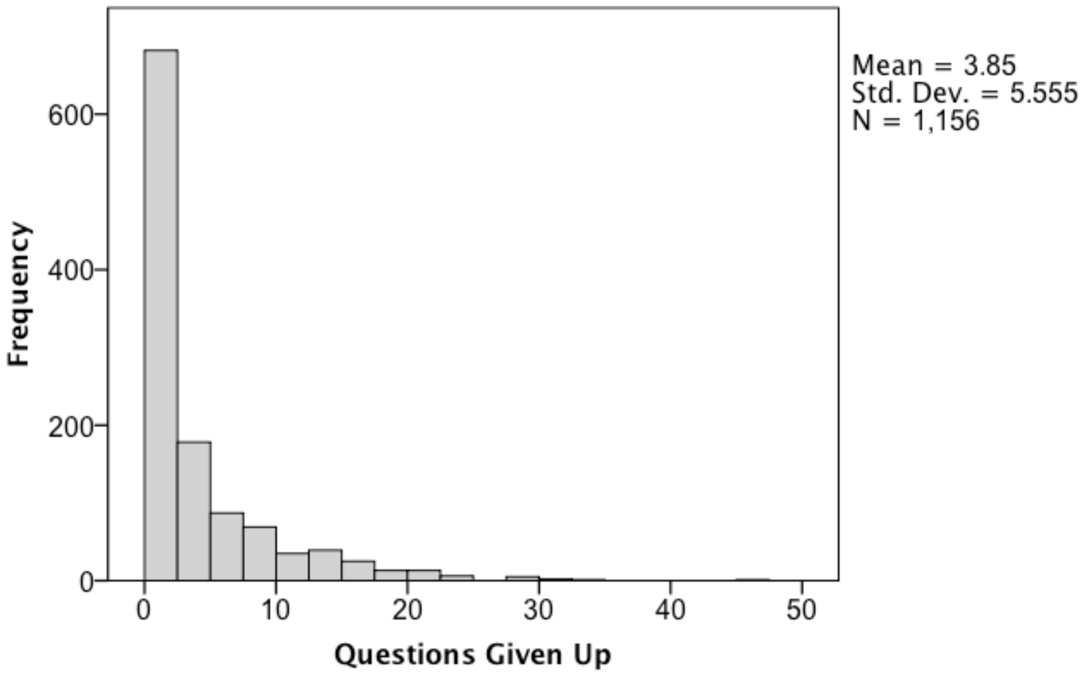

Figure 42. Distribution of the number of questions given up on, for those who submitted a response (attempted) the question.

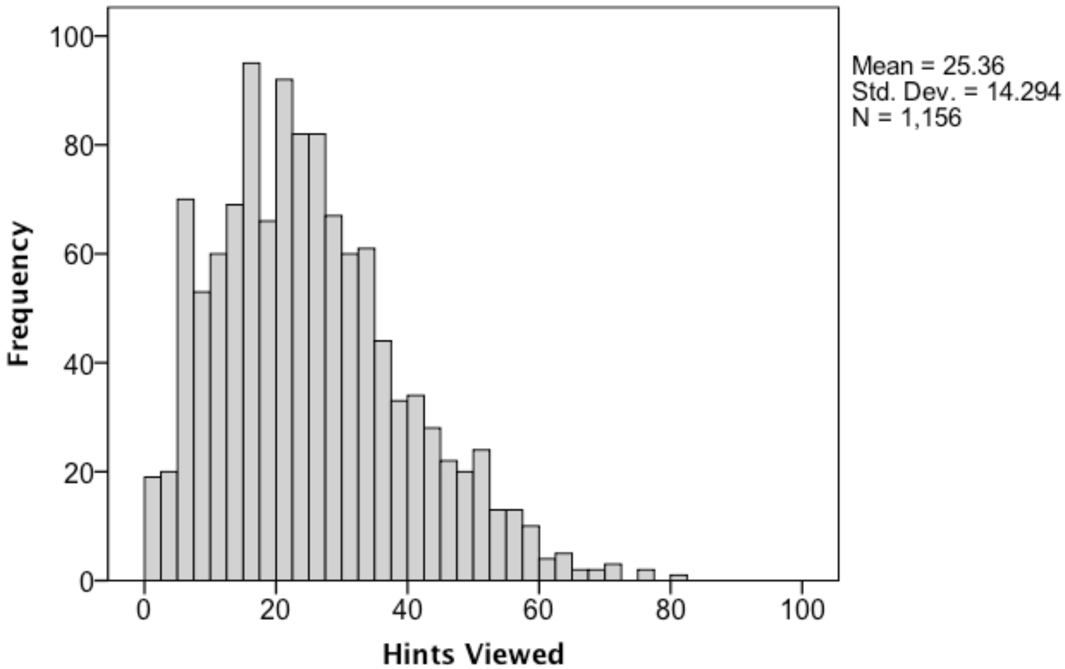

Figure 43. Distribution of the percent of items attempted for which students viewed the hint.

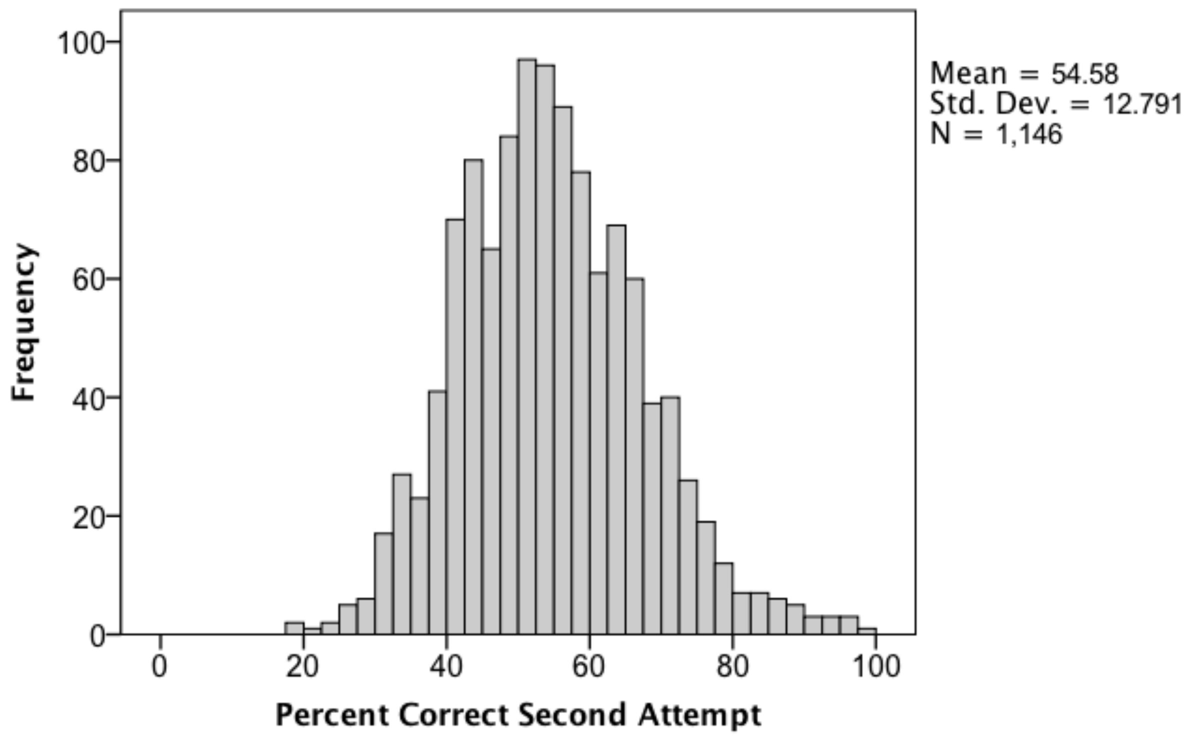

Figure 44. Distribution of the percentage of questions students answered correctly following an incorrect first attempt. Only students who submitted at least one question are included in this graphic.

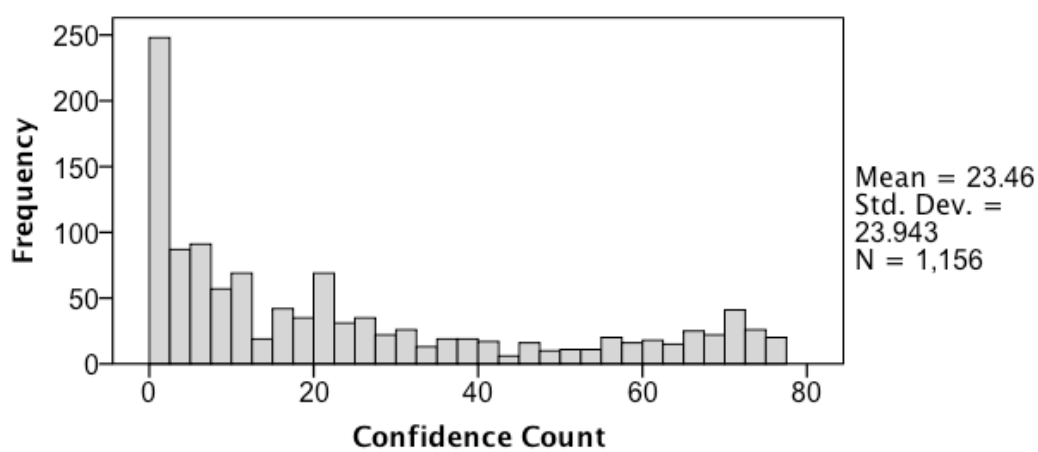

Figure 45. The distribution of the number of confidence and certainty judgments made by students.

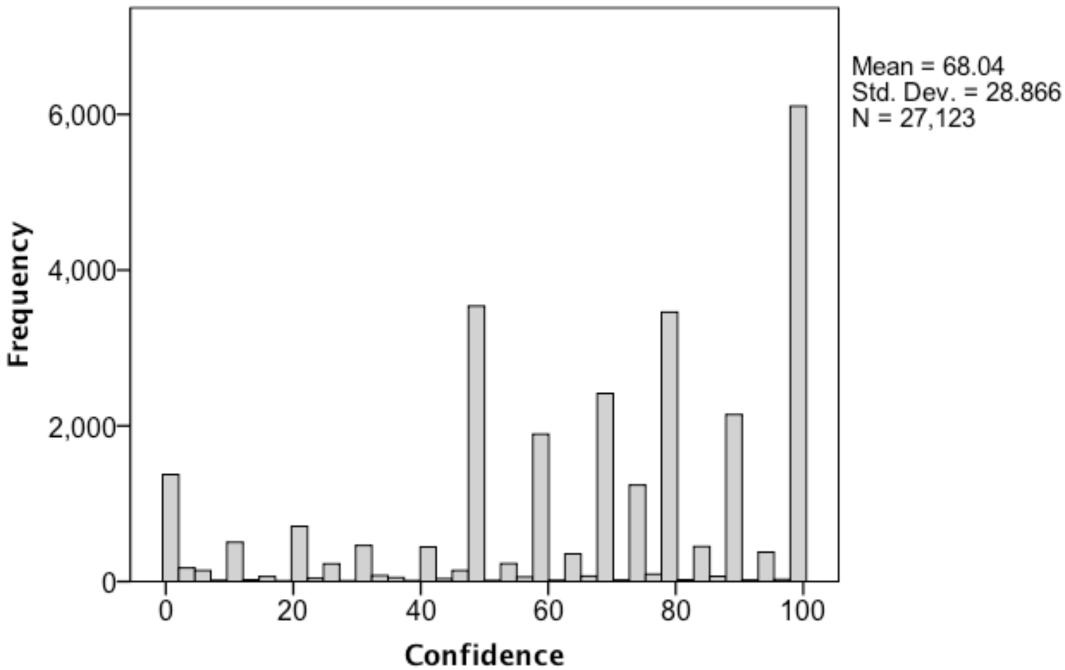

Figure 46. Histogram showing the distribution of confidence values. *N* is the number of confidence judgments made by all students over the 76 questions.

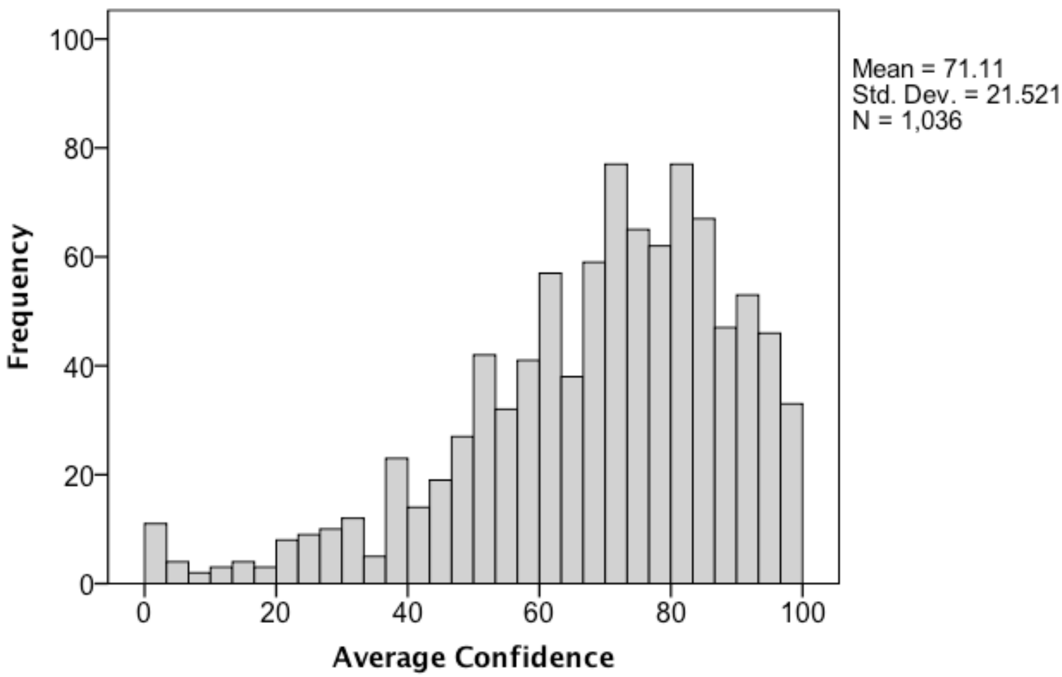

Figure 47. The distribution of students' average confidence predictions. Only students who made at least one prediction are included in the analysis.

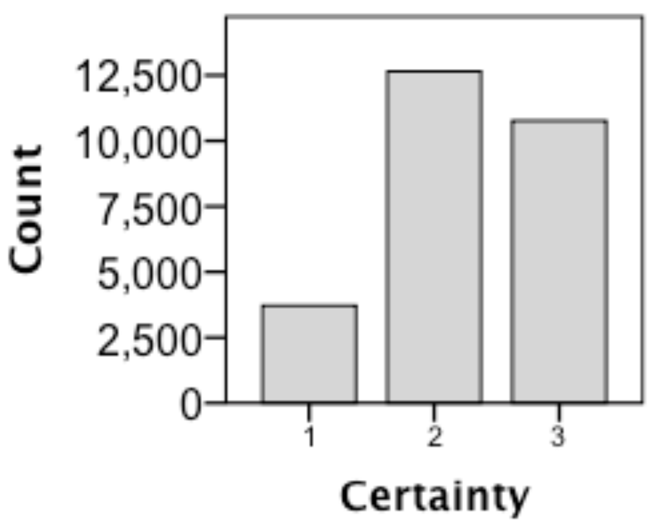

Figure 48. Distribution of certainty ratings.

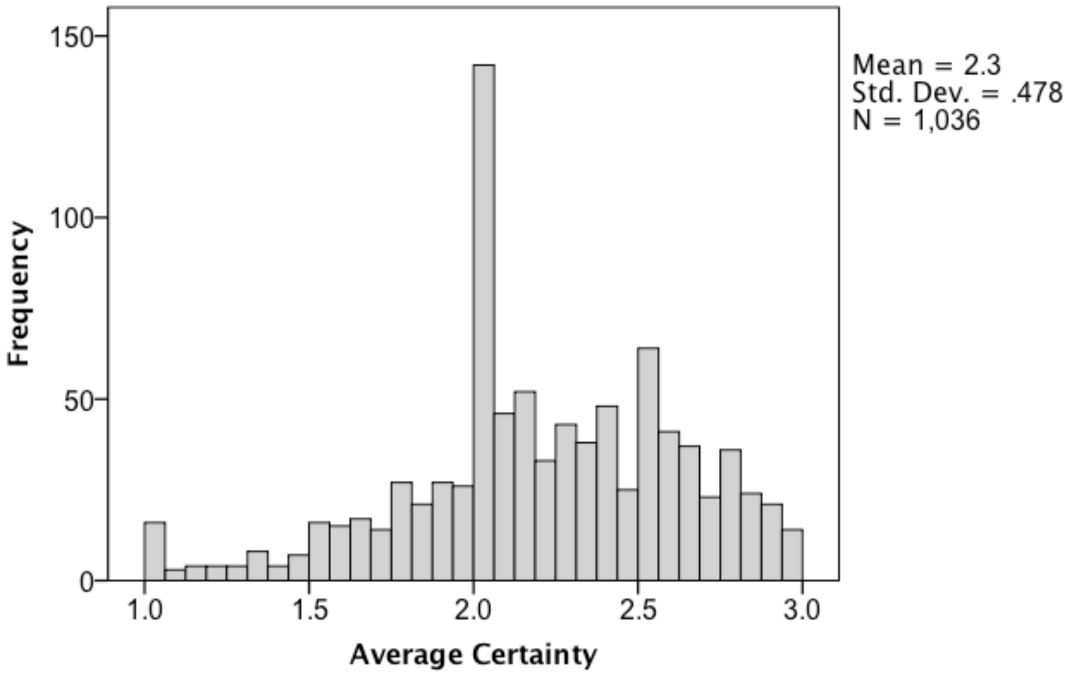

Figure 49. Distribution of average certainty judgments. Only students who made at least one prediction are included in this figure.

Appendix E.

Missing Data Analysis

To explore the impact of missing data on the descriptive statistics of the data, I constructed a table of means, standard deviations, and variances for the AGQ variables divided by groups: no missing, 1 missing, and 2-6 missing.

Table 50. Means, Standard Deviations, and Variances (σ^2) for AGQ Variables by Missing Data Group

Variable	M	SD	σ^2	M	SD	σ^2	M	SD	σ^2
	No Missing Data N = 888			Data Missing for 1 Variable N = 108-120			Data Missing for 2-6 Variables N = 15-20		
TaskApp1	5.93	1.12	1.26	5.78	1.23	1.51	5.47	1.07	1.14
TaskApp2	5.89	1.17	1.38	5.80	1.17	1.36	5.65	1.17	1.37
TaskApp3	6.01	1.11	1.24	5.89	1.22	1.48	5.89	1.05	1.10
TaskAvo1	5.83	1.15	1.33	5.62	1.19	1.42	5.60	1.55	2.40
TaskAvo2	5.83	1.28	1.65	5.64	1.49	2.23	5.00	1.50	2.24
TaskAvo3	5.80	1.32	1.73	5.88	1.33	1.76	5.13	1.30	1.70
SelfApp1	5.69	1.28	1.64	5.30	1.55	2.41	5.60	1.30	1.69
SelfApp2	5.63	1.33	1.76	5.36	1.43	2.03	5.47	1.47	2.15
SelfApp3	5.49	1.36	1.85	5.35	1.44	2.07	4.94	1.89	3.56
SelfAvo1	5.52	1.42	2.02	5.15	1.53	2.34	5.00	1.84	3.37
SelfAvo2	5.66	1.29	1.66	5.46	1.54	2.38	5.75	1.57	2.47
SelfAvo3	5.63	1.28	1.63	5.36	1.52	2.31	4.89	1.53	2.34
OtherApp1	4.81	1.66	2.75	4.90	1.64	2.69	4.80	1.96	3.85
OtherApp2	5.39	1.46	2.14	5.40	1.48	2.19	5.47	1.59	2.52
OtherApp3	5.06	1.59	2.52	5.16	1.59	2.54	5.40	1.76	3.10
OtherAvo1	5.29	1.50	2.26	5.21	1.54	2.38	5.22	1.87	3.48
OtherAvo2	5.24	1.55	2.42	5.18	1.51	2.28	5.16	1.74	3.03
OtherAvo3	5.37	1.49	2.23	5.37	1.50	2.25	5.06	1.85	3.43

The results of Little's MCAR test show statistically detectable differences between groups (χ^2 = 732.562, *df* = 568, *p* < .001) indicating that the data are not missing completely at random (MCAR). Most techniques for handling missing data work best when missing data is MCAR. Separate variance t-tests for variables with more than 1% of data missing were conducted using SPSS Missing Value Analysis. There exists a systematic relationship between the missingness on SelfApp1 and TaskAvo2 (p = .050). Missingness in SelfAvo3 was significantly related to missingness on TaskApp1 (p = .042), TaskApp3 (p = .025), TaskAvo1 (p = .003), SelfApp2 (p < .001), and SelfAvo2 (p = .032). Since some of these relationships are statistically detectable, the missing data may be missing not at random (MNAR). However, since experts doubt that this test is completely reliable (citation needed), and the consensus from the literature indicates that estimation methods may be used in MNAR cases without a high likelihood of biased results (more citations).

To further explore the differences in distributions for responses with no missing data and those with some missing data, I performed an analysis of variance with the AGQ item scores as dependent variables and the number of missing data (zero, 1, or 2 and 3) as factors. The means plots are shown in Figure 50. Since the group sizes are unequal, the null hypothesis (that the means are equal) may be rejected incorrectly. Out of the 18 variables, 4 showed a statistically detectable difference in the means between two levels according to a post-hoc Tukey test. Only approximately one would be expected by chance after a Bonferonni correction, p < .0027.

Table 51. Significant Analysis of Variance Results for Four Variables

Variable	Factors that differ	Mean Diff.	Standard Error	Sig.	Lower bound 95% CI	Upper bound 95% CI
Task-avoidance 2	No missing and more than 1 missing	.834*	0.312	0.021	0.1	1.57
Self-approach 1	No missing and 1 missing	.383*	0.132	0.01	0.07	0.69
Self-avoidance 1	No missing and 1 missing	.369*	0.144	0.028	0.03	0.71
Self-avoidance 3	No missing and more than 1 missing	.743*	0.312	0.046	0.01	1.47

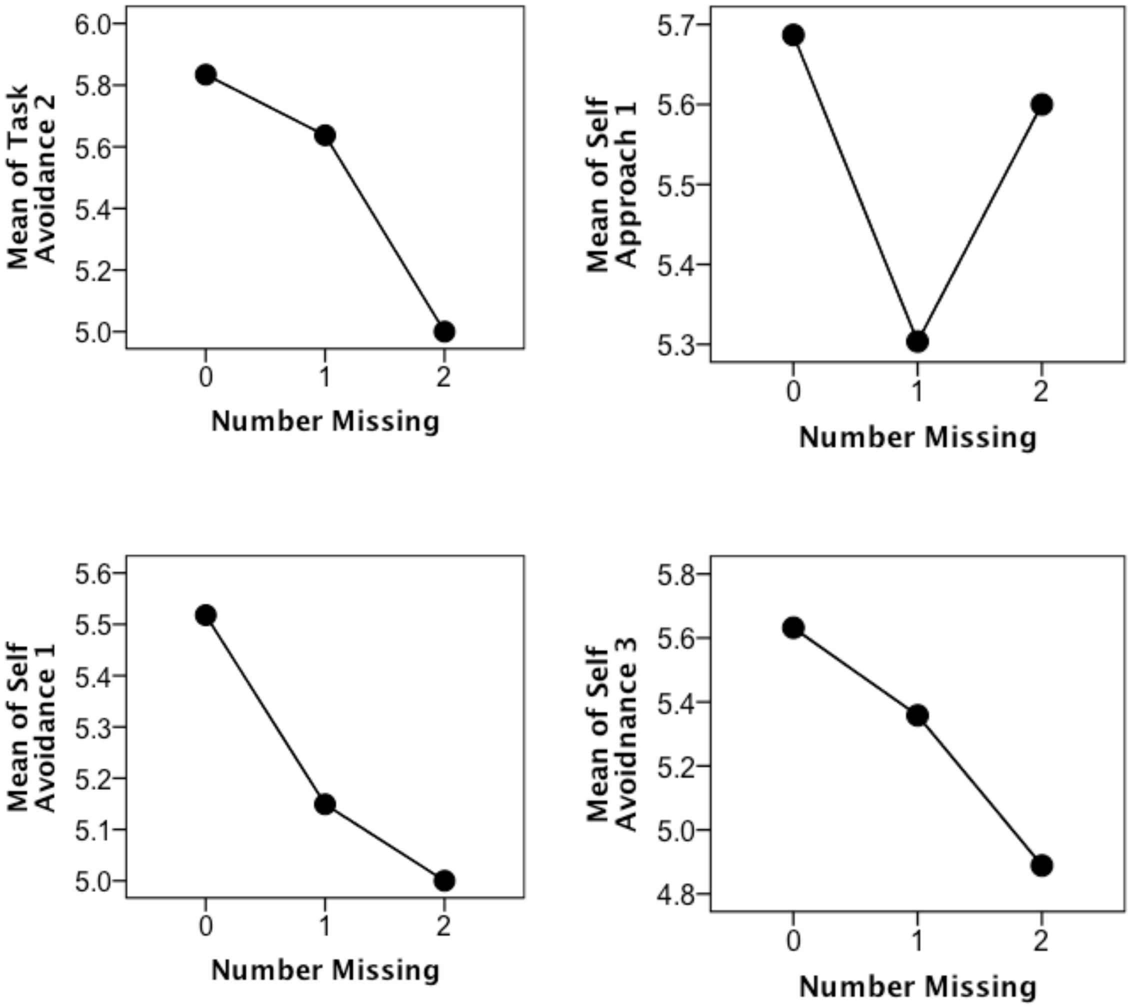

Figure 50. Means plots for the four AGQ variables with significant mean differences.

Since the AGQ is on a seven-point scale, the mean differences above could indicate that the group of students who responded to the survey and left out one or more responses are somehow different than those who responded to all of the questions. One of the goals of missing data analysis is to deal with the missing data in a way that does not distort the results of analysis using the dataset. Listwise deletion of 140 or more cases seems like a poor way to address this issue since it would decrease statistical power and distort the sample relative to the population of students enrolled in the course. Tabachnick and Fidell (2007) state that if less than 5% of the data is missing from a large dataset, any method of replacing missing data would probably have a similar result. Although no individual variable possesses more than 5% missing data points, the overall percentage missing for this data set is 6.9%. Thus, there is a strong case for using a method other than listwise deletion. I suspect that most of the missing data is because of accidental skipping, since most students only missed one item if they missed any at all. Since there are three items on each subscale, I judged it is unlikely that students would consciously object to responding to only one item.

Are the student who missed items somehow different than those who did not miss any items? Perhaps the degree of missingness correlates to ability to pay attention to detail, which is important for learning in demanding university-level courses. Since participants are likely to be only somewhat motivated to complete the questionnaire items, missing responses are not surprising.

I also explored if students with missing responses had different distributions on dependent variables such as course grade, online homework grade, Sapling Learning participation. There was no effect of missing responses on these dependent variables according to the ANOVA. Due to the nature of this research, I strongly believe that survey respondents simply missed some items by accident. The only reasonable hypothesis is that students who did not complete every item on the surveys are less conscientious than others, but without a direct measure of this it is not possible to test. Since it does not appear as though the students with missing data have different scores on other important measures, I will treat the missing data as missing at random (MAR) and ignorable.

Since the data is not MCAR, listwise deletion may introduce biased parameter estimates in the confirmatory factor analysis, and it could bias estimates of the standard errors (Allison, 2003). When a substantial portion of the dataset is deleted, standard errors will be larger than they would have been. Hypothesis tests will have lower power. Another option is pairwise deletion, but this can have the outcome of producing a not positive definite covariance matrix for analysis in the structural equation model (SEM) used in confirmatory factor analysis. Since SEM requires manipulations of the matrix, it must be positive definite. Pairwise deletion also poses a question about which sample size to use to calculate standard errors, since each correlation and covariance may be calculated with a different sample size. Maximum likelihood imputation methods such as expectation-maximization (EM) can be used for data where missing data is assumed to be MAR.

Outliers in Achievement Goal Data

Using the Mahalanobis distance, 100 cases were identified as outliers at $p < 0.01$. Upon inspection, it was found that these cases included both low (1) and high (5-7) responses on items in the hypothesized same construct. Since many participants may have successfully answered the screening question but not spent very much time and effort to complete the questionnaire, some outliers could reasonably be deleted from the dataset. To estimate the impact of these outliers on the CFA results, the analysis was run with and without the outliers.

Table 52. Model Fit Summary for CFA With and Without Outliers

Model	NPAR	CMIN	DF	P	CMIN/DF	TLI	CFI	RMSEA
With outliers	51	913.805	120	.000	7.616	.920	.938	.080
Saturated model	171	.000	0				1.000	
Independence model	18	12880.984	153	.000	84.189	.000	.000	.285
Without outliers	51	940.383	120	.000	7.837	.924	.941	.086
Saturated model	171	.000	0				1.000	
Independence model	18	13971.569	153	.000	91.317	.000	.000	.312

Note: NPAR is the number of distinct parameters; CMIN is the chi-square measurement; DF is degrees of freedom; TLI is the Tucker-Lewis index; CFI is the comparative fit index; and RMSEA is the root mean square error of approximation.

The impact of the 100 outliers on fit statistics is minimal. More concerning is that following removal of the outliers, AMOS identified a new set of outliers. To address outliers without eliminating probable valid data, only one outlier with a Mahalanobis distance very different from others were removed for subsequent analyses. The Mahalanobis distance of 169.786 is quite far away from the remaining values that are fairly closely spaced together (Table 53).

Table 53. The Largest Mahalanobis D-Squared Values

Observation number	Mahalanobis d-squared	p1	p2	Omitted or retained?
1026	169.786	0	0	Omitted
939	127.124	0	0	Retained
262	124.75	0	0	Retained
415	119.443	0	0	Retained
22	110.278	0	0	Retained
261	109.537	0	0	Retained
902	105.345	0	0	Retained
280	99.431	0	0	Retained
653	98.674	0	0	Retained
460	92.797	0	0	Retained
737	86.27	0	0	Retained
285	85.082	0	0	Retained
346	84.238	0	0	Retained
896	80.311	0	0	Retained
382	79.804	0	0	Retained
858	79.668	0	0	Retained
397	79.626	0	0	Retained

Upon inspection, observation number 1026 had several responses of 1 and 7 for items in the same construct. Observation number 939 also had wide variation in response values for items in the same construct, but it was less clear the responses are likely due to not reading or considering the statements.

Missing Data EOQ

To decide if the missing data should be treated as missing completely at random MCAR, MAR, or MNAR, I determined the univariate distributions for each variable using a dummy variable with three groups: those without any missing data (N = 788), those with one missing data point (N = 163), and those with missing data for two or more points (N = 52). I compared the three histograms for each variable as well as the normality statistics (skewness and kurtosis). Table **54** contains the descriptive statistics for each EOQ variable by group, Table **55** contains the skewness statistics for each EOQ variable by group, and Table **56** contains the kurtosis statistics for each EOQ variable by group.

Table 54. Means, Standard Deviations, and Variances (σ^2) for EOQ Items by Missing Data Group

Variable	M	SD	σ^2	M	SD	σ^2	M	SD	σ^2
	No Missing Data N = 788			Data Missing for 1 Variable N = 157-166			Data Missing for 2-6 Variables N = 46-54		
Competence 1	2.42	0.88	0.77	2.44	0.88	0.78	3.78	0.92	0.84
Competence 2	3.52	0.86	0.74	3.52	0.88	0.77	3.28	0.83	0.70
Competence 3	3.1	0.91	0.83	3.20	0.92	0.85	3.57	1.04	1.08
Competence	3.72	0.87	0.76	3.85	0.85	0.72	3.43	0.84	0.71
Learning 1	4.05	0.86	0.74	4.09	0.84	0.71	3.42	0.91	0.82
Learning 2	4.18	0.79	0.63	4.20	0.81	0.66	3.62	1.18	1.38
Learning 3	4.09	0.84	0.71	4.10	0.84	0.70	3.88	0.82	0.68
Learning 4	4.04	0.81	0.65	3.99	0.83	0.69	3.54	0.89	0.78
Learning 5	4.03	0.84	0.70	4.12	0.76	0.57	2.58	1.13	1.27
Risk 1	3.76	0.93	0.87	3.87	0.94	0.88	2.59	1.06	1.13
Risk 2	4.32	0.80	0.65	4.35	0.76	0.58	3.49	0.99	0.98
Risk 3	3.83	1.04	1.09	3.93	1.00	1.01	3.76	0.97	0.94
Risk 4	3.82	1.13	1.28	3.81	1.18	1.38	3.98	0.76	0.58
Strain 1	3.46	1.06	1.13	3.35	1.13	1.27	3.49	1.10	1.22
Strain 2	3.04	1.22	1.48	3.06	1.21	1.47	3.33	1.09	1.19
Strain 3	2.79	1.16	1.35	2.61	1.20	1.44	3.06	1.03	1.06
Strain 4	2.67	1.23	1.51	2.60	1.22	1.48	2.51	1.07	1.14
Anticipation 1	3.24	0.94	0.88	3.23	0.92	0.84	2.53	1.12	1.25
Anticipation 2	3.57	0.94	0.88	3.65	0.97	0.93	3.57	0.99	0.97
Anticipation 3	3	0.96	0.93	2.92	0.99	0.98	3.65	1.09	1.19
Anticipation 4	3.7	0.95	0.90	3.70	0.95	0.90	3.5	0.92	0.84
Anticipation 5	3.63	1.03	1.06	3.55	1.07	1.15	4.1	0.87	0.76
Covering 1	2.09	1.06	1.13	2.26	1.15	1.32	2.75	1.12	1.25
Covering 2	2.07	1.11	1.24	2.18	1.27	1.61	3.77	0.92	0.85
Covering 3	2.08	1.10	1.21	2.10	1.09	1.20	3.56	1.09	1.19
Covering 4	2.04	1.12	1.24	2.03	1.11	1.24	3.4	1.13	1.27
Covering 5	2.52	1.10	1.21	2.59	1.14	1.29	3.13	0.96	0.93
Communication 1	3.82	0.90	0.81	3.82	1.04	1.08	3.85	0.89	0.78
Communication 2	3.64	1.09	1.19	3.67	1.10	1.21	2.91	1.15	1.32
Communication 3	3.38	1.09	1.20	3.37	1.10	1.21	3.64	0.98	0.97
Communication 4	4.09	0.81	0.66	4.22	0.78	0.61	3.15	0.91	0.82
Thinking 1	3.63	0.90	0.81	3.70	0.90	0.82	3.77	0.93	0.87
Thinking 2	3.94	0.78	0.60	3.96	0.80	0.64	2.55	1.23	1.52
Thinking 3	3.63	0.93	0.86	3.74	0.91	0.83	2.51	1.12	1.26
Thinking 4	3.69	0.91	0.83	3.74	0.97	0.93	2.77	1.01	1.03
Motivation 1	3.87	0.90	0.81	3.94	0.89	0.79	3.43	1.01	1.02
Motivation 2	2.53	1.13	1.27	2.48	1.11	1.22	3.51	0.87	0.76
Motivation 3	3.93	1.11	1.24	3.85	1.13	1.27	3.42	0.82	0.67
Motivation 4	3.49	0.95	0.91	3.49	0.99	0.98	3.56	1.09	1.20
Motivation 5	3.91	1.21	1.46	3.72	1.29	1.67	3.26	1.05	1.10

Table 55. Skewness Statistics and Standard Errors for EOQ Items by Missing Data Group

Variable	Skew statistic	Skew standard error	Ratio	Skew statistic	Skew standard error	Ratio	Skew statistic	Skew standard error	Ratio
	No Missing Data N = 788			Data Missing for 1 Variable N = 157-166			Data Missing for 2-6 Variables N = 46-54		
Competence 1	0.27	0.09	3.1	-0.06	0.19	-0.3	-0.63	0.35	-1.8
Competence 2	-0.24	0.09	-2.7	-0.24	0.19	-1.3	0.38	0.35	1.1
Competence 3	-0.15	0.09	-1.7	-0.13	0.19	-0.7	-0.31	0.34	-0.9
Competence	-0.34	0.09	-3.9	-0.08	0.19	-0.4	0.13	0.34	0.4
Learning 1	-0.81	0.09	-9.3	-1.08	0.19	-5.6	-0.61	0.34	-1.8
Learning 2	-0.87	0.09	-10.0	-0.89	0.19	-4.6	-0.85	0.34	-2.5
Learning 3	-0.83	0.09	-9.5	-0.98	0.19	-5.0	-0.23	0.34	-0.7
Learning 4	-0.85	0.09	-9.8	-1.17	0.19	-6.1	-0.59	0.34	-1.7
Learning 5	-0.78	0.09	-9.0	-0.65	0.19	-3.4	0.28	0.34	0.8
Risk 1	-0.50	0.09	-5.7	-0.65	0.19	-3.4	0.28	0.33	0.8
Risk 2	-1.15	0.09	-13.2	-1.18	0.19	-6.2	-0.49	0.33	-1.5
Risk 3	-0.67	0.09	-7.7	-0.82	0.19	-4.3	-0.32	0.33	-1.0
Risk 4	-0.50	0.09	-5.8	-0.71	0.19	-3.7	-0.82	0.33	-2.5
Strain 1	-0.37	0.09	-4.3	-0.15	0.19	-0.8	-0.16	0.33	-0.5
Strain 2	-0.03	0.09	-0.4	-0.02	0.19	-0.1	-0.23	0.33	-0.7
Strain 3	0.12	0.09	1.3	0.22	0.19	1.2	-0.01	0.33	0.0
Strain 4	0.24	0.09	2.7	0.24	0.19	1.3	-0.03	0.33	-0.1
Anticipation 1	-0.06	0.09	-0.6	0.15	0.19	0.8	0.15	0.33	0.4
Anticipation 2	-0.35	0.09	-4.0	-0.50	0.19	-2.6	-0.46	0.33	-1.4
Anticipation 3	-0.10	0.09	-1.1	0.24	0.19	1.3	-0.30	0.33	-0.9
Anticipation 4	-0.54	0.09	-6.3	-0.66	0.19	-3.4	-0.24	0.33	-0.7
Anticipation 5	-0.31	0.09	-3.6	-0.09	0.19	-0.5	-0.75	0.33	-2.3
Covering 1	0.59	0.09	6.7	0.44	0.19	2.3	-0.01	0.33	0.0
Covering 2	0.66	0.09	7.6	0.61	0.19	3.2	-0.77	0.33	-2.3
Covering 3	0.65	0.09	7.5	0.61	0.19	3.2	-0.44	0.33	-1.3
Covering 4	0.64	0.09	7.4	0.75	0.19	4.0	0.08	0.33	0.2
Covering 5	0.28	0.09	3.2	0.39	0.19	2.0	-0.41	0.33	-1.3
Communication 1	-0.69	0.09	-7.9	-0.86	0.19	-4.5	-0.90	0.33	-2.8
Communication 2	-0.53	0.09	-6.0	-0.72	0.19	-3.7	0.11	0.33	0.3
Communication 3	-0.33	0.09	-3.8	-0.35	0.19	-1.8	-0.48	0.33	-1.5
Communication 4	-0.72	0.09	-8.3	-0.88	0.19	-4.6	-0.47	0.33	-1.4
Thinking 1	-0.55	0.09	-6.3	-0.47	0.19	-2.5	-0.26	0.33	-0.8
Thinking 2	-0.54	0.09	-6.2	-0.59	0.19	-3.1	0.24	0.33	0.7
Thinking 3	-0.41	0.09	-4.7	-0.22	0.19	-1.1	-0.02	0.33	-0.1
Thinking 4	-0.50	0.09	-5.7	-0.45	0.19	-2.3	-0.22	0.33	-0.7
Motivation 1	-0.63	0.09	-7.2	-0.58	0.19	-3.0	-0.28	0.33	-0.9
Motivation 2	0.30	0.09	3.4	0.41	0.19	2.2	-0.12	0.33	-0.4
Motivation 3	-0.66	0.09	-7.6	-0.51	0.19	-2.6	0.17	0.33	0.5
Motivation 4	-0.40	0.09	-4.6	-0.39	0.19	-2.1	-0.46	0.33	-1.4
Motivation 5	-0.72	0.09	-8.2	-0.48	0.19	-2.5	-0.14	0.33	-0.4

Table 56. Kurtosis Statistics and Standard Errors for EOQ Items by Missing Data Group

Variable	Kurtosis statistic	Kurtosis standard error	Ratio	Kurtosis statistic	Kurtosis standard error	Ratio	Kurtosis statistic	Kurtosis standard error	Ratio
	No Missing Data N = 788			Data Missing for 1 Variable N = 157-166			Data Missing for 2-6 Variables N = 46-54		
Competence 1	-0.09	0.17	-0.5	-0.72	0.38	-1.9	0.64	0.69	0.9
Competence 2	-0.40	0.17	-2.3	-0.40	0.38	-1.0	-0.22	0.69	-0.3
Competence 3	-0.30	0.17	-1.7	-0.30	0.38	-0.8	-1.06	0.67	-1.6
Competence	-0.15	0.17	-0.9	-0.92	0.38	-2.4	-0.47	0.67	-0.7
Learning 1	0.30	0.17	1.7	1.68	0.39	4.4	0.68	0.66	1.0
Learning 2	0.60	0.17	3.4	0.43	0.38	1.1	-0.01	0.66	0.0
Learning 3	0.41	0.17	2.4	1.09	0.39	2.8	-0.58	0.66	-0.9
Learning 4	0.79	0.17	4.5	1.96	0.38	5.2	0.37	0.66	0.6
Learning 5	0.41	0.17	2.3	0.27	0.38	0.7	-0.79	0.66	-1.2
Risk 1	-0.32	0.17	-1.8	0.23	0.38	0.6	-0.48	0.66	-0.7
Risk 2	1.13	0.17	6.5	1.76	0.38	4.7	0.14	0.66	0.2
Risk 3	-0.29	0.17	-1.7	0.28	0.38	0.7	-0.83	0.66	-1.3
Risk 4	-0.81	0.17	-4.6	-0.47	0.38	-1.2	1.08	0.66	1.6
Strain 1	-0.67	0.17	-3.8	-0.97	0.38	-2.6	-0.58	0.66	-0.9
Strain 2	-1.02	0.17	-5.8	-0.88	0.38	-2.3	-0.76	0.66	-1.2
Strain 3	-0.81	0.17	-4.6	-0.86	0.38	-2.3	-0.44	0.66	-0.7
Strain 4	-0.91	0.17	-5.2	-0.85	0.38	-2.3	-0.79	0.66	-1.2
Anticipation 1	-0.44	0.17	-2.5	-0.23	0.38	-0.6	-0.68	0.66	-1.0
Anticipation 2	-0.44	0.17	-2.6	-0.19	0.38	-0.5	-0.28	0.66	-0.4
Anticipation 3	-0.32	0.17	-1.9	-0.35	0.38	-0.9	-0.81	0.66	-1.2
Anticipation 4	-0.06	0.17	-0.4	0.28	0.38	0.7	-0.76	0.65	-1.2
Anticipation 5	-0.90	0.17	-5.2	-1.24	0.38	-3.3	-0.01	0.65	0.0
Covering 1	-0.48	0.17	-2.7	-0.71	0.38	-1.9	-0.55	0.65	-0.8
Covering 2	-0.49	0.17	-2.8	-0.93	0.38	-2.5	0.65	0.65	1.0
Covering 3	-0.51	0.17	-3.0	-0.51	0.38	-1.3	-0.45	0.65	-0.7
Covering 4	-0.70	0.17	-4.0	-0.41	0.38	-1.1	-1.04	0.65	-1.6
Covering 5	-0.70	0.17	-4.0	-0.62	0.38	-1.6	0.29	0.64	0.5
Communication 1	0.33	0.17	1.9	0.36	0.38	0.9	1.22	0.64	1.9
Communication 2	-0.52	0.17	-3.0	-0.20	0.39	-0.5	-0.69	0.64	-1.1
Communication 3	-0.58	0.17	-3.3	-0.60	0.38	-1.6	-0.19	0.64	-0.3
Communication 4	0.29	0.17	1.6	0.98	0.38	2.6	0.28	0.64	0.4
Thinking 1	-0.04	0.17	-0.2	-0.24	0.38	-0.6	-0.78	0.64	-1.2
Thinking 2	0.21	0.17	1.2	0.50	0.38	1.3	-0.54	0.64	-0.8
Thinking 3	-0.26	0.17	-1.5	-0.76	0.38	-2.0	-1.02	0.64	-1.6
Thinking 4	-0.16	0.17	-0.9	-0.53	0.38	-1.4	-0.61	0.64	-0.9
Motivation 1	0.02	0.17	0.1	-0.33	0.38	-0.9	-0.15	0.64	-0.2
Motivation 2	-0.91	0.17	-5.2	-0.67	0.38	-1.8	-0.59	0.64	-0.9
Motivation 3	-0.62	0.17	-3.5	-0.86	0.38	-2.3	-0.38	0.64	-0.6
Motivation 4	-0.45	0.17	-2.6	-0.57	0.38	-1.5	-0.48	0.64	-0.8
Motivation 5	-0.73	0.17	-4.2	-1.13	0.38	-3.0	-0.27	0.64	-0.4

Appendix F.

Item Correlation Tables

Table 57 Pearson Correlations Between AGQ Item Scores

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Task Approach 1	1																	
Task Approach 2	.667	1																
Task Approach 3	.733	.679	1															
Task Avoidance 1	.689	.551	.604	1														
Task Avoidance 2	.549	.497	.613	.555	1													
Task Avoidance 3	.503	.427	.445	.509	.476	1												
Self Approach 1	.454	.405	.458	.421	.371	.483	1											
Self Approach 2	.433	.413	.498	.393	.426	.359	.524	1										
Self Approach 3	.427	.463	.472	.398	.359	.404	.718	.462	1									
Self Avoidance 1	.328	.379	.409	.398	.457	.427	.530	.480	.532	1								
Self Avoidance 2	.416	.384	.414	.482	.443	.514	.584	.428	.506	.548	1							
Self Avoidance 3	.431	.359	.414	.533	.463	.497	.639	.506	.563	.656	.665	1						
Other Approach 1	.399	.333	.349	.306	.289	.246	.222	.245	.227	.204	.228	.220	1					
Other Approach 2	.433	.418	.456	.384	.350	.349	.292	.318	.291	.273	.320	.327	.624	1				
Other Approach 3	.404	.345	.361	.344	.304	.276	.268	.259	.255	.211	.274	.290	.802	.712	1			
Other Avoidance 1	.373	.303	.330	.392	.378	.376	.285	.287	.280	.399	.407	.422	.596	.628	.662	1		
Other Avoidance 2	.361	.331	.377	.346	.501	.349	.270	.293	.260	.370	.369	.378	.502	.623	.564	.645	1	
Other Avoidance 3	.380	.356	.359	.399	.413	.406	.320	.333	.318	.399	.426	.444	.523	.740	.614	.736	.675	1

Table 58. Bivariate Correlations Between EOQ Items (1-40 with 1-8)

Item	1	2	3	4	5	6	7	8
1. Competence 1	---							
2. Competence 2	.314**	---						
3. Competence 3	.415**	.361**	---					
4. Competence	.131**	.311**	.246**	---				
5. Learning	.098**	.319**	.225**	.397**	---			
6. Learning 2	.026	.295**	.151**	.341**	.596**	---		
7. Learning 3	.127**	.370**	.285**	.368**	.708**	.541**	---	
8. Learning 4	.074*	.238**	.222**	.285**	.523**	.359**	.578**	---
9. Learning 5	.045	.301**	.213**	.362**	.677**	.508**	.684**	.576**
10. Risk 1	.030	.190**	.092**	.256**	.336**	.306**	.367**	.369**
11. Risk 2	.036	.230**	.149**	.319**	.443**	.409**	.459**	.323**
12. Risk 3	-.004	.101**	.109**	.265**	.375**	.290**	.351**	.395**
13. Risk 4	-.117**	.129**	.044	.133**	.235**	.235**	.214**	.133**
14. Strain 1	-.039	.004	-.110**	-.030	-.129**	-.044	-.128**	-.068*
15. Strain 2	.041	-.002	-.104**	-.039	-.157**	-.052	-.133**	-.086**
16. Strain 3	.012	-.054	-.069*	-.101**	-.165**	-.128**	-.159**	-.040
17. Strain 4	.012	-.078*	-.091**	-.106**	-.189**	-.150**	-.203**	-.098**
18. Anticipation 1	-.175**	-.059	-.201**	-.010	-.033	.016	-.079*	-.017
19. Anticipation 2	-.003	.173**	.094**	.178**	.237**	.219**	.234**	.187**
20. Anticipation 3	-.047	.024	-.041	-.015	.031	.040	.045	.108**
21. Anticipation 4	.017	.247**	.049	.162**	.302**	.372**	.263**	.231**
22. Anticipation 5	-.134**	-.043	-.117**	.031	.058	.114**	.058	.261**
23. Covering 1	.167**	-.046	.066*	-.083**	-.210**	-.216**	-.222**	-.157**
24. Covering 2	.126**	-.028	.007	-.089**	-.211**	-.212**	-.217**	-.154**
25. Covering 3	.048	-.093**	.028	-.138**	-.193**	-.144**	-.187**	-.110**
26. Covering 4	.190**	-.055	.143**	-.052	-.235**	-.216**	-.256**	-.165**
27. Covering 5	.053	-.121**	.002	-.118**	-.242**	-.196**	-.221**	-.125**
28. Communication 1	-.050	-.205**	-.083**	-.201**	-.256**	-.255**	-.271**	-.266**
29. Communication 2	.017	.197**	.051	.157**	.303**	.262**	.266**	.200**
30. Communication 3	-.064*	.005	-.039	.019	.165**	.124**	.094**	.059
31. Communication 4	.107**	.350**	.280**	.367**	.436**	.397**	.401**	.305**
32. Thinking 1	.199**	.352**	.279**	.350**	.361**	.291**	.341**	.269**
33. Thinking 2	.132**	.426**	.297**	.407**	.438**	.393**	.414**	.316**
34. Thinking 3	.207**	.441**	.315**	.453**	.373**	.328**	.360**	.280**
35. Thinking 4	.209**	.408**	.298**	.339**	.326**	.321**	.342**	.256**
36. Motivation 1	.199**	.437**	.270**	.353**	.394**	.362**	.399**	.303**
37. Motivation 2	.106**	.070*	.189**	.080*	.140**	.034	.162**	.107**
38. Motivation 3	-.116**	.136**	.082**	.188**	.308**	.290**	.286**	.169**
39. Motivation 4	.181**	.256**	.190**	.214**	.294**	.260**	.340**	.228**
40. Motivation 5	-.055	.143**	.098**	.172**	.228**	.238**	.236**	.169**

Note: Risk 4, Motivation 2, Motivation 3, and Motivation 5 have been reverse coded.
*Correlation is significant to the 0.05 level.
**Correlation is significant to the 0.01 level.

Table 59. Bivariate Correlations Between EOQ Items (9-40 with 9-16)

Item	9	10	11	12	13	14	15	16
9. Learning 5	---							
10. Risk 1	.348**	---						
11. Risk 2	.393**	.294**	---					
12. Risk 3	.353**	.374**	.359**	---				
13. Risk 4	.176**	.073*	.282**	.111**	---			
14. Strain 1	-.067*	-.037	-.060	-.117**	-.245**	---		
15. Strain 2	-.075*	-.094**	-.128**	-.181**	-.253**	.508**	---	
16. Strain 3	-.090**	-.039	-.142**	-.132**	-.255**	.348**	.401**	---
17. Strain 4	-.126**	-.036	-.168**	-.120**	-.317**	.383**	.454**	.603**
18. Anticipation 1	.018	.168**	-.025	.044	-.196**	.308**	.286**	.240**
19. Anticipation 2	.328**	.220**	.208**	.146**	-.038	.075*	.074*	.109**
20. Anticipation 3	.120**	.145**	.017	.104**	-.139**	.183**	.154**	.214**
21. Anticipation 4	.300**	.218**	.218**	.187**	.026	.055	.050	.028
22. Anticipation 5	.110**	.183**	.083**	.153**	-.055	.192**	.124**	.161**
23. Covering 1	-.203**	-.100**	-.216**	-.139**	-.246**	.160**	.213**	.273**
24. Covering 2	-.174**	-.070*	-.240**	-.152**	-.373**	.293**	.348**	.332**
25. Covering 3	-.168**	-.057	-.170**	-.127**	-.184**	.112**	.126**	.204**
26. Covering 4	-.208**	-.107**	-.183**	-.141**	-.346**	.165**	.239**	.243**
27. Covering 5	-.184**	-.114**	-.143**	-.164**	-.294**	.286**	.319**	.388**
28. Communication 1	-.298**	-.235**	-.199**	-.172**	-.088**	-.017	.002	.051
29. Communication 2	.255**	.123**	.142**	.140**	.036	.005	-.022	-.085**
30. Communication 3	.104**	.032	.007	.089**	-.112**	.086**	.041	.002
31. Communication 4	.404**	.194**	.377**	.154**	.288**	.037	-.070*	-.072*
32. Thinking 1	.376**	.186**	.281**	.119**	.085**	.023	.013	-.061
33. Thinking 2	.406**	.229**	.463**	.184**	.193**	-.013	-.059	-.096**
34. Thinking 3	.361**	.192**	.333**	.136**	.118**	.005	-.042	-.071*
35. Thinking 4	.337**	.177**	.342**	.112**	.204**	-.005	-.067*	-.082**
36. Motivation 1	.406**	.240**	.408**	.175**	.208**	-.051	-.109**	-.155**
37. Motivation 2	.116**	.050	.100**	.127**	.209**	-.555**	-.482**	-.375**
38. Motivation 3	.229**	.079*	.318**	.129**	.695**	-.217**	-.272**	-.257**
39. Motivation 4	.248**	.165**	.184**	.147**	.100**	.062	-.009	-.030
40. Motivation 5	.180**	.103**	.269**	.153**	.404**	-.306**	-.307**	-.271**

Note: Risk4, Motivation2, Motivation3, and Motivation5 have been reverse coded.
*Correlation is significant to the 0.05 level.
**Correlation is significant to the 0.01 level.

Table 60. Correlations Between EOQ Items (17-40 with 17-24)

Item	17	18	19	20	21	22	23	24
17. Strain 4	---							
18. Anticipation 1	.262**	---						
19. Anticipation 2	.078*	.271**	---					
20. Anticipation 3	.167**	.394**	.257**	---				
21. Anticipation 4	-.029	.175**	.372**	.230**	---			
22. Anticipation 5	.101**	.396**	.277**	.312**	.226**	---		
23. Covering 1	.351**	.095**	.024	.067*	-.020	-.010	---	
24. Covering 2	.487**	.144**	.066*	.066*	-.055	.018	.490**	---
25. Covering 3	.173**	.151**	.073*	.157**	.015	.042	.229**	.213**
26. Covering 4	.349**	.187**	.039	.122**	-.032	.040	.500**	.483**
27. Covering 5	.522**	.173**	.020	.079*	-.026	.083**	.468**	.437**
28. Communication 1	.128**	-.039	-.162**	-.079*	-.134**	-.118**	.310**	.201**
29. Communication 2	-.138**	.011	.125**	.057	.194**	.074*	-.184**	-.146**
30. Communication 3	.015	.180**	.147**	.106**	.140**	.151**	-.070*	-.028
31. Communication 4	-.159**	-.034	.177**	.018	.161**	.046	-.132**	-.173**
32. Thinking 1	-.107**	-.025	.215**	.028	.204**	-.001	-.024	-.069*
33. Thinking 2	-.124**	-.067*	.228**	-.029	.213**	-.025	-.110**	-.095**
34. Thinking 3	-.120**	-.087**	.170**	.010	.157**	-.064*	-.072*	-.062*
35. Thinking 4	-.062	-.100**	.125**	-.025	.136**	-.054	-.064*	-.081*
36. Motivation 1	-.167**	-.129**	.154**	-.053	.160**	-.050	-.123**	-.118**
37. Motivation 2	-.407**	-.346**	-.082**	-.181**	-.021	-.176**	-.178**	-.227**
38. Motivation 3	-.329**	-.204**	.014	-.116**	.076*	-.037	-.272**	-.375**
39. Motivation 4	-.083**	.019	.155**	.075*	.180**	.021	.009	-.054
40. Motivation 5	-.345**	-.273**	-.013	-.194**	.031	-.048	-.264**	-.350**

Note: Risk4, Motivation2, Motivation3, and Motivation5 have been reverse coded.

*Correlation is significant to the 0.05 level.

**Correlation is significant to the 0.01 level.

Table 61. Correlations Between EOQ Items (25-40 with 25-32)

Item	25	26	27	28	29	30	31	32
25. Covering 3	---							
26. Covering 4	.294**	---						
27. Covering 5	.241**	.442**	---					
28. Communication 1	.120**	.209**	.264**	---				
29. Communication 2	-.091**	-.329**	-.271**	-.272**	---			
30. Communication 3	-.032	-.170**	-.113**	-.164**	.516**	---		
31. Communication 4	-.152**	-.151**	-.140**	-.227**	.196**	.011	---	
32. Thinking 1	-.106**	-.043	-.119**	-.227**	.173**	.039	.366**	---
33. Thinking 2	-.136**	-.119**	-.145**	-.229**	.162**	.013	.491**	.500**
34. Thinking 3	-.153**	-.087**	-.144**	-.189**	.177**	.036	.443**	.490**
35. Thinking 4	-.151**	-.070*	-.157**	-.166**	.132**	-.021	.445**	.453**
36. Motivation 1	-.201**	-.130**	-.175**	-.195**	.158**	-.024	.426**	.407**
37. Motivation 2	-.121**	-.169**	-.277**	-.012	.012	-.107**	.034	.088**
38. Motivation 3	-.211**	-.348**	-.295**	-.087**	.089**	-.019	.311**	.136**
39. Motivation 4	-.040	-.028	-.038	-.171**	.110**	.035	.263**	.301**
40. Motivation 5	-.242**	-.277**	-.248**	-.040	.020	-.092**	.224**	.102**

Note: Risk4, Motivation2, Motivation3, and Motivation5 have been reverse coded.
*Correlation is significant to the 0.05 level.
**Correlation is significant to the 0.01 level.

Table 62. Correlations Between EOQ Items (33-40 with 33-40)

Item	33	34	35	36	37	38	39	40
33. Thinking 2	---							
34. Thinking 3	.622**	---						
35. Thinking 4	.585**	.607**	---					
36. Motivation 1	.568**	.544**	.611**	---				
37. Motivation 2	.119**	.111**	.097**	.193**	---			
38. Motivation 3	.278**	.177**	.218**	.239**	.212**	---		
39. Motivation 4	.221**	.225**	.203**	.175**	.023	.064*	---	
40. Motivation 5	.206**	.163**	.163**	.275**	.309**	.377**	.087**	---

Note: Risk4, Motivation2, Motivation3, and Motivation5 have been reverse coded.
*Correlation is significant to the 0.05 level.
**Correlation is significant to the 0.01 level.

Appendix G.

Sapling Learning Question Statistics

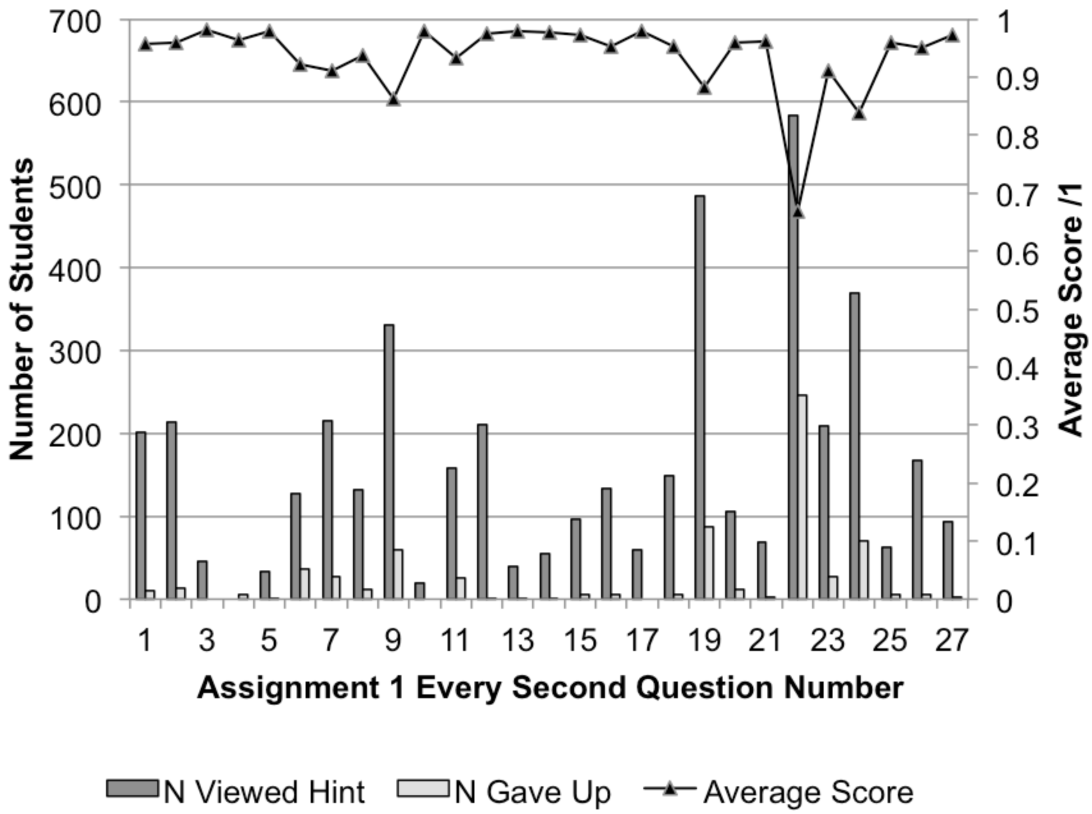

Figure 51. Number of students who viewed hint and gave up on questions from assignment 1 on the left-hand *y*-axis (bars). Average score for each question is shown as a line graph on the right-hand *y*-axis.

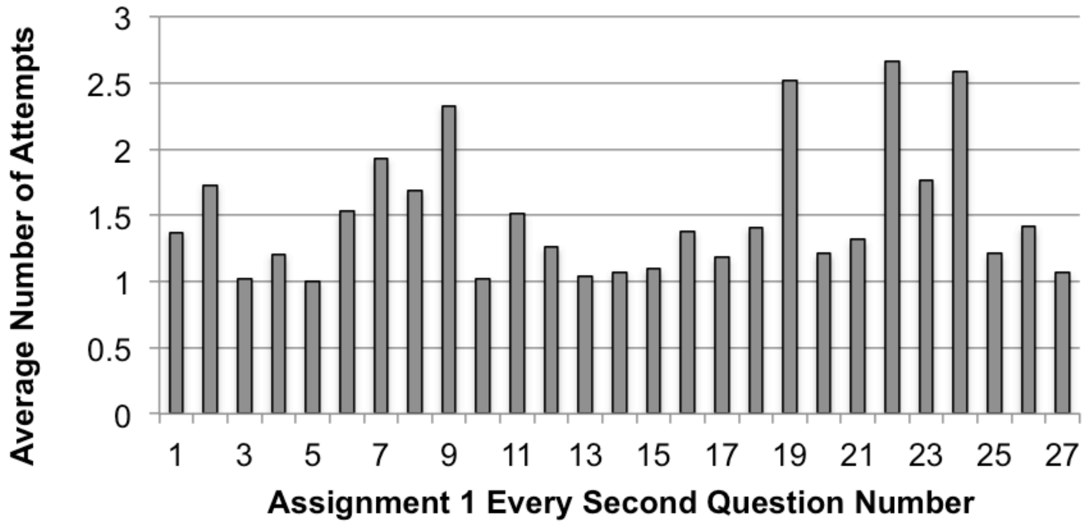

Figure 52. Average number of attempts for questions from assignment 1.

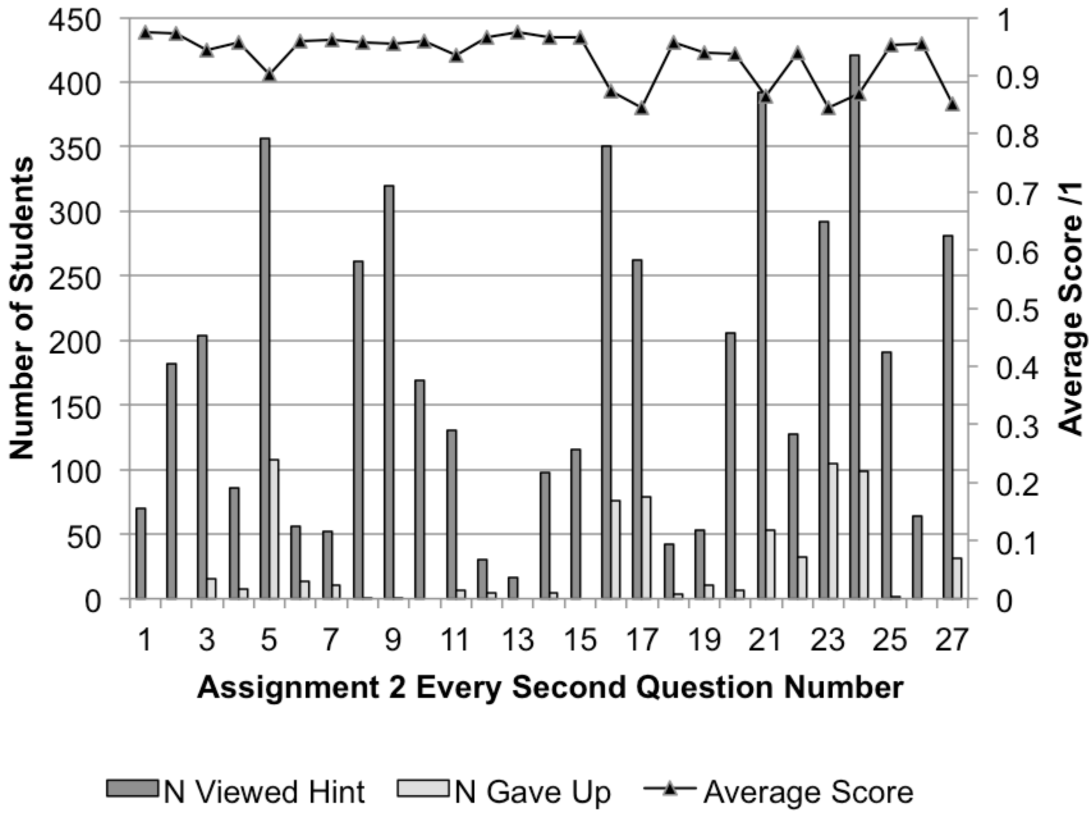

Figure 53. Number of students who viewed hint and gave up on questions from assignment 2 (left-hand *y*-axis) and average question score (right-hand *y*-axis).

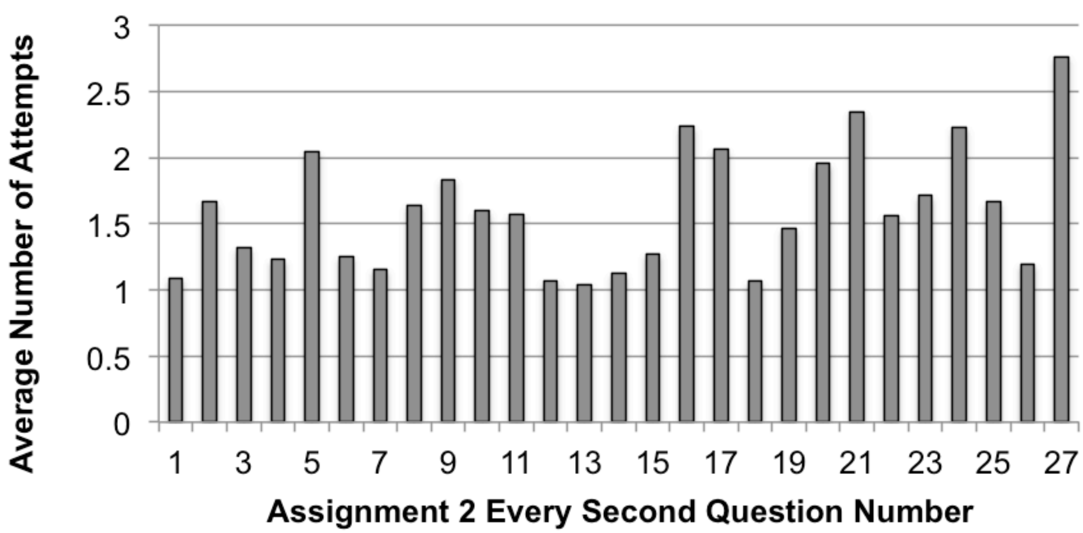

Figure 54. Average number of attempts for questions from assignment 2.

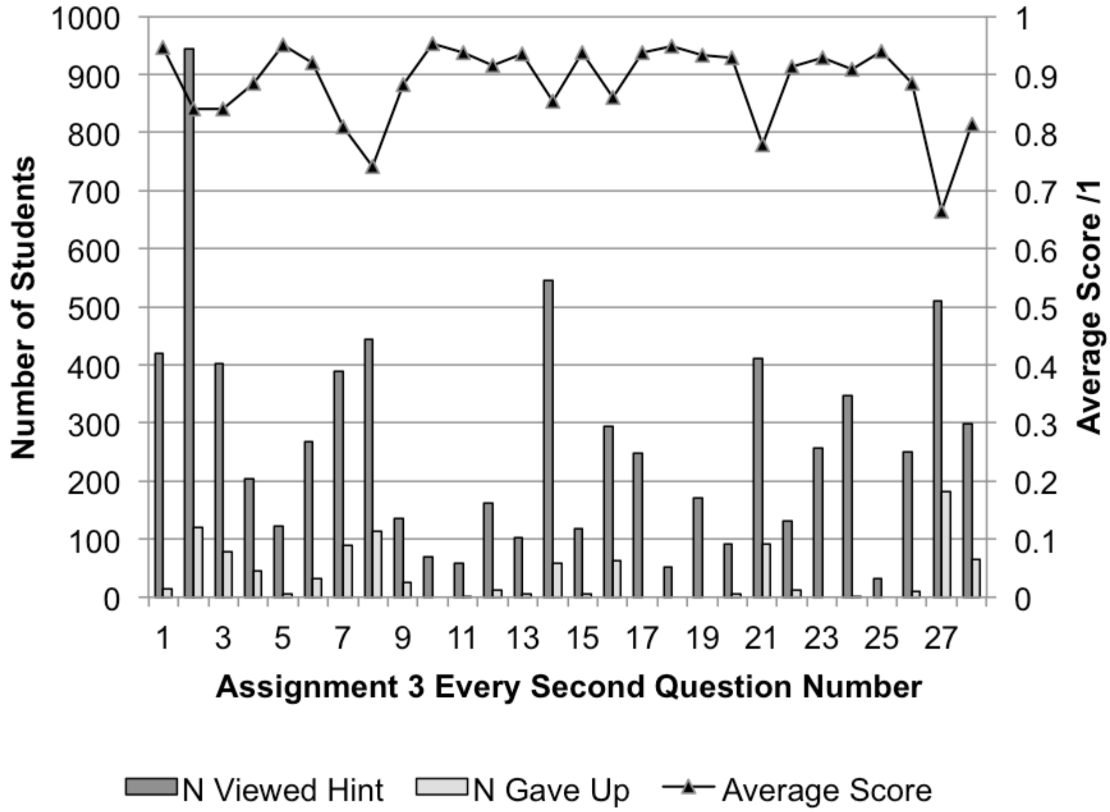

Figure 55. Number of students who viewed hint and gave up on questions from assignment 3 (left-hand *y*-axis) and average question score (right-hand *y*-axis).

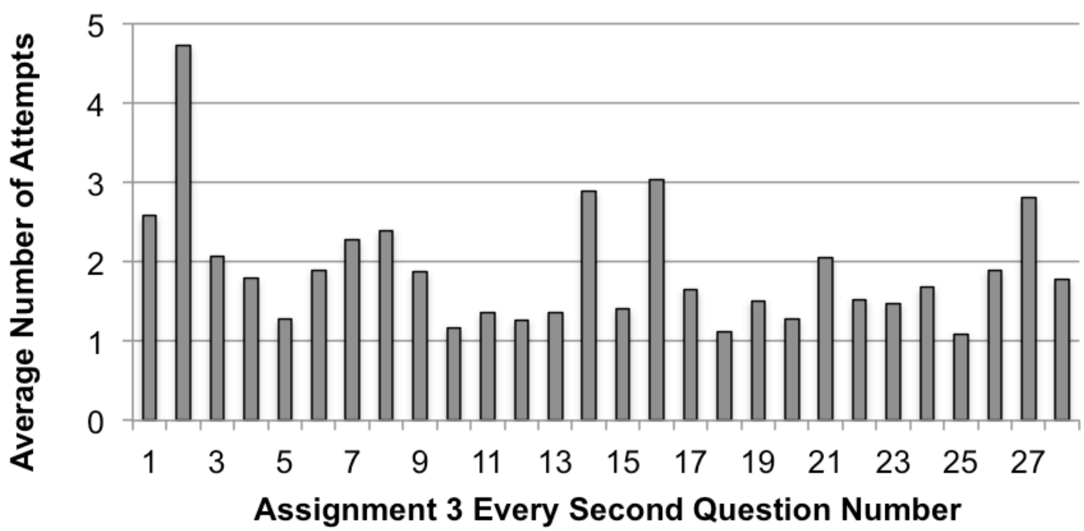

Figure 56. Average number of attempts for questions from assignment 3.

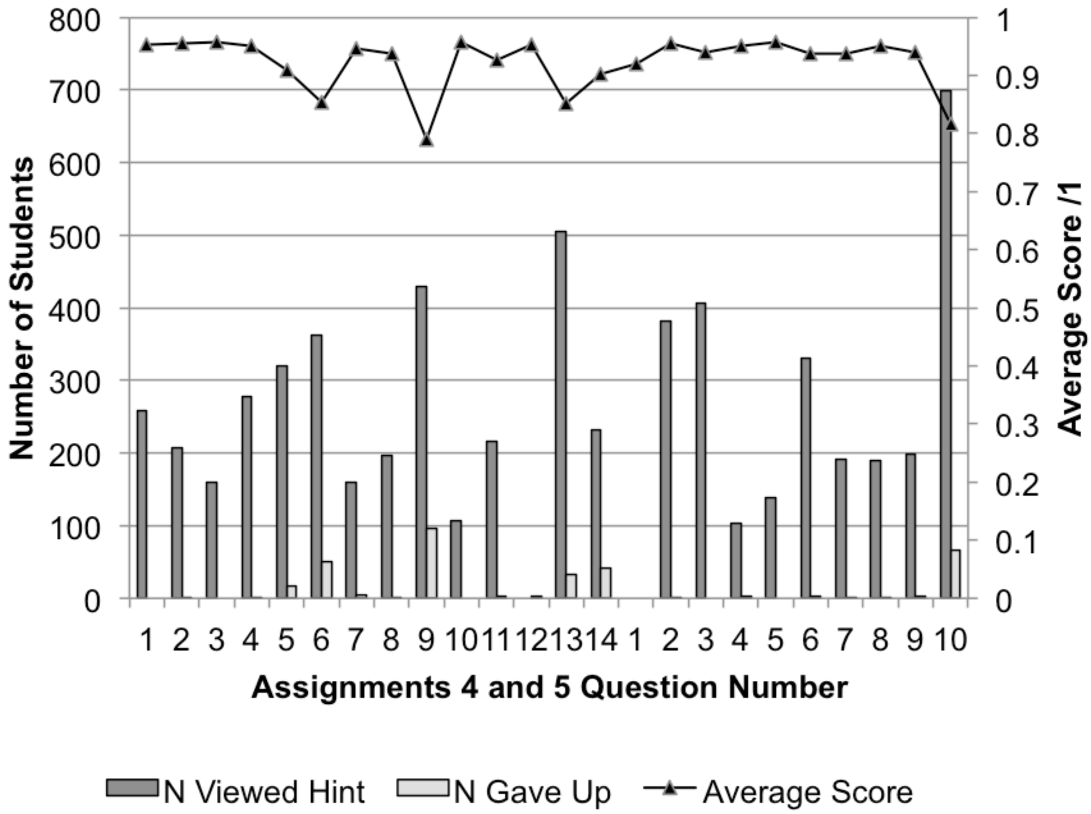

Figure 57. Number of students who viewed hint and gave up on questions from assignments 4 and 5 (left-hand *y*-axis) and average question score (right-hand *y*-axis).

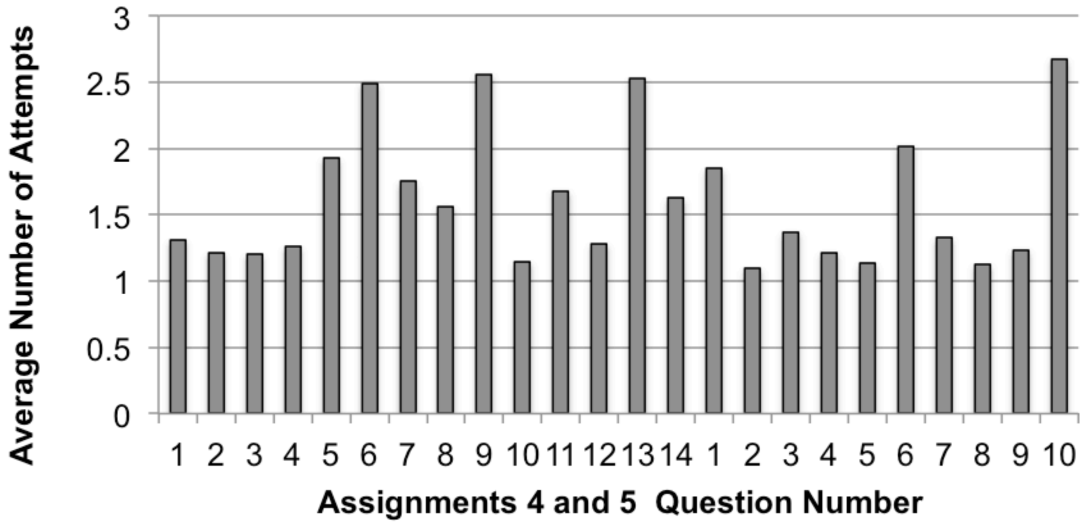

Figure 58. Average number of attempts for questions from assignments 4 and 5.

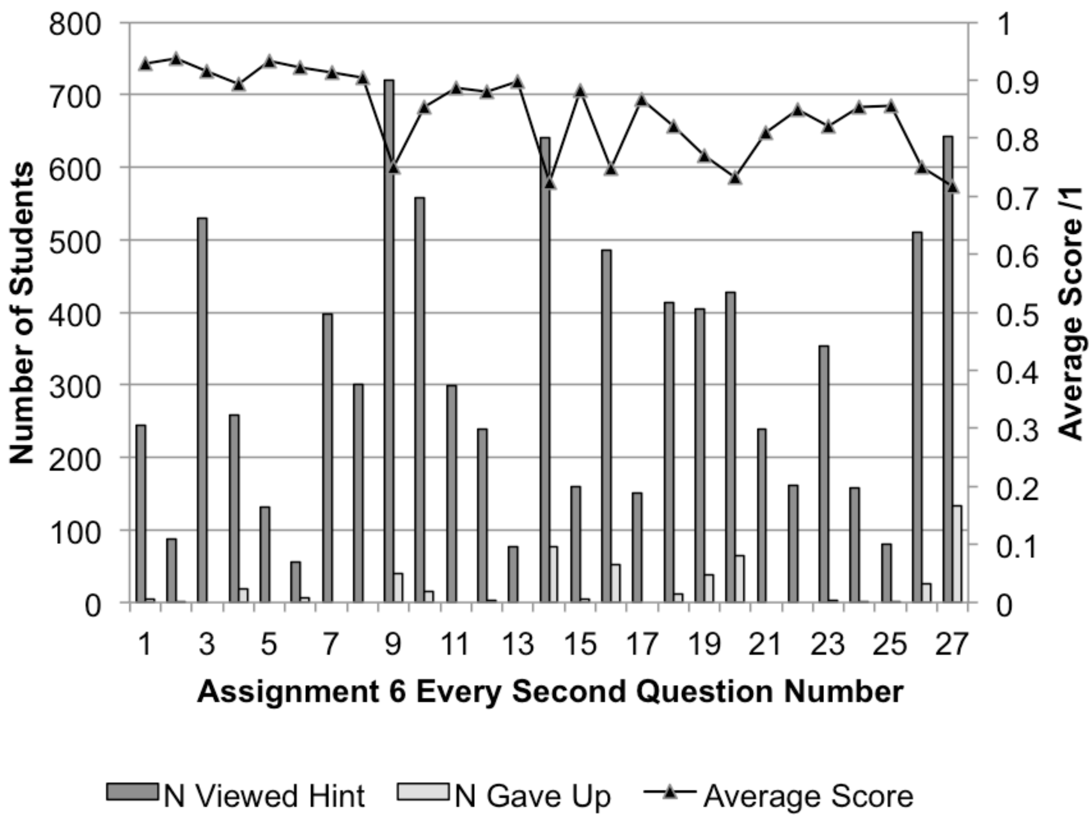

Figure 59. Number of students who viewed hint and gave up on questions from assignment 6 (left-hand *y*-axis) and average question score (right-hand *y-axis*).

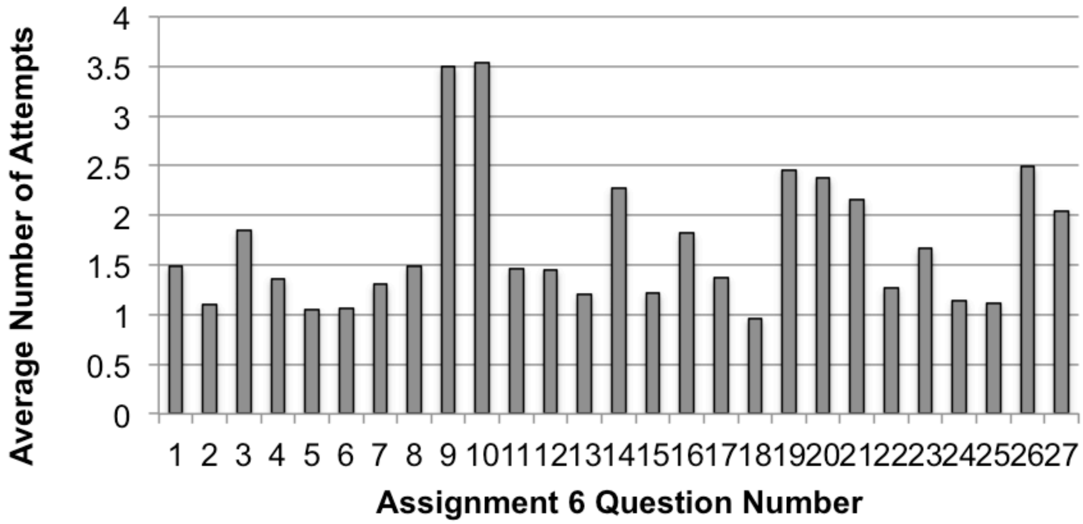

Figure 60. Average number of attempts for questions from assignment 6.

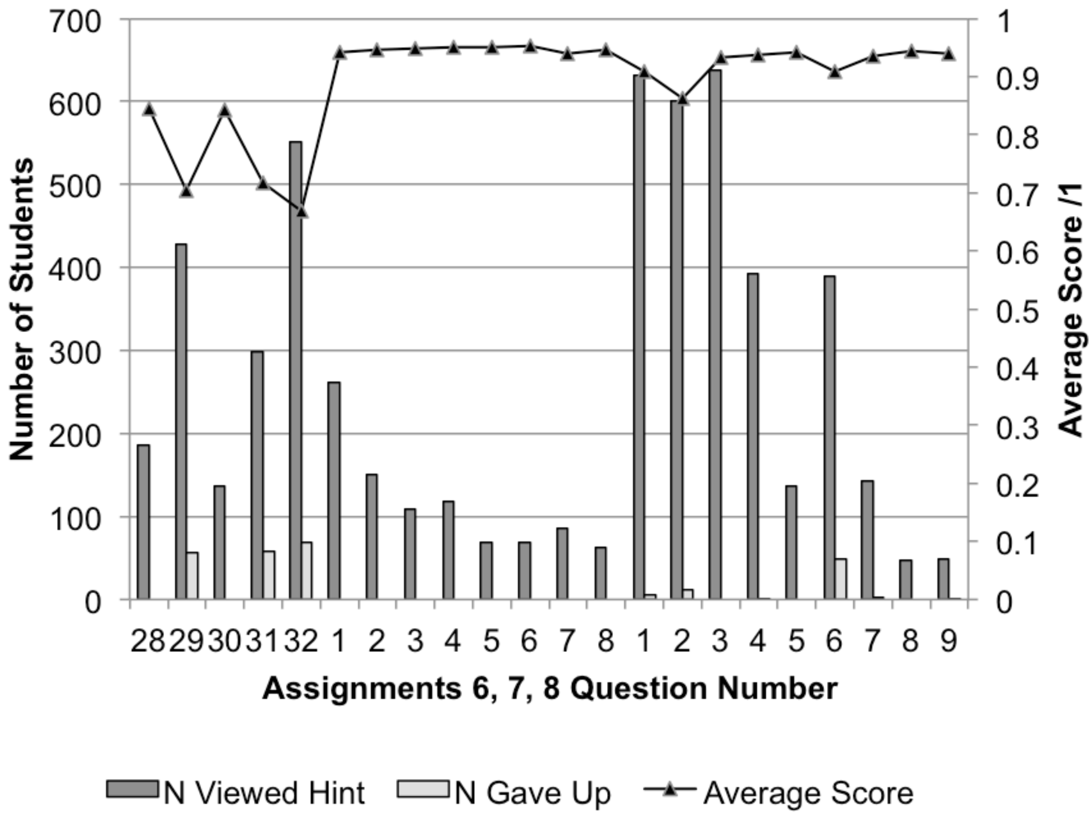

Figure 61. Number of students who viewed hint and gave up on questions from assignments 6, 7, and 8 (left-hand y-axis) and average question score (right-hand y-axis).

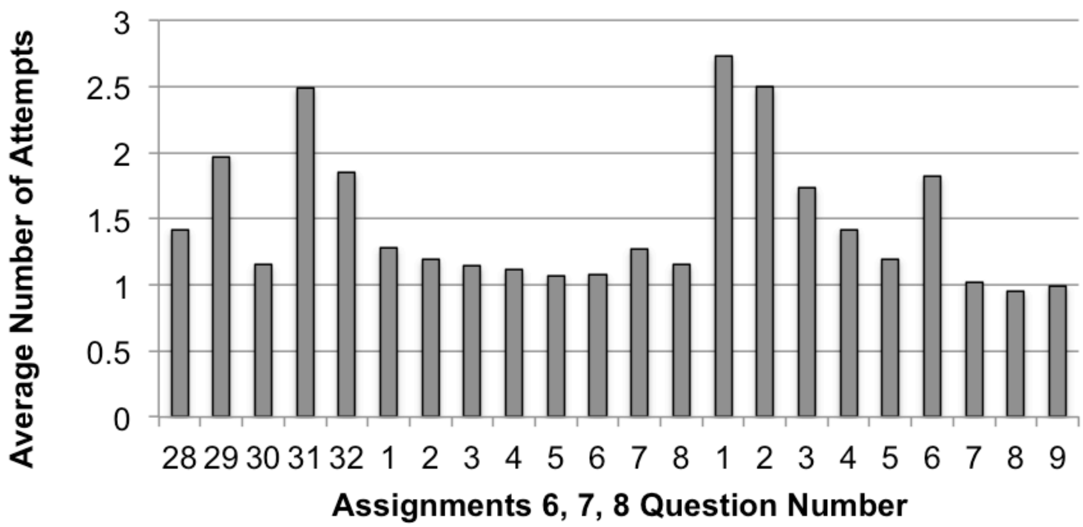

Figure 62. Average number of attempts for questions from assignments 6 7, and 8.

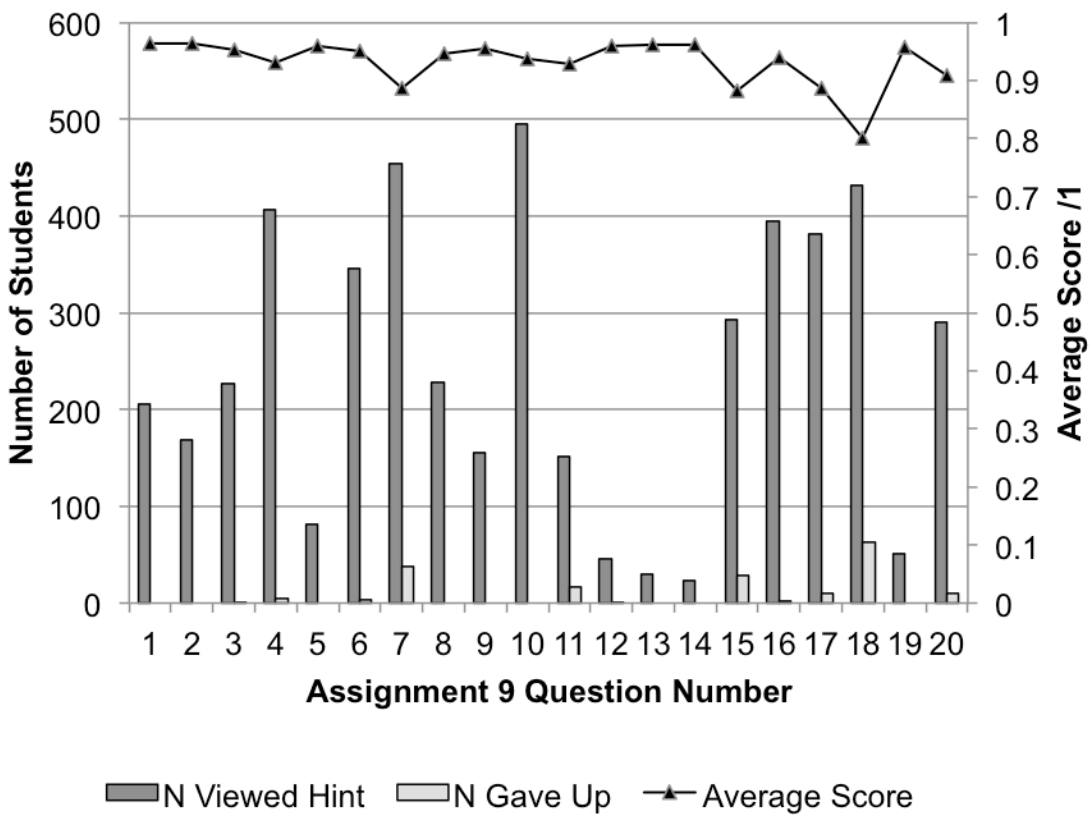

Figure 63. Number of students who viewed hint and gave up on questions from assignment 9 (left-hand *y*-axis) and average question score (right-hand *y*-axis).

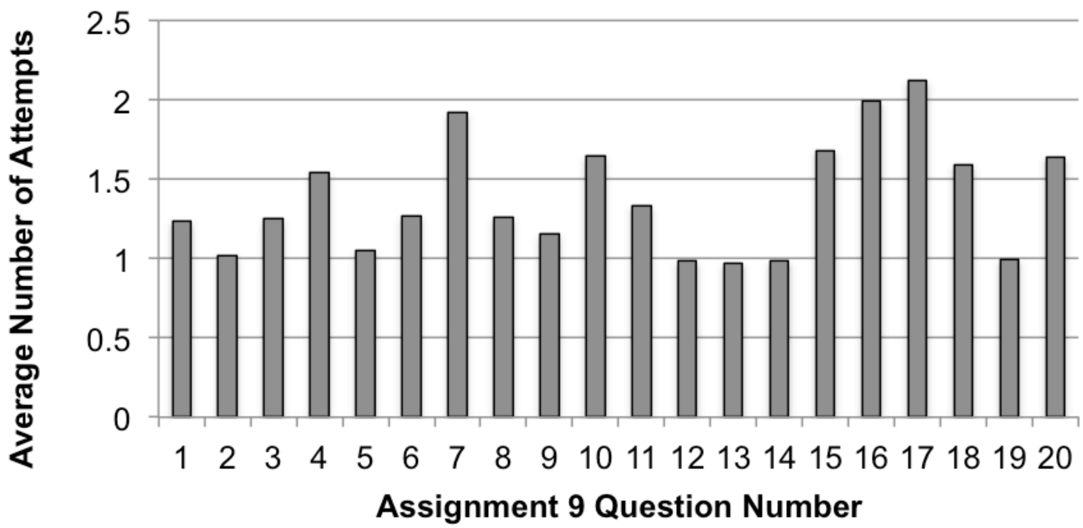

Figure 64. Average number of attempts for questions from assignment 9.

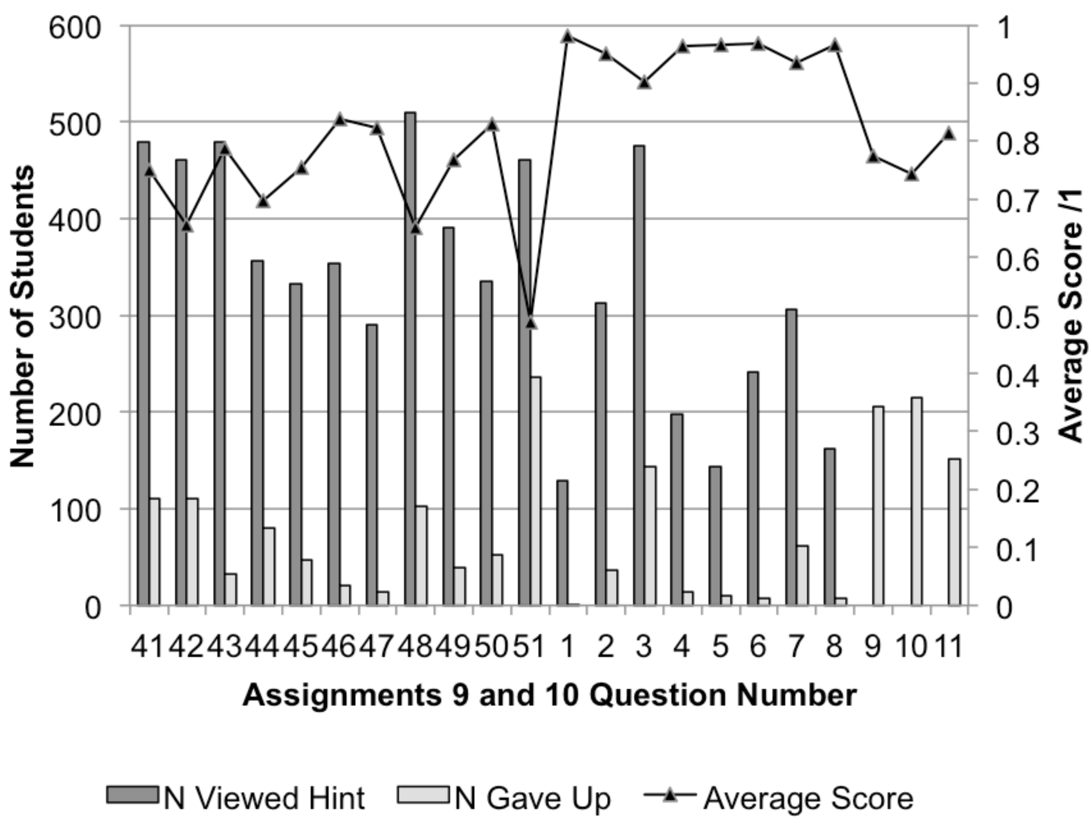

Figure 65. Number of students who viewed hint and gave up on questions from assignments 9 and 10 (left-hand *y*-axis) and average question score (right-hand *y*-axis).

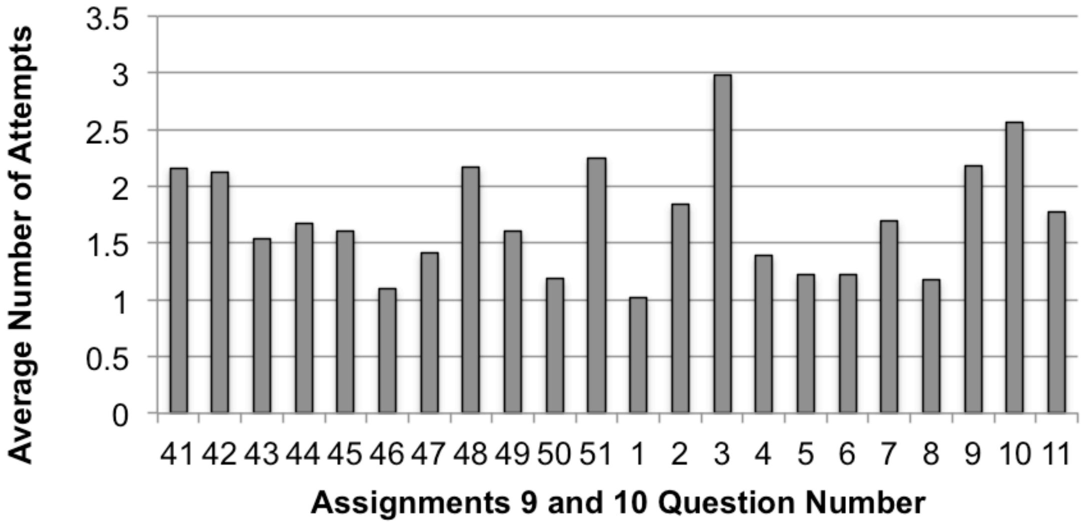

Figure 66. Average number of attempts for questions from assignments 9 and 10.

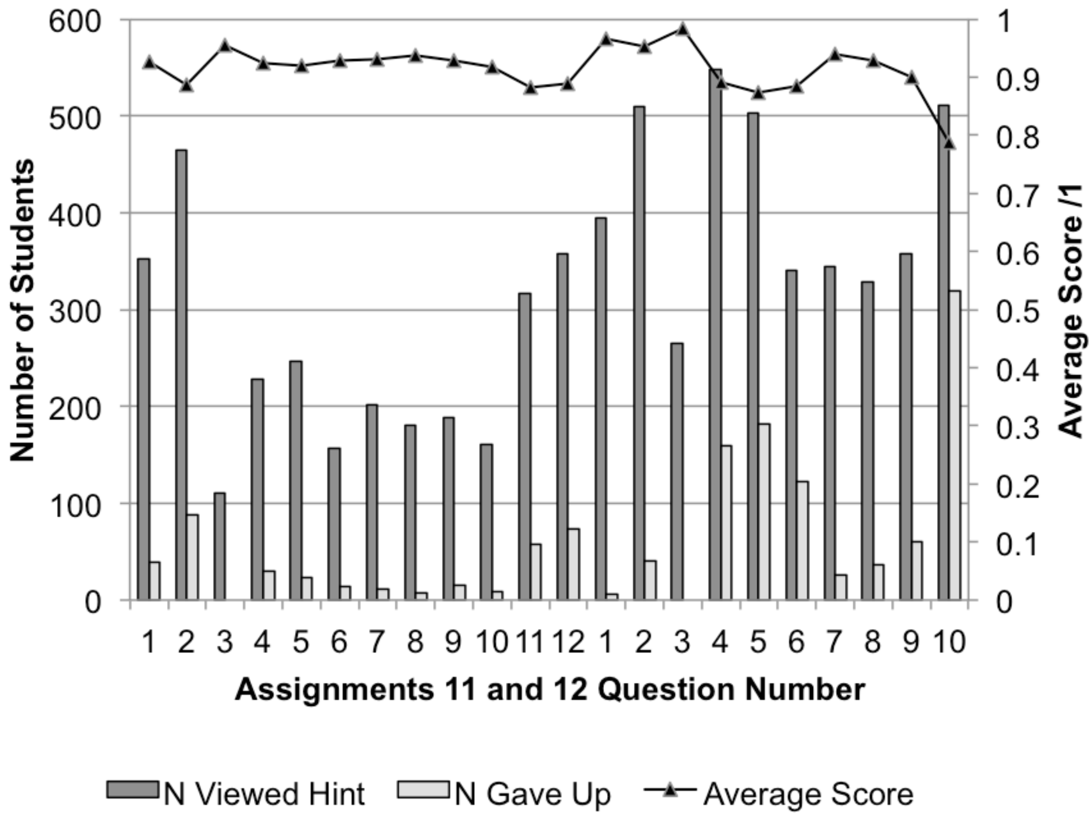

Figure 67. Number of students who viewed hint and gave up on questions from assignments 11 and 12 (left-hand *y*-axis) and average question score (right-hand y-axis).

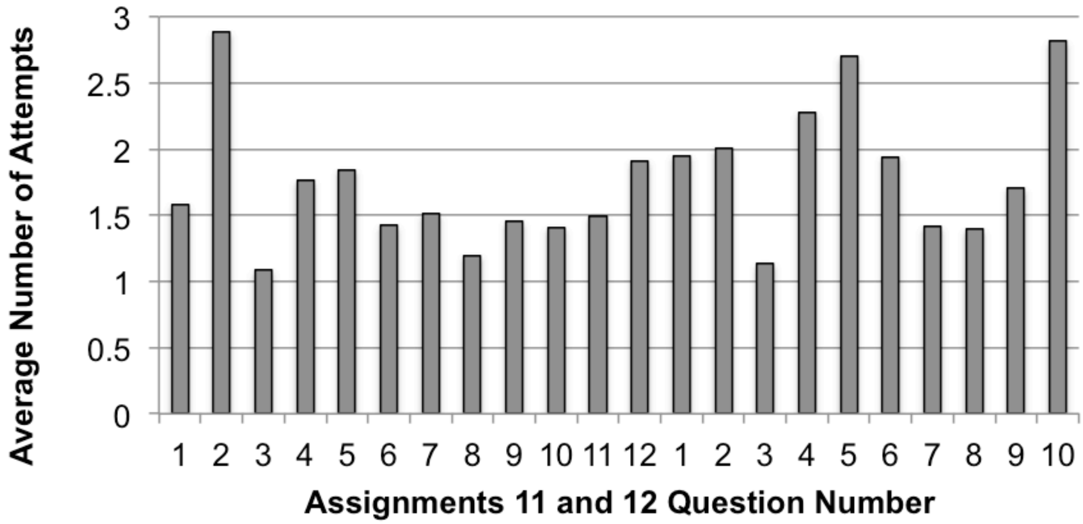

Figure 68. Average number of attempts for questions from assignments 11 and 12.

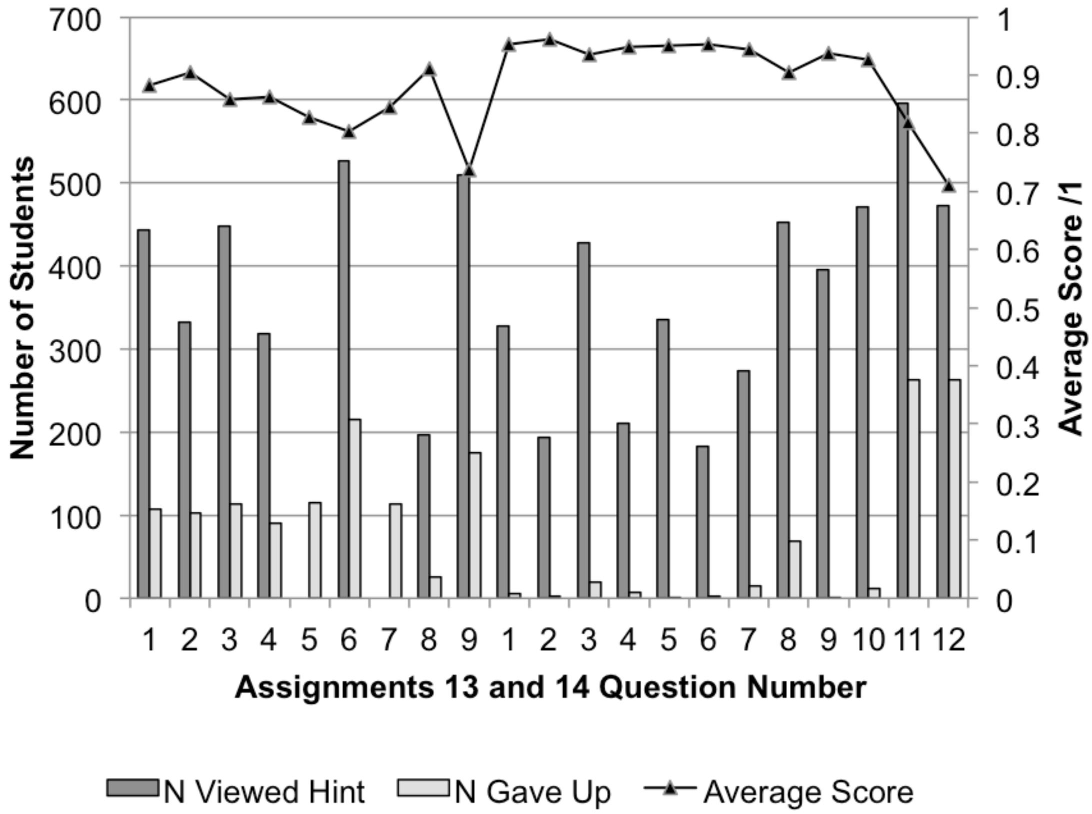

Figure 69. Number of students who viewed hint and gave up on questions from assignments 13 and 14 (left-hand *y*-axis) and average question score (right-hand *y*-axis).

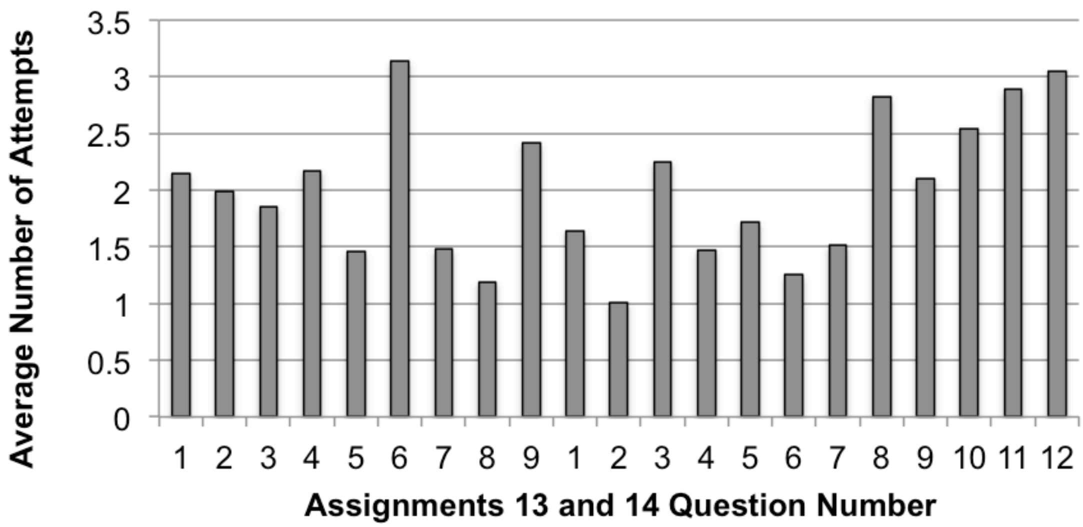

Figure 70. Average number of attempts for questions from assignments 13 and 14.

Appendix H.

Average Confidence and Certainty Judgments for Sapling Learning Questions.

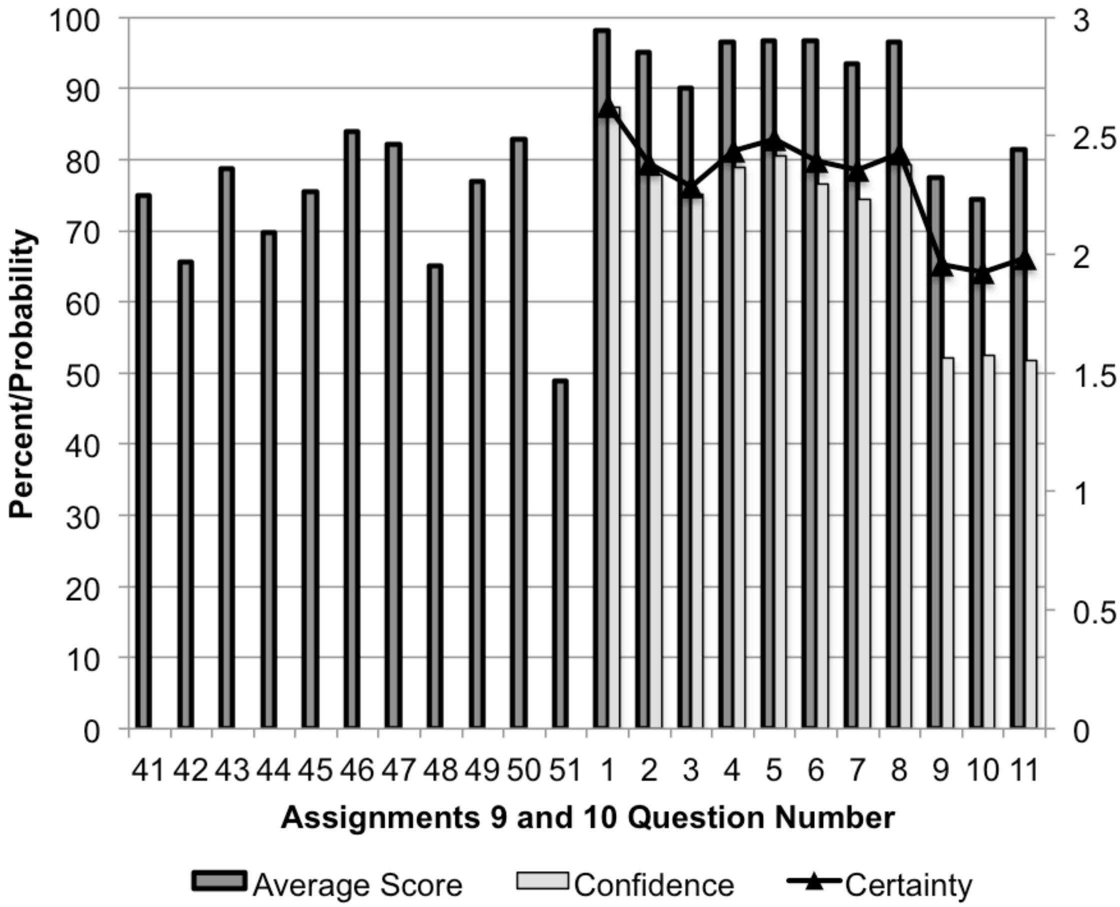

Figure 71. Confidence judgments (bars) and certainty judgments (points) for questions in assignment 10.

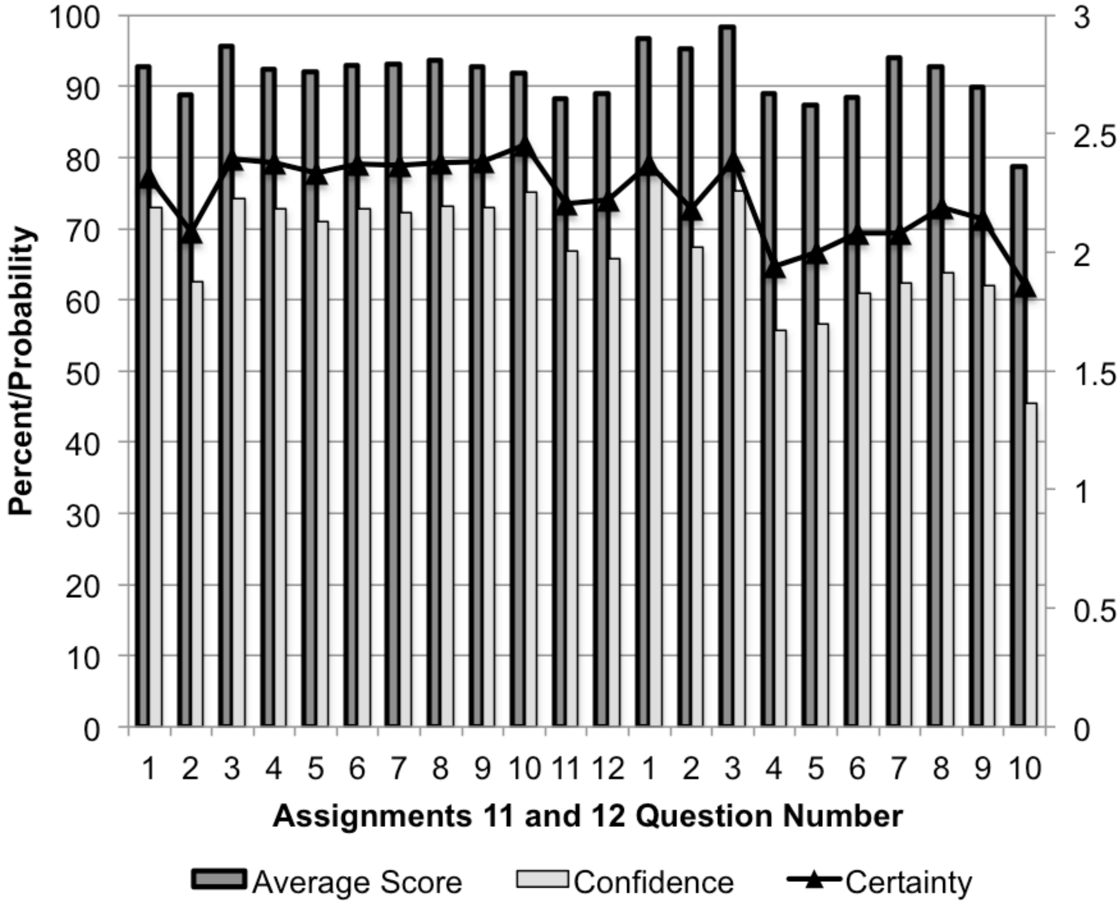

Figure 72. Confidence judgments (bars) and certainty judgments (points) for questions in assignments 11 and 12.

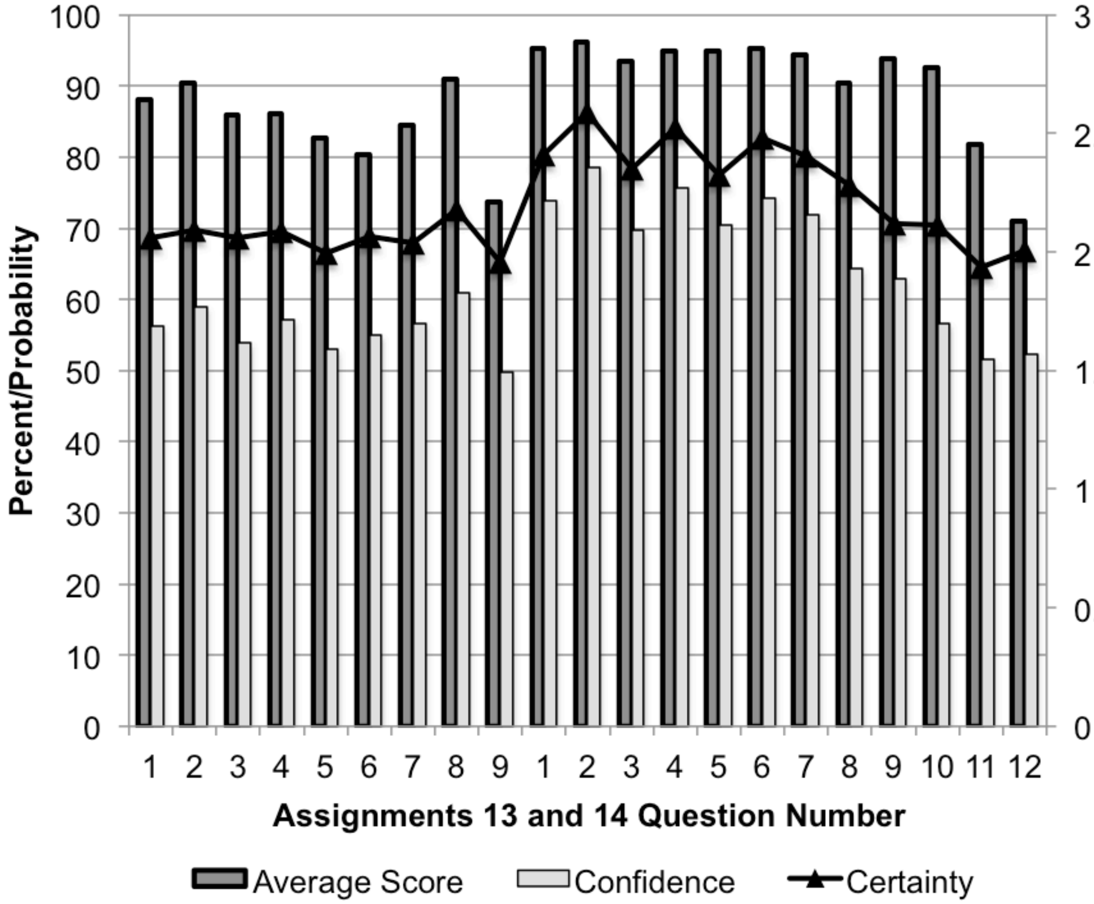

Figure 73. Confidence judgments (bars) and certainty judgments (points) for questions in assignments 13 and 14.

Appendix I.

Analysis of Variance Results for Online Homework Behaviour Profiles

Table 63. Course Examination Scores Grouped by Online Homework Behaviour Cluster

Variable		Sum of Squares	df	Mean Square	F	Sig.
Midterm 1 Percent	Between Groups	12752.622	5	2550.524	14.94	0.000
	Within Groups	191198.101	1120	170.713		
	Total	203950.723	1125			
Midterm 2 Percent	Between Groups	29117.512	5	5823.502	18.399	0.000
	Within Groups	350692.377	1108	316.509		
	Total	379809.89	1113			
Final Exam Percent	Between Groups	25625.401	5	5125.08	18.058	0.000
	Within Groups	318444.254	1122	283.818		
	Total	344069.656	1127			
Weighted Exam	Between Groups	22919.259	5	4583.852	19.866	0.000
	Within Groups	258893.243	1122	230.743		
	Total	281812.501	1127			

Table 64. Sapling Performance Scores Grouped by Online Homework Behaviour Cluster

Variable		Sum of Squares	df	Mean Square	F	Sig.
Sapling Percent	Between Groups	26586.09	5	5317.218	34.098	0.000
	Within Groups	174963.026	1122	155.939		
	Total	201549.116	1127			
Average Item Score	Between Groups	0.861	5	0.172	144.991	0.000
	Within Groups	1.333	1122	0.001		
	Total	2.194	1127			
Transformed Average Item Score	Between Groups	40.646	5	8.129	80.22	0.000
	Within Groups	113.7	1122	0.101		
	Total	154.346	1127			
Transformed Sapling Percent	Between Groups	29.601	5	5.92	40.586	0.000
	Within Groups	163.667	1122	0.146		
	Total	193.268	1127			

Table 65. Confidence, Certainty, and Learning From Errors Grouped by Online Homework Behaviour Cluster

Variable		Sum of Squares	df	Mean Square	F	Sig.
Average Confidence	Between Groups	13669.993	5	2733.999	6.032	0.000
	Within Groups	465902.668	1028	453.213		
	Total	479572.661	1033			
Average Certainty	Between Groups	8.81	5	1.762	7.994	0.000
	Within Groups	226.15	1026	0.22		
	Total	234.961	1031			
Learning From Errors (Correct 2nd attempt)	Between Groups	36975.615	5	7395.123	56.773	0.000
	Within Groups	146149.457	1122	130.258		
	Total	183125.072	1127			